Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Τμήμα Βιολογίας

Σχέσεις δομής-λειτουργίας και εξελικτικές σχέσεις μεταφορέων πουρινών και πυριμιδινών των οικογένειων NCS1 και NCS2

Αιμιλία Κρυπωτού

Διδακτορική Διατριβή

Ιανουάριος 2016

National and Kapodistrian University of Athens

Faculty of Biology

Structure-function and evolutionary relationships in purine/pyrimidine transporters of the NCS1 and NCS2 families

Ph.D. thesis of Emilia Krypotou

January 2016

Advising Committee:

George Diallinas (supervisor)

Stavros Hamodrakas

Efstathios Frillingos

Defence Committee:

George Diallinas

Stavros Hamodrakas

Efstathios Frillingos

Claudio Scazzocchio

Bernadette Byrne

Emmanuel Mikros

Dimitris Hatzinikolaou

To everyone standing by me

during a “bumpy ride”,

Acknowledgements

When you decide to do a Ph.D. you imagine that is everything about hard work and studying, but you could not be more wrong… A Ph.D. is love and passion for science, 100% devotion to something that will follow you everywhere for at least 3-4 years and probably just a handful of your closest people will understand why you spent so many hours in that "stupid" lab!

For me the Ph.D., was an intense experience that included lots of good, bad, happy and stressful moments, but definitely taught me how to do "pure" science without fancy equipment, how to be inventive, how to use critical thinking, how to be patient and persistent at the same time (until you finally get a perfect western blot!!!), how to supervise and most importantly how to be supervised. After 5 years in Diallinas' lab, I am thankful to many people that have stood by my side scientifically or in general, people that we shared good or bad moments and helped me reach my goals.

So, first of all I owe a great "ευχαριστώ" to my supervisor, George, for everything he taught me, for our long conversations about science, art, food and life in general, for his patience to listen to all my ideas (even the stupid ones), for his persistent efforts to make me a better scientist, for his talent to transfer his passion for science to others. Our collaboration was a long journey that has provided me all the tools necessary to advance in the scientific field.

Then, I have to acknowledge the unique opportunity I had to meet and closely collaborate with a unique person; a person with an endless amount of energy and passion for life and science, Claudio Scazzocchio. I am gratefull to him for his collaboration and for all the "4 a.m." emails, expression of his "science-holism".

Next, I have to thank Sotiris Amillis, one of the most patient people I have ever met. He is always (=after his morning coffee) willing to help you in any experiment, and to talk for everything (science, politics, history, and literally everything). It has been a pleasure to collaborate with him, as I learned a lot of thinks working by his side.

During my PhD, I had the chance to collaborate with a lot of post- and undergraduate students working in Diallinas' lab and I am thankful to all of them, as each one has contributed more or less to my work or becoming the person I am today. However, I will specifically refer to few people that we spent countless hours together and collaborated closely. So, by chronological order I would like to thank: Vasso Kosti for her great collaboration during our shared project and for her pleasant company during our morning coffees; Joana Sa Pessoa for meeting and becoming friends with a great Portuguese; Vassilis Yalelis for having the chance to meet an incredible guy and spending countless hours with him discussing science or "philosophy of science"; Katerina Galanopoulou for making my workdays better just by being a great friend; Alexandros Kokotos for standing by my side every step of the way…; Minos Evangelinos for being an endless source of humor during the long work hours; Jacob Bobonis for being at first an impossible student -which made me a better teacher- and then a great collaborator; Iliana Karvela-Kalogeraki for being

supportive and encouraging when it was needed; Gerasimos Anagnostopoulos for being the happiest and most optimist person I have met; Georgia Sioupouli for her great collaboration.

Furthermore, I am grateful to everyone that I collaborated with and contributed to my work, especially members of Mikros' lab and Alexandros Pittis. Finally, I have to specially thank everyone from the Microbiology team for being available and willing to help me or advise me every time I turned to them and for all the fun we had during these years in our small corridor.

Περίληψη Διδακτορικής Διατριβής

Οι πρωτεΐνες των οικογενειών NCS1 και NCS2/NAT ανήκουν στην κατηγορία συμμεταφορέων H^+/Na^+ και συμμετέχουν στην πρόσληψη θρεπτικών συστατικών, όπως πουρίνες, πυριμιδίνες, βιταμίνες και σχετικών μεταβολιτών. Μέλη της οικογένειας NCS1 εντοπίζονται σε βακτήρια, μύκητες και φυτικούς οργανισμούς, ενώ της NCS2/NAT σχεδόν σε όλες τις μορφές ζωής. Σκοπός της συγκεκριμένης εργασίας είναι η δομική και λειτουργική μελέτη μελών των δύο αυτών οικογενειών στον μύκητα *Aspergillus nidulans*, έναν οργανισμό-μοντέλο που έχει χρησιμοποιηθεί ευρέως στη μελέτη μεταφορέων και κυτταρικών μηχανισμών.

Η συγκεκριμένη μελέτη χωρίζεται σε τέσσερα διακριτά κεφάλαια τα οποία αφορούν: (*i*) τη λειτουργική και δομική ανάλυση του κύριου μεταφορέα πουρινών AzgA, (*ii*) το λειτουργικό χαρακτηρισμό των NCS2/NAT πρωτεϊνών του *Aspergillus brasiliensis* στον *A. nidulans,* (*iii*) τη λειτουργική και δομική ανάλυση του μεταφορέα πουρινών/κυτοσίνης FcyB, (*iv*) τη λειτουργική και δομική ανάλυση του μεταφορέα ουρακίλης FurD και τη διερεύνηση της λειτουργίας αγνώστων ομολόγων της Fur οικογένειας. Αναλυτικότερα:

(*i*) Ο AzgA είναι ο κύριος μεταφορέας υποξανθίνης, αδενίνης και γουανίνης στον *A. nidulans* και "ιδρυτικό" μέλος της οικογένειας AzgA με μέλη σε βακτήρια, μύκητες και φυτά. Η οικογένεια AzgA, αν και αποτελεί μία φυλογενετικά διακριτή οικογένεια, προβλέπεται να σχετίζεται δομικά με την οικογένεια NCS2/NAT. Βάση της μελέτης μας αποτέλεσε το θεωρητικό μοντέλο της δομής του AzgA, το οποίο χτίστηκε μέσω μοντελοποίησης χρησιμοποιώντας την κρυσταλλοδομή μιας NCS2/NAT πρωτεΐνης, του μεταφορέα ουρακίλης UraA της *Escherichia coli*. Με βάση το θεωρητικό μοντέλο και υπολογισμούς πρόσδεσης υποστρωμάτων εντοπίσαμε αμινοξέα, τα οποία προβλέπονται να αλληλεπιδρούν με τα φυσιολογικά υποστρώματα του μεταφορέα. Με μια σειρά ορθολογικά σχεδιασμένων αμινοξικών αντικαταστάσεων δημιουργήσαμε στελέχη που εξέφραζαν μεταλλαγμένες μορφές του AzgA τις οποίες αναλύσαμε λειτουργικά. Η λειτουργική ανάλυση των στελεχών αυτών περιελάμβανε δοκιμασίες ανάπτυξης παρουσίας υποστρωμάτων και τοξικών αναλόγων, κινητικές μελέτες μέσω πρόσληψης και ανταγωνιστικής αναστολής της πρόσληψης ραδιοσημασμένου υποστρώματος, εντοπισμό της υποκυτταρικής τοπολογίας σημασμένων με την πράσινη φθορίζουσα πρωτεΐνη (GFP) μεταλλαγμάτων, καθώς και ανάλυση της σταθερότητας των μεταλλαγμένων μορφών του AzgA. Τα αποτελέσματα οδήγησαν στην αναγνώριση αμινοξέων απαραίτητων για τη λειτουργία του AzgA, τα οποία εντοπίζονται κυρίως στα διαμεμβρανικά τμήματα TMS3, TMS8 και TMS10. Συγκεκριμένα, τα αμινοξέα Asn131 (TMS3), Asp339 (TMS8) και Glu394 (TMS10) προβλέπονται να αλληλεπιδρούν άμεσα με τα υποστρώματα, ενώ το Asp342 (TMS8) προτείνεται ότι είναι απαραίτητο στη μεταφορά των υποστρωμάτων. Η ανάλυσή μας επιβεβαιώνει το θεωρητικό δομικό μοντέλο του AzgA και επιπλέον επιβεβαιώνει ότι οι οικογένειες AzgA και NCS2/NAT είναι δομικά συγγενείς.

(*ii*) Τα μέλη της NCS2/NAT οικογένειας στους μύκητες μπορούν να χωριστούν σε δύο συγγενείς κλάδους, οι οποίοι ορίζονται από τους μεταφορείς ξανθίνης-ουρικού οξέος του *A. nidulans*, UapA και UapC. Πρόσφατες φυλογενετικές

αναλύσεις, αναγνώρισαν μία διακριτή ομάδα NCS2/NAT πρωτεϊνών σε περιορισμένο αριθμό μυκήτων, που ονομάστηκαν UapD. Καθώς καμία UapD πρωτεΐνη δεν έχει χαρακτηριστεί λειτουργικά μέχρι τώρα, κλωνοποιήσαμε και εκφράσαμε στον *A. nidulans* ένα γονίδιο του *A. brasiliensis* που κωδικοποιεί μία UapD πρωτεΐνη (AbUapD). Επίσης, εκφράσαμε την UapC πρωτεΐνη του *A. brasiliensis* (AbUapC) στον *A. nidulans*, ως μάρτυρα της ετερόλογης έκφρασης. Η πρωτεΐνη AbUapC αποδείχθηκε, όπως αναμενόταν, ότι δρα ως συμμεταφορέας H^+/ξανθίνης-ουρικού οξέος και εντοπίζεται στην πλασματική μεμβράνη. Αντίθετα, δεν ανιχνεύθηκε καμία λειτουργία του AbUapD σχετική με την μεταφορά πουρινών-πυραμιδινών. Επιπλέον, η πρωτεΐνη AbUapD, σημασμένη με GFP, εντοπίστηκε στο περιπυρηνικό ενδοπλασματικό δίκτυο και σε χυμοτόπια, χαρακτηριστικό πρότυπο ελαττωματικά διπλωμένων πρωτεϊνών. Τέλος, χιμαιρικές πρωτεΐνες UapA/AbUapD επίσης εντοπίζονται σε αντίστοιχες εσωτερικές δομές, υποδεικνύοντας ότι η αλληλουχία του UapD περιλαμβάνει στοιχεία που προωθούν την, *a priori*, πρωτεϊνική αποδόμηση. Το παράδοξο αυτό αποτέλεσμα της εξελικτικής διατήρησης ασταθών *ψευδο-μεταφορέων* σχολιάζεται περαιτέρω στην παρούσα εργασία.

(*iii*) Στους μύκητες η NCS1 οικογένεια περιλαμβάνει μέλη που χωρίζονται σε δύο υποοικογένειες, την Fcy και την Fur. Ο FcyB, το μόνο χαρακτηρισμένο μέλος της Fcy-υποοικογένειας στον *A. nidulans*, είναι μεταφορέας υποξανθίνης-αδενίνης-γουανίνης και κυτοσίνης. Χρησιμοποιώντας το θεωρητικό δομικό μοντέλο του FcyB που χτίστηκε με μοντελοποίηση με βάση την κρυσταλλοδομή του μεταφορέα βενζυλ-υδαντοΐνης Mhp1 και υπολογιστικές δοκιμασίες πρόσδεσης υποστρωμάτων, εντοπίσαμε αμινοξέα που προβλέπονται να αλληλεπιδρούν με τα φυσιολογικά υποστρώματα του FcyB. Με πειραματική προσέγγιση αντίστοιχη με αυτήν που ακολουθήθηκε στην ανάλυση του AzgA, κατασκευάσαμε και χαρακτηρίσαμε λειτουργικά μια σειρά FcyB μεταλλαγμάτων. Αναγνωρίσαμε αμινοξέα στα διαμεμβρανικά τμήματα TMS1, TMS3, TMS6 και TMS8 τα οποία προβλέπονται να συμμετέχουν στην πρόσδεση ή/και μεταφορά των υποστρωμάτων, αλλά και στην εξειδίκευση του FcyB. Συγκεκριμένα, τα Ser85 (TMS1), Trp159, Asn163 (TMS3), Trp259 (TMS6), Asn350, Asn351, Pro353, Asn354 (TMS8) συμμετέχουν άμεσα ή έμμεσα στην πρόσδεση των υποστρωμάτων, ενώ τα Asn163, Asn350, Asn351 και Asn354 καθορίζουν και την εξειδίκευση του FcyB.

(*iv*) Η υποοικογένεια Fur περιλαμβάνει 7 μέλη στον *A. nidulans* (Fur-A, -B, -C, -D, -E, -F, -G), εκ των οποίων τα δύο έχουν χαρακτηριστεί ως μεταφορείς αλλαντοΐνης (FurA) και ουρακίλης (FurD). Απαλοιφή των υπόλοιπων 5 *fur* γονιδίων δεν αποκάλυψε κάποια λειτουργία σχετική με την μεταφορά πουρινών, πυριμιδινών, αλλαντοΐνης, ή άλλων σχετικών μεταβολιτών. Υπερεκφράσαμε τα *fur-B, -C, -E, -F, -G* μέσω ενός ισχυρού υποκινητή συνεχούς έκφρασης, με σκοπό την εύρεση κάποιας σχετικής με αυτά λειτουργίας. Δοκιμασίες ανάπτυξης των στελεχών που εξέφραζαν τις υπερεκφρασμένες Fur πρωτεΐνες, έδειξαν ότι τα FurC, FurF και FurG είναι ειδικά για 5-φθοροουρακίλη, ενώ το FurE για αλλαντοΐνη, ουρικό οξύ, 5-φθοροουρακιλη, 5-φθοροκυτοσίνη και 5-φθοροουριδίνη. Δεν ανιχνεύσαμε κανέναν φαινότυπο που να σχετίζεται με λειτουργία του FurB. Μέσω πειραμάτων πρόσληψης υποστρώματος, ανιχνεύθηκε μετρήσιμη πρόσληψη ραδιοσημασμένης ουρακίλης μόνο για το FurE. Περαιτέρω κινητική ανάλυση του FurE, έδειξε ότι προσδένει με υψηλή συγγένεια το ουρικό οξύ και με χαμηλή συγγένεια την αλλαντοΐνη και την ουρακίλη. Αναλύοντας τα θεωρητικά δομικά μοντέλα των FurA, FurD και FurE που χτίστηκαν με βάση την κρυσταλλοδομή του Mhp1, και με δεδομένα από δοκιμασίες πρόσδεσης υποστρωμάτων, εντοπίσαμε αμινοξέα που πιθανά εμπλέκονται στην πρόσδεση των υποστρωμάτων και τον καθορισμό εξειδίκευσης των Fur. Για να επιβεβαιώσουμε τις

προβλέψεις μας, κατασκευάσαμε και αναλύσαμε λειτουργικά μια σειρά μεταλλαγμάτων στο FurD και το FurE. Η λειτουργική, αυτή, ανάλυση αναγνώρισε αμινοξέα στα διαμεμβρανικά τμήματα TMS1, TMS3, TMS6 και TMS8 που σχετίζονται με τη λειτουργία ή/και την εξειδίκευση του FurD ή του FurE. Συγκεκριμένα, τα Asn54 (TMS1), Trp130 (TMS3), Asn341 (TMS8) δείχθηκαν απαραίτητα για τη λειτουργία του FurD, ενώ, τα Asn249 και Tyr250 (TMS6) δείχθηκε ότι καθορίζουν την εξειδίκευση του FurD. Στο FurE, τα Pro251 και Lys252, που αντιστοιχούν με τα Asn249 και Tyr250 του FurD, δείχθηκαν ότι καθορίζουν την εξειδίκευση, επιβεβαιώνοντας τη σημασία του TMS6 στην εξειδίκευση των μεταφορέων Fur. Τέλος, εντοπίσαμε ένα αμινοξύ στο TMS10 (Met389), το οποίο αν και δεν είναι μέρος του κέντρου πρόσδεσης υποστρωμάτων, καθορίζει το προφίλ εξειδίκευσης του FurD. Το εύρημα αυτό συνάδει με πρόσφατα δεδομένα που προτείνουν ότι οι μεταφορείς εμπεριέχουν δυναμικά στοιχεία που λειτουργούν ως ελεγχόμενες πύλες εισόδου και εξόδου των υποστρωμάτων.

Ph.D. Thesis Summary

The NCS1 and NCS2/NAT families include proteins characterized as H^+/Na^+ symporters specific for the uptake of nutrients, such as purines, pyrimidines, vitamins and related metabolites. The NCS1 family includes members in bacteria, fungi and plants, whereas NCS2/NAT members are present in almost all domains of life. The current study concerns the structural and functional analysis of proteins belonging in NCS1 and NCS2 families in the fungus *Aspergillus nidulans*, a model-organism widely used in studies related to transporters and cellular mechanisms.

This study is divided in four distinct chapters: (*i*) functional and structural analysis of the major purine transporter AzgA, (*ii*) functional characterization of *Aspergillus brasiliensis* NCS2/NAT proteins expressed in *A. nidulans*, (*iii*) functional and structural analysis of the purine/cytosine transporter FcyB, (*iv*) functional and structural analysis of the uracil transporter FurD and investigation of "orphan" Furs function. In more detail:

(*i*) AzgA is the major hypoxanthine, adenine and guanine transporter in *A. nidulans* and the founding member of the AzgA family, including members in bacteria, fungi and plants. The AzgA family, although it forms a phylogenetically distinct group, is predicted to be structurally related to the NCS2/NAT family. In fact, our study was based on a theoretical model of the AzgA structure, built through homology modelling on the crystal structure of a NCS2/NAT protein, namely the *Escherichia coli* uracil transporter UraA. Based on the AzgA theoretical model and substrate docking calculations, we identified amino acid residues predicted to interact with the substrates. We performed a series of rationally designed amino acid substitutions and functionally analyzed the strains expressing the mutant versions of AzgA. The functional analysis of the strains included growth testing in the presence of substrates and toxic analogues, kinetic studies through radiolabeled substrate transport measurements and competitive inhibition assays, detection of the subcellular topology of GFP-tagged versions of AzgA, and protein stability analysis of mutant proteins. Our data identify residues critical for AzgA function, localized in transmembrane segments TMS3, TMS8 and TMS10. Particularly, residues Asn131 (TMS3), Asp339 (TMS8) and Glu394 (TMS10) are predicted to interact directly with substrates, while Asp342 (TMS8) is necessary for substrate transport. Our analysis confirmed the theoretical AzgA model and in addition suggests that AzgA and NCS2/NAT families are structurally related.

(*ii*) Fungal NCS2/NAT members can be classified in two very closely related clades, those exemplified by the UapA and UapC uric acid-xanthine transporters of *A. nidulans*. Recent phylogenetic analysis however, identified a clearly distinct NCS2/ NAT group, present only in a limited number of fungi, which we called UapD. As none UapD member has been functionally characterized before, we cloned and expressed in *A. nidulans* a gene encoding a UapD protein from *A. brasiliensis* (AbUapD). As a control, we also expressed the *A. brasiliensis* UapC protein (AbUapC) in *A. nidulans*. AbUapC proved, as expected, a H^+/xanthine-uric acid

symporter, localized in the plasma membrane. In contrast, no function related to purine-pyrimidine transport was detected for AbUapD. Importantly, AbUapD tagged with GFP localized in the perinuclear endoplasmic reticulum and vacuoles, a characteristic picture related to partially misfolded proteins. In addition, chimeric UapA/AbUapD proteins were, also, localized in similar internal structures, suggesting that the UapD sequence includes 'elements' that trigger, *a priori*, protein turnover. This unusual phenomenon of an apparently unstable *pseudo-transporter*, which is however evolutionary conserved, is discussed further in the present Thesis.

(*iii*) Fungal NCS1 proteins are divided in two subfamilies, Fcy and Fur. FcyB, the only characterized member of the Fcy-subfamily in *A. nidulans*, is a hypoxanthine-adenine-guanine-cytosine transporter. Based on the FcyB theoretical model, built through homology modeling on the crystal structure of the *Microbacterium liquefaciens* benzyl-hydantoin transporter Mhp1, and substrate docking calculations, we identified residues predicted to interact with substrates. Following experimental approaches similar to those described for AzgA, we constructed and functionally characterized a series of FcyB mutants. We identified residues in transmembrane segments TMS1, TMS3, TMS6 and TMS8, which are predicted to participate in transport function and specificity of FcyB. Specifically, we obtained evidence that residues Ser85 (TMS1), Trp159, Asn163 (TMS3), Trp259 (TMS6), Asn350, Asn351, Pro353, Asn354 (TMS8) participate directly or indirectly in substrate binding, while Asn163, Asn350, Asn351 and Asn354 are critical for FcyB substrate specificity.

(*iv*) The Fur subfamily includes 7 members in *A. nidulans* (Fur-A, -B, -C, -D, -E, -F, -G), out of which two have been characterized as allantoin (FurA) or uracil (FurD) transporters. Knock-out of the rest 5 *fur* genes did not reveal any function related to transport of purines, pyrimidines, allantoin or relative compounds. We over-expressed *fur-B, -C, -E, -F, -G* via a strong constitutive promoter, in order to detect their putative function. Growth test analysis showed that FurC, FurF and FurG are specific for 5-fluorouracil, whereas FurE is specific for allantoin, uric acid, 5-fluorouracil, 5-fluorocytosine and 5-fluorouridine. We did not detect any phenotype related to FurB function. Using direct uptake assays, measurable radiolabeled uracil uptake was detected only in case of FurE. Further kinetic analysis of FurE showed it has relatively high affinity for uric acid, but low affinity for allantoin and uracil. Analysis of the theoretical structural models of FurD, FurA and FurE built, based on the crystal structure of Mhp1, and the docking assays performed, in an effort to identify the structural basis for their differences in substrate specificity, we identified residues predicted to participate in substrate binding and also affecting the specificity of Furs. To validate our predictions, we constructed and analyzed a series of relative FurD or FurE mutants. Our functional analysis confirmed that residues in TMS1, TMS3, TMS6 and TMS8 are related to the function or/and specificity of FurD or FurE. Specifically, in FurD, residues Asn54 (TMS1), Trp130 (TMS3), Asn341 (TMS8) were shown critical for function, whereas Asn249 and Tyr250 (TMS6) determine substrate specificity. In FurE, residues Pro251 and Lys252, corresponding to Asn249 and Tyr250 of FurD, were shown to determine substrate selection,

confirming the significance of TMS6 in Fur specificity. Importantly, we also identified a residue in TMS10 (Met389), which determined the specificity profile of FurD, even though it is not an element of the major substrate binding site. This finding is in agreement with recent data supporting the novel concept stating that transporters include dynamic channel-like gates, which control substrate selection.

Table of Contents

1 Introduction

1.1 Transport proteins: transporters *vs.* channels

The purpose of this Ph.D. thesis is the study of purine/pyrimidine transporters of the model fungus *Aspergillus nidulans* in respect to their structural, transport and specificity properties. An extensive delineation concerning transport proteins, as well as their structure and transport mechanism follows, in order to introduce the reader to the subject and to outline the state of the art in the field of transporters.

Various nutrient molecules, ions and water are necessary for the development and proper function of the cells. Membranes are totally impermeable to ions and polar molecules, surrounding like a natural barrier all cells and cellular organelles. The only molecules permeating the plasma membrane, albeit at low rates, are lipid-soluble molecules such as steroids, O_2, CO_2 and some very small polar molecules such as water, ethanol and glycerol. Specific transmembrane proteins have evolved in order to mediate the movement of molecules and ions necessary for the cells (Garrett and Grisham, 2012). Two great categories of transmembrane transport proteins exist: **transporters** (also known as permeases or carriers) and **channels**. Both categories share several features, such as specificity or selectivity, compact secondary structure built mostly by several α-helices and elements that regulate their topology and function, but also have major differences in their mechanism of function. Transport systems may be passive or active. Passive transport does not require direct energy expenditure and utilizes the concentration gradients of substrates. Active transport requires energy coupling that can be an ATP hydrolysis or ion movement in order to transport a compound against its concentration gradient. All transport processes are summarized in **Figure 1**.

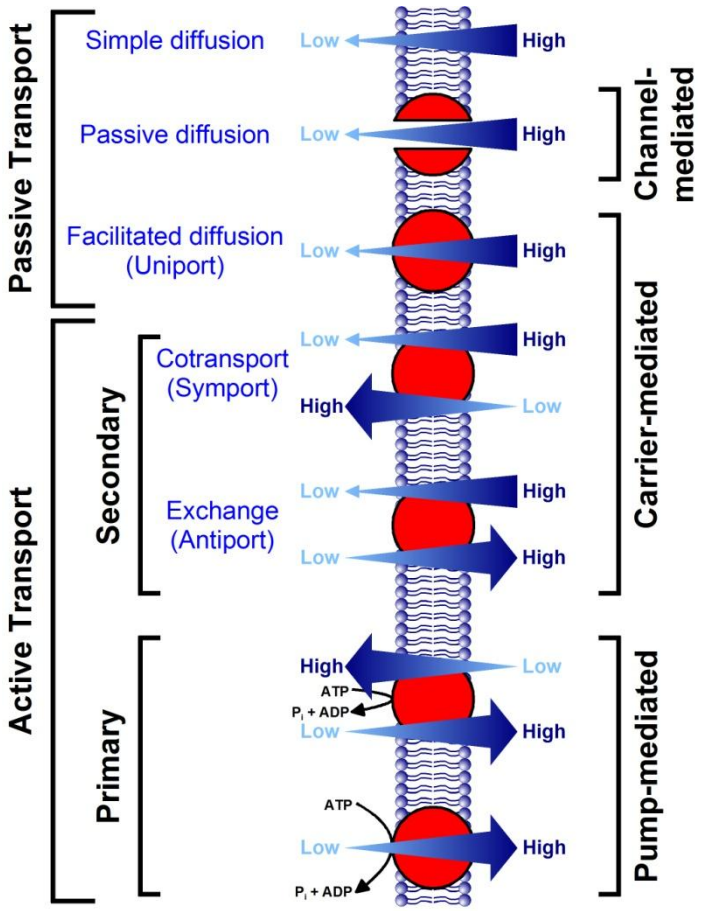


Figure 1. Summary of membrane transport processes. All kinds of transport processes and transport proteins are shown. Channels mediate passive diffusion, while transporters (carriers or pumps) mediate facilitated diffusion and active transport. Active transport can be divided in primary transport that is coupled to binding and hydrolysis of ATP and secondary transport that is coupled to ion movement on the same (symport) or opposite (antiport) direction. Blue arrows indicate the concentration gradient of the transported molecule. The Figure is adapted by *http://www.physiologyweb.com/*.

Channels are proteins that function as monomers or homo-oligomers and facilitate the passive diffusion of ions or small molecules (e.g. water or acetylcholine) downstream their concentration (from a high concentration to a lower) or down an electrochemical concentration; thus little energetic interaction between the channel and the transported molecule is required. The passive translocation of ions in the pore leads to an ion flux approximating a rate of 10^6-10^7 ions/sec, one of the major characteristics that a transporter lacks. Channels are not "holes" in the membrane but are highly selective for their substrates; their selectivity is achieved by the existence of filters (*selectivity filters*), which are composed of residues, specific for the molecules transported through the channel (Moran et al. 2015, and references therein). A constantly accessible channel would lead to massive uncontrolled movements of ions and other molecules with catastrophic consequences for the cell. Specific *gating* elements exist most often at both sides of the membrane, which control the accessibility to the selectivity filters and thus to the channel pore (**Figure 2**). Both gates can be open simultaneously, which contrasts the situation in transporters, where gates open alternatively (Gadsby, 2009; and references therein). The external signals

responsible for the change between an open and a closed state of the gates include changes in transmembrane voltage, binding of ligands or mechanical stress. Recently, cases of ion channels have been reported that also possess an additional "hydrophobic gating" that is tuneable to local changes in the diameter and/or hydrophilicity of the channel pore due to liquid-vapour transitions of water within the pore (Aryal et al. 2015, and references therein).

Transporters are proteins that can mediate substrate transport downstream its concentration, called *facilitators*, or active transport of molecules upstream their concentration (from low concentration to higher) and thus their function is energy-coupled, known as *active transporters*. Based on the kind of transport-coupled energy, an active transporter can be categorized as a *primary* or a *secondary active transporter*. Facilitators (also known as uniporters) transport a substrate downstream its electrochemical gradient without a coupled energy release, but contrasting with channels, they bind their substrate and undergo conformational changes in order to transport it (**Figure 2**). Primary active transport is succeeded by direct coupling of the transported molecule to the binding and hydrolysis of ATP, while in secondary active transport the movement of one molecule against its electrochemical gradient is coupled to the movement of another molecule (usually ion) down its electrochemical gradient. This mechanism works for molecules moving the same direction (*symport*) or the opposite (*antiport*). The mechanism underlying transporter function seems to be more complex compared with that of a channel and requires larger conformational changes from an open to a closed (occluded) state. A plethora of genetic data from our laboratory and recent structural evidence from other groups has shown that transporters also have gates. It seems that two gates at both sides of the membrane are necessary in order to *sequentially* control the access and release of substrates to and from a major, centrally located, substrate binding site. This mechanism results in a much slower transport rate of about 10^2-10^5 molecules/sec compared with the "current"-like transport rate of channels (Dubyak 2004; and references therein). Up to recently, specificity in transporters was generally thought to depend solely on the topological fitness and the strength of interactions of substrates with specific residues, when present in a major, freely accessible, binding pocket (le Coutre and Kaback, 2000). Recently however, this long lasting dogma has been challenged, mostly from our laboratory, with the discovery of mutations in transporter gating domains that could also lead to major specificity modifications (Diallinas, 2014). This novel concept underlies a large part of the work reported in the current Ph.D. thesis.

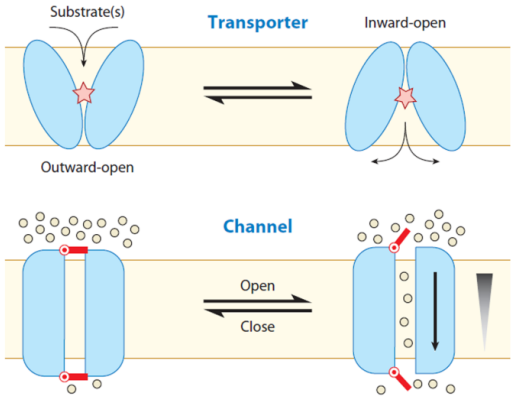


Figure 2. Transporter *vs.* channel. Schematic representation of a transporter and a channel, highlighting their major differences. A transporter undergoes conformational changes in order to transport a substrate from one side of the membrane to the other, while a channel includes a pore that is controlled by gates (red-colored), which when they are open, molecules can move like a current through the pore from both sides depending on their concentration gradient. The Figure is adapted from (Shi, 2013).

Transporters and channels are involved in a large number of physiological processes of organisms: nutrition, homeostasis, signalling, neurotransmission, detoxification, cell migration and polarity. It is not, thus, surprising that their malfunction is associated with many diseases: cystic fibrosis, neurodegeneration, diabetes, stroke, metastasis in cancer etc. In addition, transporters are related to phenomena of drug or antibiotic resistance, both as gateways of efflux proteins. Over the years many drugs have been used to target transport proteins (e.g. drug efflux proteins in bacteria and cancer cells), most of them selected through massive screenings. On the other hand, no effort has been made to rationally design or screen highly specific drugs based on knowledge coming from the comparative study of transporters that could target transport proteins implicated in diseases or are present in pathogens (Diallinas, 2014).

1.2 Transport mechanism: the alternating access model

The predominant model concerning the function mechanism of transporters is the *alternating access model*. The first description of this kind of mechanism was by Jardetzky (1966), who suggested an allosteric rearrangement of the transporting protein that mediates the translocation of a molecule towards a specific direction, and not by ordinary diffusion. He also predicted the existence of a cavity that is alternatively accessible from either side of the membrane, acquiring different binding and release affinities depending on the conformational state. This idea was born to describe the Na^+/K^+ transporter function mechanism, but has been generalised to other transport processes that constitute the alternating access model.

A general description of this model includes distinct conformational stages that a transporter acquires in order to perform a complete transport cycle. This

hypothesis predicts that the transporter is accessible to the periplasmic side of the membrane (outward open) in order to upload the substrate, passes through an intermediate state inaccessible or closed at either side (occluded) and then the transporter releases the substrate in the cytoplasmic side of the membrane (inward open). The inverse cycle is needed for the transporter to return to its original outward open conformation. Thus at least three distinct states exist that include large conformational changes of the transporter.

Based on the protein structures and biochemical data available, three distinct mechanisms have been described that fall within the alternating model:

(i) the rocker-switch mechanism (**Figure 3a**): The substrate binding site of the protein, located in the middle, is exposed on one side of the membrane (e.g. the periplasm), the substrate binds in the binding site and then with a hinge or rocker-switch like movement around the region of the binding site the protein opens towards the other side of the membrane (e.g. the cytoplasm) while simultaneously closing the previously open side of the protein. This mechanism is mostly proposed for the MFS (Major Facilitator Superfamily) family based on the crystal structures available in different conformations for 15 MFS proteins (see below). Although the rocker-switch mechanism in its strict definition proposes only two conformations, outward-open and inward-open, recent data predict the existence of an intermediate occluded conformation, closed in both sides, that is acquired while in the transition of the two alternate open conformations (see also later).

(ii) the gated pore mechanism (**Figure 3b**): The substrate binding site is located in the middle, similarly to the previous mechanism, but the access to the binding site is controlled by two gates, one outer (periplasmic side) and one inner (cytoplasmic side). The outer gate opens and one (in case of antiporters) or two (in case of symporters) substrates bind in the binding site cavity. Then, the outer gate closes triggering the opening of the inner gate, in order the substrate or substrates to be released in the cytoplasm. An occluded intermediate conformation has been captured in several crystal structures facing either outward or inward. The opening of the two gates occurs alternately and not at the same time, one must close in order for the other to open. The molecular mechanism that underlies the gating process seems to vary among proteins and has not yet been studied in detail. The movement of flexible helices is predicted to trigger the rigid movement of a bundle of helices and their coordination regulates the opening and the closure of the gates. Members of different families with similar or relatively similar folds are predicted to use this transport mechanism, such as members of the APC (Amino acid-Polyamine-Cation), NCS1 (Nucleobase Cation Symporter 1), and NAT/NCS2 (Nucleobase Ascorbate Transporter/Nucleobase Cation Symporter 2) families (see below).

(iii) the modified "moving-carrier" mechanism (**Figure 3c**): The "moving-carrier" model was first proposed by Widdas (1952) to describe the placental glucose transfer in sheep, and describes the transport process as a moving protein that carries glucose from one side of the membrane to the other. Later on, this mechanism was extended to constitute the alternating access mechanism. The crystal structures of an aspartate:Na^+ symporter, Glt_{Ph}, of the EAAT (Excitatory Amino Acid Transporter)

family from *Pyrococcus horikoshii* (Boudker et al., 2007; Yernool et al., 2004) predict a transport mechanism similar to what Widdas had described. In particular, contrasting with the first two mechanisms described herein, where substrate remains during the whole transport process at the same cavity where it was originally bound, in the mechanism predicted for Glt_{Ph} the substrate binds onto a protein domain that moves along with the substrate during the transport cycle. The moving (transport) domain retains the same conformation and slides along the rest protein in order to obtain/release the substrate at either side of the membrane. Similarly to the gated mechanism, two gates regulate the opening of the binding cavity and occluded positions of the "moving" protein are intermediate between the open conformations.

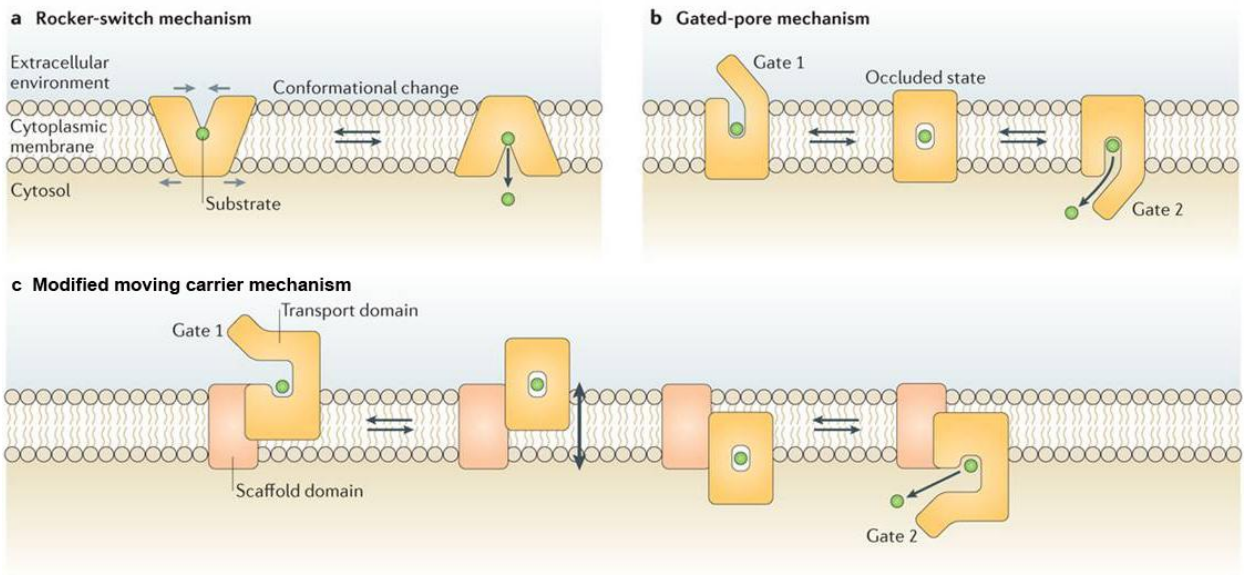


Figure 3. Diagrammatic representation of three distinct models of the alternating access mechanism of transport. (a) The rocker-switch mechanism: The substrate binds in a cavity within the protein. Opening of the protein towards the cytoplasm occurs with the simultaneous closure towards the periplasm with a hinge-like movement. **(b)** The gated pore mechanism: Opening of the outer gate allows the binding of the substrate, closure of the outer gate creates an occluded intermediate with the binding site inaccessible from both sides. The occluded state is followed by the opening of the inner gate and the release of the substrate to the cytoplasm. **(c)** The modified moving-carrier mechanism: The substrate binds on the transport domain when the outer gate is open. Closure of the outer gate is followed by movement of the transport domain in respect to the scaffold domain. The cycle is completed with the opening of the inner gate and release of the substrate in the cytoplasm. The models depict transport of a substrate from the periplasm towards the cytoplasm; the same mechanism applies *vice versa.* The Figure is adapted from (Slotboom, 2014).

The mechanisms described above are deduced from extant experimental data, but it does not follow that all transporters function through one of these mechanisms. On the contrary, some proteins might express characteristics of more than one of the above mechanisms, or even function through a completely different transport cycle.

1.3 Secondary active transporters

This category includes proteins that catalyze transport of a substrate upstream its concentration gradient by coupling the transport process with an energy force other than ATP or similar compound hydrolysis, as mentioned earlier. This subclass includes 123 protein families according to *www.tcdb.org* (Saier et al., 2014); these families are categorized in superfamilies by the same classification system. The two largest superfamilies are: the Major Facilitator Superfamily (MFS) and the Amino acid-Polyamine-organoCation (APC) superfamily.

1.3.1 The Major Facilitator Superfamily (MFS)

The MFS is the largest known superfamily of secondary transporters found in the biosphere ubiquitously in all organisms. According to *www.tcdb.org*, the MFS superfamily includes eight families among which are: the "canonical" MFS family, from which the superfamily was named after, the Glycoside-Pentoside-Hexuronide:Cation symporter (GPH), ATP:ADP Antiporter (AAA) family, with the MFS family the one with the most characterized members (Saier et al., 2014).

1.3.1.1 The MFS family

The MFS is an evolutionarily old, large and diverse family present in all domains of life, consists of 82 subfamilies, 17 of them without any functionally characterized members. It includes more than 10,000 sequenced members, specific for: sugars, polyols, drugs, neurotransmitters, Krebs cycle metabolites, amino acid peptides, osmolites, siderophores, iron-siderophores, nucleosides, organic and inorganic anions (*www.tcdb.org*). Most proteins consist of 400-600 amino acid residues which form 12, 14 or 24 putative transmembrane α-helical segments (*www.tcdb.org*). All known MFS transporters appear to function as monomers, although they can be isolated as homo-oligomers. The functional role of oligomerization, if any, remains unknown although some evidence exists that in eukaryotes it might be related to subcellular trafficking and turnover (Martzoukou et al., 2015).

The structures of several MFS transporters have been elucidated as shown in **Table 1**. It is apparent that the last 3 years a boom of crystal structures has occurred, in particular 47 different structures of the MFS are available in *www.rcsb.org* on July 2015, and many of them concern different conformations of the same protein.

The core MFS fold, as shown by the crystal structures, contains 12 TMSs, which are organized into N- and C-domains that consist of 6 TMSs each. The two domains are related to each other by a pseudo-twofold symmetry axis that is perpendicular to the membrane. Each domain comprises a pair of inverted structural repeats, TMSs 1-3 and TMSs 4-6 in the N-domain, and TMSs 7-9 and TMSs 10-12 in the C-domain. The two repeats within each domain are intertwined to form two tightly

packed three-helix bundles. Occasionally, some MFS transporters may contain extra TMSs additional to the core MFS fold that participate in stabilizing the overall structure. Studies concerning the origin of the MFS topology support that the probable origin of the 6-TMS unit involved triplication of a 2-TMS hairpin structure (Reddy et al., 2012). Although the mechanism of transport of the MFS proteins has been proposed to be the rocker-switch model, intermediate occluded structures have also been obtained, casting some doubt on this issue (see also below).

Table 1. Available crystal structures of MFS proteins as in July 2015.

Transporter	Substrate	Organism	Reference
LacY	lactose:H^+ symporter	*E. coli*	(Abramson et al., 2003; Guan et al., 2007)
GlpT	glycerol-3-phosphate: phosphate antiporter	*E. coli*	(Huang et al., 2003)
EmrD	multidrug:H^+ antiporter	*E. coli*	(Yin et al., 2006)
FucP	L-fucose:H^+ symporter	*E. coli*	(Dang et al., 2010)
$PepT_{So}$	Peptide:H^+ symporter	*Shewanella oneidensis*	(Newstead et al., 2011)
$PepT_{Sh}$	Peptide:H^+ symporter	*Streptococcus thermophilus*	(Solcan et al., 2012)
YbgH	Peptide:H^+ symporter	*E. coli*	(Zhao et al., 2014)
XylE	D-xylose:H^+ symporter	*E. coli*	(Quistgaard et al., 2013; Sun et al., 2012)
Glut1	glucose facilitator	*Homo sapiens*	(Deng et al., 2014)
NRT1.1	dual affinity nitrate transporter	*Arabidopsis thaliana*	(Sun et al., 2014)
MelB	melibiose/Na^+ symporter	*Salmonella typhimurium*	(Ethayathulla et al., 2014)
NarK	nitrate/nitrite exchanger	*E. coli*	(Zheng et al., 2013) (Fukuda et al., 2015)
NarU	nitrate transporter	*E. coli*	(Yan et al., 2013)
PiPT	phosphate transporter	*Piriformospora indica*	(Pedersen et al., 2013)
	glucose/H^+ symporter	*Staphylococcus epidermidis*	(Iancu et al., 2013)

1.3.1.1.1 The case of LacY

The lactose permease of *E. coli*, LacY, catalyzes the symport of galactoside:H^+ and is the most extensively studied member of the MFS (Smirnova et al., 2011). All 417 amino-acyl side chains have been mutated and functional analysis revealed that only about ten side chains are irreplaceable for transport function. Even though studies on

LacY had started in the 70s the first crystal structure was released three decades later (Abramson et al., 2003). Today, there are six different crystals for LacY of wild type and mutant proteins, caught in three different conformations: inward open (Abramson et al., 2003; Guan et al., 2007; Mirza et al., 2006), inward partially occluded (Chaptal et al., 2011) and outward almost occluded (Kumar et al., 2014).

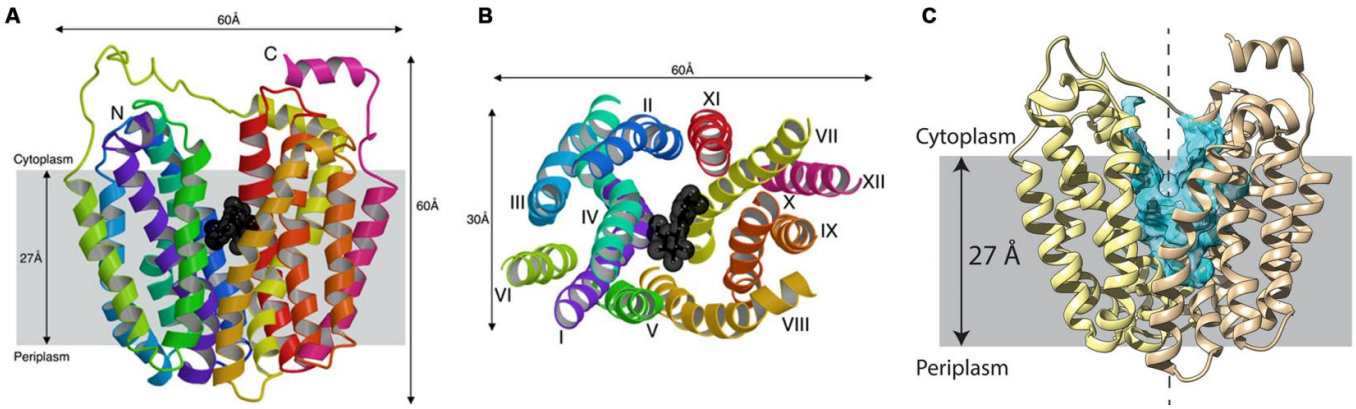


Figure 4. The structure of LacY in inward open conformation bound with substrate. (A) Side view of LacY. **(B)** Bottom view of LacY from the cytoplasmic side. **(C)** N- (yellow) and C-terminal (tan) bundles of LacY. The bound substrate (TDG: b-D-galactopyranosyl 1-thio-b, D-galactopyranoside) is represented with black spheres. The grey shaded area represents the membrane. The Figure is adapted from (Kaback, 2015; Kaback et al., 2011).

The structure of LacY, similar to other MFS transporters, consists of twelve TMSs mostly α-helical, organized into two pseudo-symmetrical six-helix bundles, the N- and the C-terminal bundles (**Figure 4**). The sugar and H^+ binding sites are located almost in the middle of the protein at the apex of the hydrophilic cavity. The residues participating in the sugar binding are mostly located in the N-terminal bundle, while the ones for the H^+ binding in the C-terminal (Smirnova et al., 2011). The first crystal structures released for LacY in an inward facing conformation and the structures available from other MFS proteins in an outward facing conformation were used in computational experiments to predict the outward facing conformation, as well as, the transition from the one to the other through an occluded state (Stelzl et al., 2014). A partially occluded state was also captured in the latest released crystal structure of LacY (Chaptal et al., 2011; Kumar et al., 2014) indicating that an occluded intermediate state is present in the transport mechanism of LacY. Interestingly, the existence of flexible gates similar to the ones proposed for the gated model of transport has been predicted computationally, challenging the strict rocker-switch model that involves only rigid body movements (Stelzl et al., 2014).

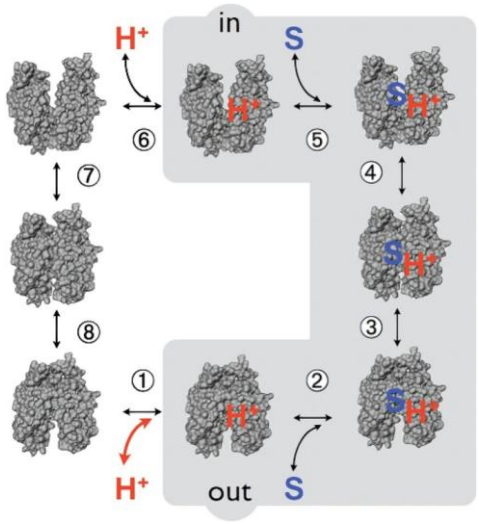


Figure 5. The transport cycle of LacY. The outward facing protein is protonated (1). The substrate binds in the binding site (2). Conformational changes lead to an intermediate occluded state (3) and then an inward open conformation (4). Substrate dissociates towards the cytoplasm (5) and then LacY deprotonates (6). Conformational changes lead to the initial state of the protein facing the periplasmic side of the membrane (7-8). The Figure is adapted from (Kaback, 2015).

Based on biochemical and structural data, a six-step kinetic model has been proposed for LacY (**Figure 5**): (1) protonation of LacY, (2) substrate binding, (3-4) global conformational change resulting in an inward open conformation bound with the H^+ and the substrate, (5) dissociation of the substrate, (6) deprotonation of LacY, (7-8) a conformational change allowing substrate and proton binding sites to become accessible to the outer side of the membrane (Madej et al. 2012, Kaback 2015; and references therein). During these steps a periplasmic/cytoplasmic pathway opens with simultaneous closing of the inward/outward-facing cavity resulting to a sugar-binding side alternatively accessible to either side of the membrane (rocker-switch model of transport). Biochemical studies have shown that the closing of the cytoplasmic cavity in LacY is reciprocally related to opening of the periplasmic cavity (Smirnova et al., 2014). Observations have led to the prediction of a chemiosmotic coupling of lactose/H^+ symport (Kaback, 2015). Symport in the absence or presence of an electrochemical gradient of H^+ is the same overall reaction. It appears that sugar binding and dissociation, and not the electrochemical gradient of H^+, are the driving force for alternating access. This model resembles that of an enzyme, as an induced fit state causes the transition to an occluded intermediate (Kaback, 2015). The necessary energy is actually released by the dissociation of the sugar, while the H^+ electrochemical gradient controls the rate of the transport process (Kaback, 2015).

1.3.2 The amino acid-polyamine-organoCation (APC) superfamily

The APC superfamily is the second largest superfamily of secondary carriers currently known. Members of the APC superfamily are present in all domains of life. Currently 13 families are categorized in this superfamily, with members specific for

nucleobases, amino acids, metal ions, purines, pyrimidines, vitamins, peptides etc (Vastermark et al., 2014). Some of the families included in this superfamily are: the *sensu stricto* Amino acid-Polyamine-Organocation Family (APC), the Nucleobase:Cation Symporter family 1 (NCS1), the Nucleobase Ascorbate Transporter/Nucleobase: Cation Symporter family 2 (NAT/NCS2), the Sulphate Permease family (SulP) and the Neurotransmitter:Sodium Symporter family (NSS). This superfamily includes a large number of crystal structures from members belonging to different families. Most members exhibit a common 5+5 topology, but two families (SulP and NCS2) have a 7+7 topology (Vastermark et al., 2014).

1.3.2.1 The NSS family

The NSS (Neurotransmitter:Sodium Symporter) family has members present in all domains of life and are specific for neurotransmitters, amino acids (glycine and GABA), osmolytes, monoamines (serotonin, dopamine and norepinephrine) and related nitrogenous substances (*www.tcdb.org*). Most proteins support Na^+ and in some cases Cl^- or K^+ symport. NSS members participate in neurotransmission and their malfunction is associated with neurological diseases. NSS transporters are targets for psychostimulants, anti-depressants and other drugs. One of the most important members of the family is LeuT, a bacterial amino acid symporter with several crystallographic and biochemical data available. LeuT is in fact the paradigmatic model of the APC superfamily with several crystal structures available, which not only established a common fold for the APC members, but also provided important details on the mechanism of substrate binding and transport.

1.3.2.1.1 The case of LeuT

LeuT is a Na^+-coupled, non-polar amino acid symporter from the bacterium *Aquifex aeolicus*. Although, it was named after leucine (Leu), LeuT is specific for several aliphatic/aromatic amino acids, including Gly, Ala, Met and Tyr, and can also bind, but not transport, Trp (Singh and Pal 2015; and references therein). The first crystal structure of LeuT (Yamashita et al., 2005) revealed that the protein comprises 12 mostly α-helical TMSs. The core domain of TMS1-10 is organized into two inverted structural repeats, each containing 5 consecutive TMSs. The two repeats are related by an antiparallel pseudo twofold in the membrane plane (**Figure 6**). The first TMSs of each repeat (TMS1 and 6) consist of two short α-helical segments connected with an unwound segment in the middle. The binding sites of the two Na^+ ions and the amino acid substrate are located between TMSs 1, 3, 6 and 8.

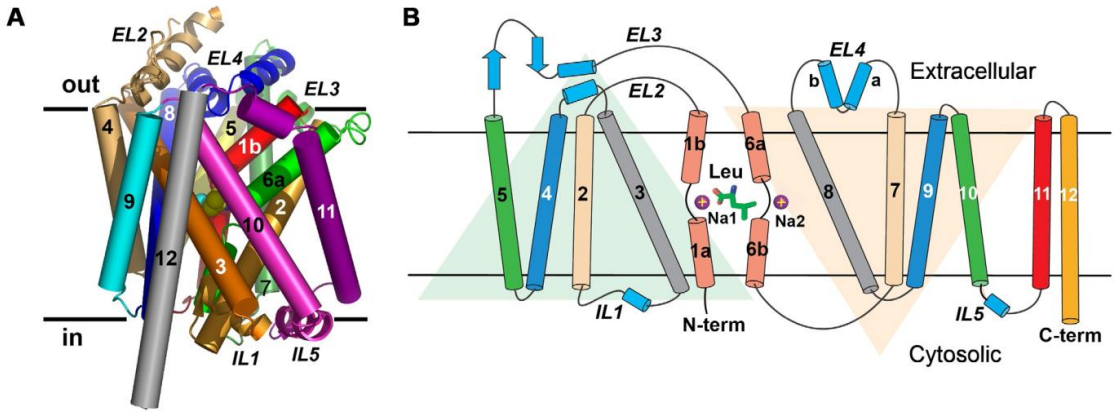


Figure 6. The LeuT crystal structure. (A) Side view of LeuT structure in an outward occluded conformation bound with substrate (Leu). Transmembrane segments (TMSs) are depicted with colored cylinders. **(B)** LeuT topology representation. Leu and two sodium ions (purple circles) are shown in the binding site. Faint triangles illustrate the component helices of the 5+5 inverted repeat, each pair of which is shown in the same colour. The membrane boundaries are shown with black lines. The Figure is adapted from (Singh, 2014; Singh and Pal, 2015).

Several crystal structures in different conformational states are available for LeuT: outward occluded (Piscitelli and Gouaux, 2012; Yamashita et al., 2005), outward occluded with an inhibitor bound on the external gate (Zhou et al., 2007), outward open trapped with Trp (Singh et al., 2008), outward open and inward open (Krishnamurthy and Gouaux, 2012). The LeuT fold has been also observed for members of the APC family (AdiC, ApcT and GadC), the betaine/carnitine/choline transporter family (BCCT) (BetP and CaiT), the NCS1 family (Mhp1) and the solute-sodium symporter family (SSS) (vSGLT) (Shi 2013; and references therein).

The transport mechanism predicted by the available structures of LeuT (**Figure 7**) and LeuT-fold proteins is the gated pore model of the alternating access mechanism. Initially the transporter is open to the periplasm with its binding site accessible, successive binding of the substrates from the outer gate towards the main binding site is predicted (Celik et al., 2008; Shi et al., 2008). Crystal structures of a mutant that can transport Trp (a competitive inhibitor for LeuT) (Piscitelli and Gouaux, 2012) revealed that the transition from the outward open to the occluded state is feasible only when the ligand is properly bound in the binding cavity. The transition involves movement of TMS1b and 6a towards TMS3 and 10 (**Figure 7A** and **B**), which is inhibited by the "wrong" substrate due to steric clash effects; e.g. LeuT bound with a Trp is locked in the outward open state (**Figure 7F**) (Singh et al., 2008). The conformational transition from the outward-occluded to the inward-open state is achieved by multiple adjustments in individual TMSs, including bending of TMS2 and 7 and independent movements of TMS1a and 6b (**Figure 7B-D**) (Krishnamurthy and Gouaux, 2012). Our knowledge of the transport mechanism has been perfected by the acquisition of an inward open structure for LeuT (Krishnamurthy and Gouaux, 2012). In particular, changes in the substrate/ion binding sites are predicted to be coupled with large structural changes that simultaneously open the inner gate and close the outer one. This is indicated by the

cascade of TMS movements, initiated by TMS1a and 5, which cause movement of TMS7 and 1b and finally close the outer gate. This sequence of events, implies that the two gates are linked, ensuring that one gate cannot open unless the other is closed.

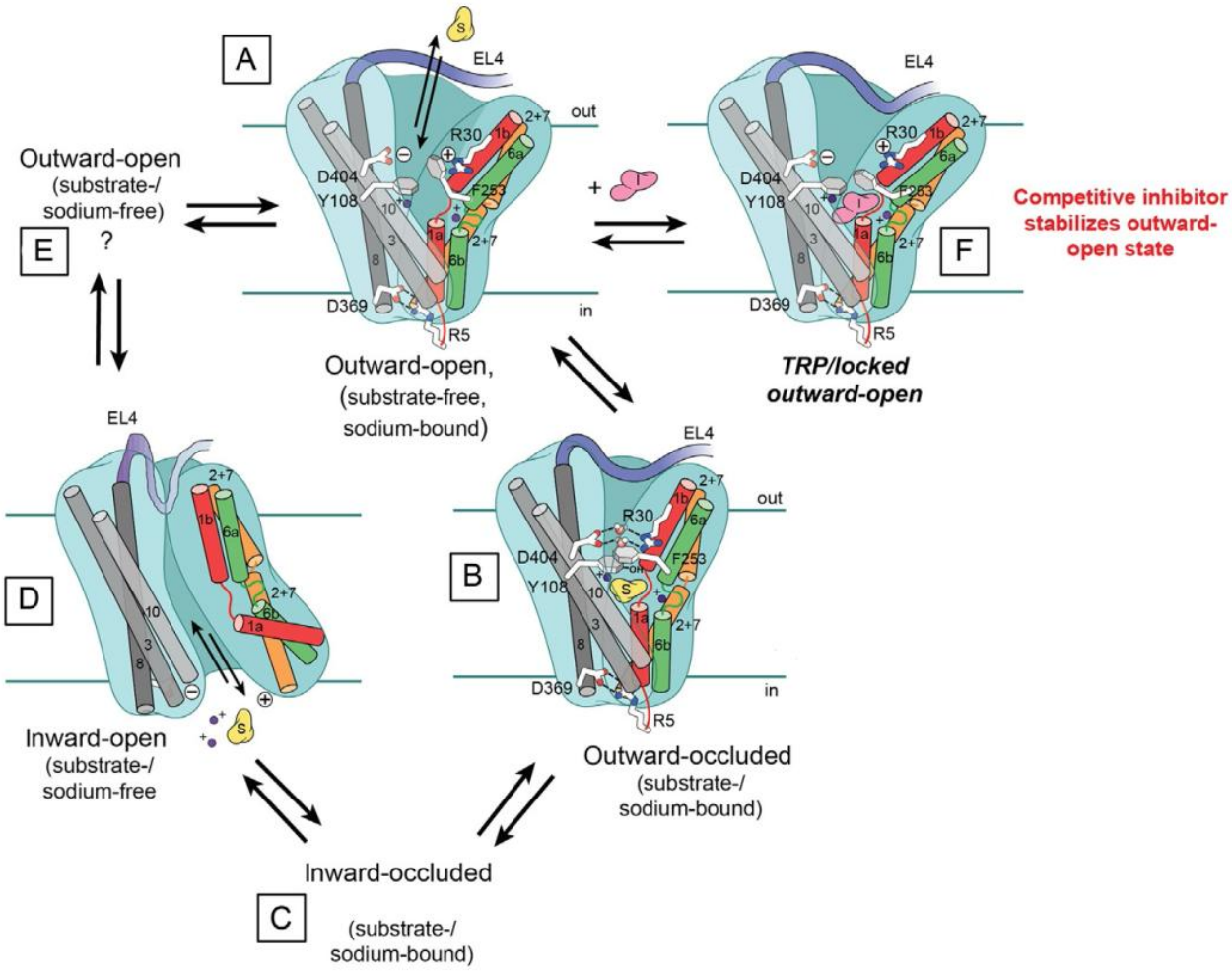


Figure 7. Transport mechanism of LeuT based on crystal structures and kinetic data. TMSs participating in the conformational changes are shown, while TMS5 is not shown for clarity. Amino acid residues participating in substrate binding are also shown. **(A)** Substrate binds in the ion bound outward open conformation. **(B)** Movement of TMSs 1b and 6a lead to the outward occluded conformation. An inward occluded state **(C)** mediates the opening of the inner gate **(D)**, which is mediated by movement of TMSs 1a, 2, 6b and 7. The substrate and ions are released in the cytoplasm and the protein returns to an outward open conformation **(E)** becoming accessible to the periplasm. The structure of the locked outward open conformation bound with a competitive inhibitor (Trp) is also shown **(F)**. The Figure is adapted from (Singh and Pal, 2015).

1.3.2.2 The NCS1 family

The NCS1 (Nucleobase Cation Symporter 1) family includes members from archaea, bacteria, fungi and plants and are specific for the uptake of purines, pyrimidines, vitamins and related metabolites (*www.tcdb.org*). NCS1 proteins are usually 419-635 amino acyl residues long and possess twelve putative α-helical TMSs (*www.tcdb.org*). Several members of the family have been functionally characterized, including eukaryotic members from *Saccharomyces cerevisiae*, *A. nidulans*, *Candida albicans*, *A. fumigatus* and *Arabidopsis thaliana*. The only crystal structure available from this family is the benzyl-hydantoin transporter Mhp1 from *Microbacterium liquefaciens*.

1.3.2.2.1 The case of Mhp1

Mhp1 is a Na^+ symporter specific for 5-aryl substituted hydantoins from the bacterium *M. liquefaciens* (Simmons et al., 2014). Mhp1 is the only member of the NCS1 family with a crystal structure solved (Weyand et al., 2008) and the only one captured throughout a complete transport cycle: outward open (Weyand et al., 2008), outward occluded (Weyand et al., 2008) and inward open (Shimamura et al., 2010), thus providing important clues on the mechanism underlying transport catalysis.

Mhp1 is structurally similar to LeuT and other members of the APC superfamily, such as vSGLT, even though they share a rather low identity on the protein sequence level (15-16%) (Weyand et al., 2008). The structure, shown in **Figure 8**, consists of 12 TMSs mostly α-helical. The first ten constitute the conserved motif of the LeuT-fold, where TMSs 1-5 and 6-10 are related by a pseudo two-fold axis along the plane of the membrane. The α-helical structure of TMS 1 and 6 appear to be disrupted by an unwound segment in the middle, similarly to LeuT. The structure could be divided in two domains, the *four-helix-bundle* (TMSs 1, 2, 6 and 7) and the *hash motif* (TMSs 3, 4, 8 and 9) (Shimamura et al., 2010; Weyand et al., 2008). TMS 5 and two short helices link the two motifs, while TMS10 links the hash motif with the C-terminal TMSs 11 and 12. According to (Shimamura et al., 2010) TMSs 5 and 10 appear to participate in significant movements during transition from one state to the other, and thus were named the flexible helices. The substrate-binding site is located between TMSs 1, 3, 6 and 8. The occluded conformation (Weyand et al., 2008) is bound with a substrate that lies in between two Trp residues (Trp117 in TMS3 and Trp220 in TMS6) and is further stabilized through H-bonding by residues Gln121 (TMS3) and Asn318 (TMS6).

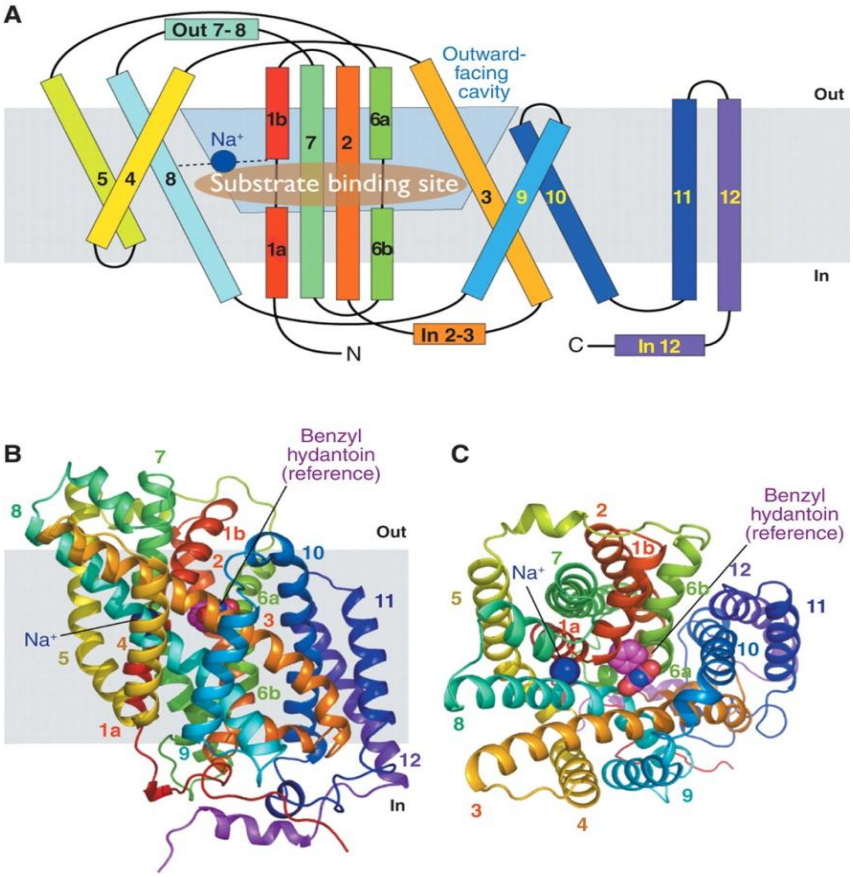


Figure 8. The structure of Mhp1. (A) Mhp1 topology depicted by colored cylinders. The substrate and Na^+ binding site are highlighted in brown and blue respectively. The outward facing cavity is blue-shaded. Mhp1 is shown in an outward open conformation from side view **(B)** and top view **(C)** from the periplasmic side. Benzyl hydantoin is shown in magenta color. The Figure is adapted from (Weyand et al., 2008).

The transport model followed by Mhp1, shown in **Figure 9**, is the gated pore model of the alternating mechanism. During the transition from the outward open to the outward occluded conformation, the N-terminal half of TMS10 appears to bend over the binding cavity and the substrate (**Figure 9** arrow A) (Simmons et al., 2014; Weyand et al., 2011). At the same time, the hash motif shifts towards the bundle through a rigid body rotation of 30^o occurring around an axis in line with TMS3 (**Figure 9** arrow B) (Weyand et al. 2011; and references therein). This movement results into rendering the binding cavity inaccessible from the outside and accessible from the inside. An additional movement of TMS5 allows further opening of the cavity towards the interior of the cell and a small extracellular helix moves and seals completely the outer surface of the protein. Using a system of thin and thick gates where TMSs 5 and 10 act as the thin gates, the substrate-binding cavity is sealed alternatively from the inside and outside (Krishnamurthy et al., 2009). The role of the thick gate that controls the switching from the two extreme conformations is played by the hash motif that moves in relation to the bundle. The movement of the hash motif and the bundle is similar to the rocking bundle proposed for LeuT (Forrest et al., 2011).

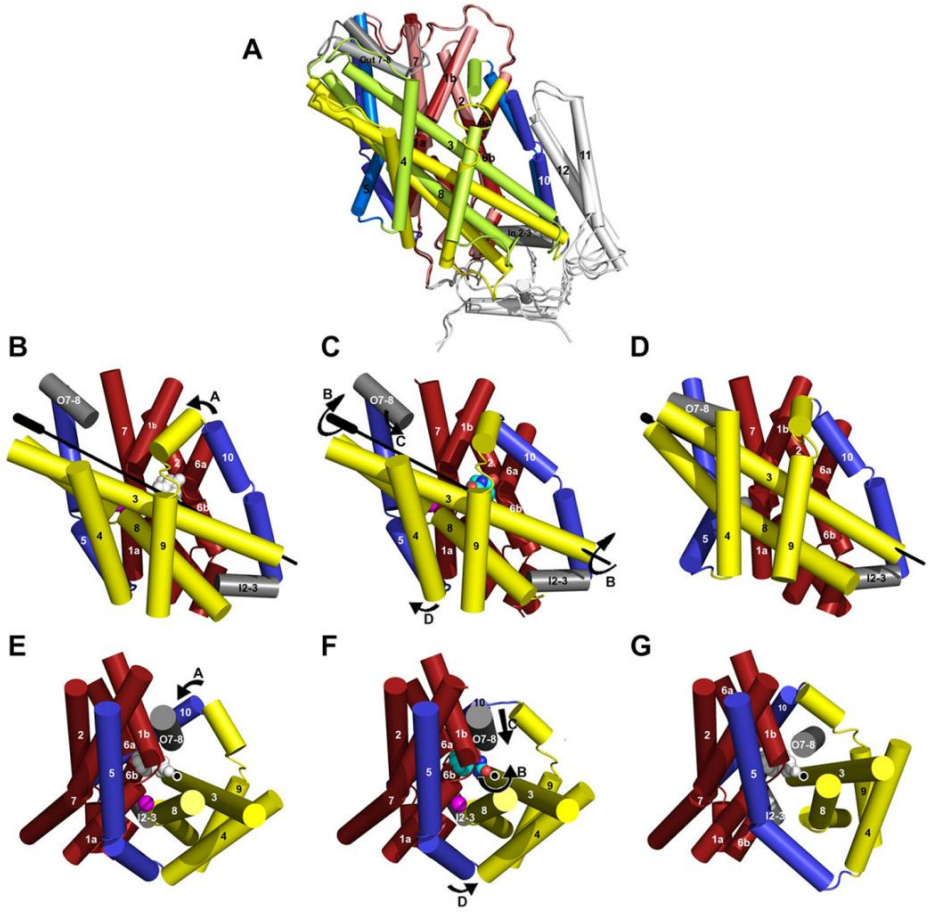


Figure 9. Transport mechanism of Mhp1. TMSs are represented in colored cylinders. Bundle motif (TMSs 1, 2, 6, 7) is in red color. Hash motif (TMSs 3, 4, 8, 9) is in yellow. The flexible helices TMSs 5 and 10 are in blue color. **(A)** Superposition of the two extreme conformations of Mhp1, for the outward facing (colored as mentioned above) and inward facing (colored in salmon for the bundle, light green for the hash motif and light blue for the flexible helices). Side views of outward open **(B)**, occluded **(C)** and inward open **(D)** conformations are shown, TMSs11-12 are omitted for clarity. Top views of the respective conformations are shown in **(E)**, **(F)** and **(G)** respectively. The TMS10 movement from the outward open to the occluded conformation is denoted with arrow A. The hash motif rotation in respect to the bundle motif during transition from outward to inward conformation is denoted with arrow B. The Figure is adapted from (Shimamura et al., 2010).

Very recent studies on Mhp1 revealed an additional role of TMS10 in substrate binding (Simmons et al., 2014). Transport of a bulky hydantoin analogue, L-5-(2-naphthylmethyl) hydantoin (NMH), could not be carried out because of steric clash of a Leu in TMS10 with the substrate, inhibiting the transition from the open state to the occluded. Substitution of the specific amino acid (Leu) with an Ala allowed the formation of the occluded state and as a result the transport of the bulky substrate. Thus, it seems that TMS10 affects the substrate specificity. The importance of this finding will be further discussed in relation to results presented in this Thesis later.

1.3.2.3 The NAT/NCS2 family

The NAT/NCS2 (Nucleobase Ascorbate Transporter/Nucleobase Cation Symporter 2) family includes members from all domains of life. The NAT family was defined by the cloning and characterization of the UapA and UapC uric acid-xanthine transporters of *A. nidulans* and the UraA and PyrP uracil transporters of *E. coli* and *Bacillus subtilis* respectively (Gournas et al. 2008, Diallinas and Gournas 2008; for reviews and references therein). The NAT proteins are H^+ or Na^+ symporters and specific for purines, pyrimidines, purine analogues including well established analogues used in pharmacological practise (e.g. allopurinol and oxypurinol) or ascorbic acid (Diallinas and Gournas, 2008; Faaland et al., 1998; Gournas et al., 2008; Tsukaguchi et al., 1999). This family was recently included in the APC superfamily (Wong et al., 2012). Members of the NAT family were initially proposed to have 12 TMSs, but the crystal structure of a bacterial NAT protein, the uracil transporter UraA from *E. coli*, revealed that the structure comprises 14 TMSs (Lu et al., 2011).

1.3.2.3.1 The case of UraA

UraA is a H^+/uracil symporter from *E. coli* and the only member of the NAT family with a crystal structure (Lu et al., 2011). Although the NAT family is a member of the APC superfamily, it appears to differ from LeuT, Mhp1 and the other APC members with similar folds. The UraA structure consists of 14 TMSs, mostly α-helical (**Figure 10**). One distinct difference from the LeuT-fold and similar folds is the existence of two short antiparallel β-strands in TMS3 and TMS10 breaking the continuity of their α-helical structure. The 14 TMSs are arranged into two structural repeats, TMS1-7 and TMS8-14 and are related to each other by a rotation of ~180^o around an axis parallel to the membrane. The structure can be divided in two apparent domains, the core (TMS1-4 and 8-11) and the gate (TMS5-7 and 12-14) domain. The binding site of uracil is located in between TMSs 1, 3, 8 and 10.

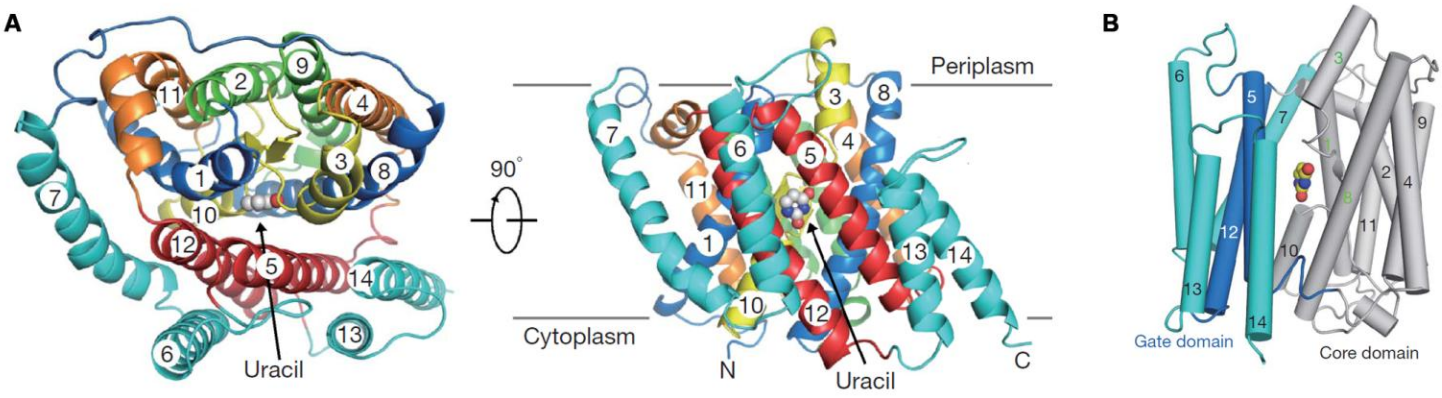


Figure 10. The UraA crystal structure. (A) Top and side views of the overall structure of UraA. Bound uracil is shown. **(B)** TMSs are represented with cylinders. The two domains gate and core are highlighted in blue and grey color respectively. The Figure is adapted from (Lu et al., 2011).

The captured conformation is in an inward open state and the absence of structures of similar fold limits our knowledge concerning the transport mechanism. Based on computational methods, Lu et al. 2011 predict that during the outward conformation, protonation of specific Glu residues in the binding site and binding of the uracil lead to movement of the gate domain in relation to the core domain until closure of the outer surface of the cavity, with simultaneous opening of the inner surface of the cavity (**Figure 11**). Dissociation of the H^+ is predicted to result in dissociation of uracil in the cytoplasm. Based on this model, a role of substrate specificity is given to the core domain, while transport is succeeded via movements of the gate domain.

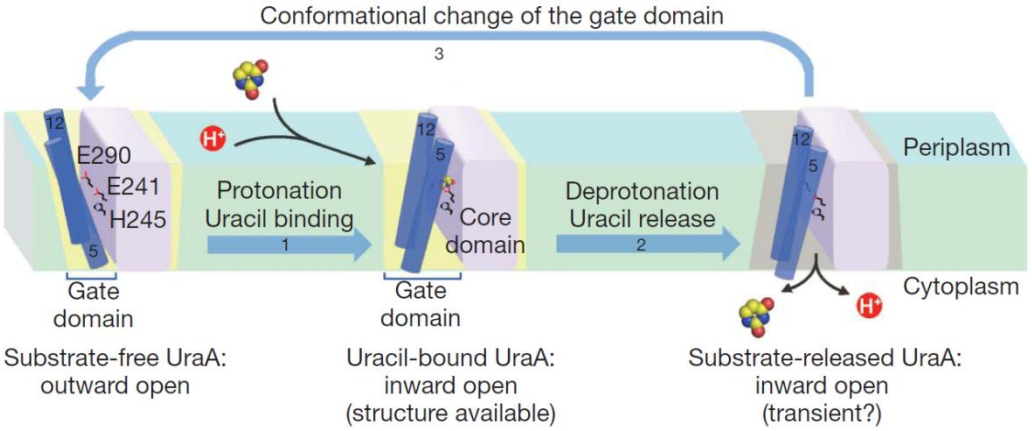


Figure 11. Putative transport mechanism of UraA. Movements of the gate domain in respect to the core domain are shown during the conformational changes of the protein. Two Glu residues (Glu241 and Glu290) are shown, which are predicted to participate in H^+ binding. The Figure is adapted from (Lu et al., 2011).

1.3.3 Common structural characteristics and evolutionary relationships among transporters

Even though transporters belonging to different families differ in their primary structure and their specificity, they share several common structural features and occasionally they even display a similar fold. Most protein structures comprise two or more structural repeats (**Figure 12**) (e.g. two 3 α-helical TMSs in LacY-fold, two 5 α-helical TMSs in LeuT fold). Many hypotheses have been proposed for the existence of a common ancestor-protein that consisted of few TMSs and after gene duplications and rearrangements led to the variety of structures available today (Keller et al., 2014; Västermark and Saier, 2014). Experimental data are absent to support these hypotheses, but computational analysis methods have been used to compare, mix and match and finally "recreate" proteins. For example, in the case of two MFS transporters, rearrangements of structural motifs consisting of 3-helices in the fucose transporter FucP from *E. coli* led to a protein with high homology to LacY, contrasting the weak similarity they share in linear sequence alignment (Madej et al., 2013). These observations are consistent with the hypothesis that MFS might have

evolved through duplications of an initial triple-helix bundle (Madej et al., 2013). Structural, functional and sequence diversity could be due to multiple events of mutations and multiple independent triple-helix bundle fusions during time (Madej et al., 2013).

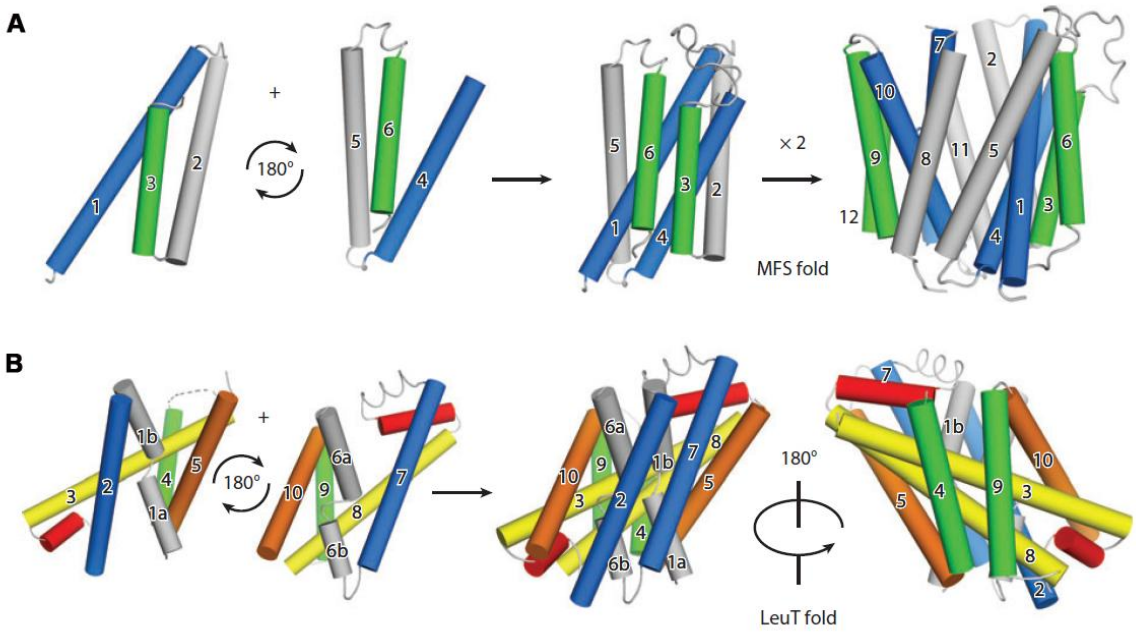


Figure 12. Structural repeats of MFS- and LeuT-fold. (A) The MFS-fold includes four structural repeats, each repeat comprises three TMSs. **(B)** The LeuT-fold includes two structural repeats, each repeat comprises five TMSs. The Figure is adapted from (Shi, 2013).

Members of the APC superfamily acquiring the LeuT fold comprise of two 5 α-helical bundle repeats (5+5), while UraA comprises of two 7 α-helical bundle repeats (7+7). Although these two folds seem rather different and probably evolutionary distant, extensive analysis and structural alignments reveal that they share common characteristics, and the 7+7 architecture might be an extension by duplications of the 5+5 one (Västermark and Saier, 2014). The 7+7 architecture was shown to contain a common repeat unit and the two extra TMSs might have occurred by a hairpin loop duplication (Västermark and Saier, 2014). Similarly to the observations made for the MFS family, an original 5-TMS unit might have existed in the common ancestral protein. Duplication of this unit, as well as additional hairpin loop duplications on either side of the 10 TMS unit could have recurred independently many times resulting in topological diversity, as seen in the variable structures today (Västermark and Saier, 2014).

Another common feature arising from the repeat-based structures is the presence of a pseudosymmetry in the protein (**Figure 12**) (Shi, 2013). The pseudosymmetry is usually twofold and the symmetry axis can be parallel or perpendicular to the membrane. When the symmetry axis is parallel, a subunit/domain is the asymmetric unit and the two asymmetric units are placed against each other through a flat interface to form an intact transporter. But, in the case of a

perpendicular to the membrane axis, the asymmetric units are inverted repeats that are intertwined through extensive interfaces to form a subdomain or a folded protein (Shi, 2013). The ubiquitous presence of pseudosymmetry in proteins (also seen in channels) could be due to functional selection during evolution and also for folding and insertion in the membrane (Shi, 2013).

Another common structural feature is that the binding sites of the ions and the substrate in many cases are formed by TMSs that contain unwound regions (e.g. LeuT and Mhp1) or with a different structure (e.g. UraA) or linking two antiparallel half helices within the membrane (e.g. Glt_{Ph}). This feature could be related with the properties of the amino acid residues needed to interact and coordinate the ion binding (Shi 2013; references therein).

All common structural characteristics shared by the transporters could be related to common features of transport mechanism. Even though, the transport mechanism might have adjusted to the individual structural features of each protein, the same basic properties underlie it.

1.4 Protein turnover regulated by conformational changes

Protein degradation is crucial for a cell. Non-functional proteins must be degraded in order to prevent cellular toxicity and preserve the proper cellular function (Goldberg, 2003). Transporters are retrieved from the plasma membrane through endocytosis; they are internalized, targeted with ubiquitin and finally end up in the vacuole or lysosome for degradation (for a review see MacGurn et al. 2012). Transporter endocytosis has been extensively studied in fungi (*S. cerevisiae* and *A. nidulans*) and mammalian transporters participating in neurotransmission (Diallinas 2014; and references therein).

Even though the mechanism of endocytosis has been studied extensively the initial signal triggering the process of transporter degradation was unclear. Studies on UapA from *A. nidulans* (Gournas et al., 2010; Karachaliou et al., 2013), the amino acid permeases Gap1 and Can1 (Cain and Kaiser, 2011; Ghaddar et al., 2014; Van Zeebroeck et al., 2014) and the uracil permease Fur4 from *S. cerevisiae* (Blondel et al., 2004; Keener and Babst, 2013; Lauwers et al., 2010) have revealed that transporter endocytosis can be elicited by the presence of a substrate (*substrate induced endocytosis*) or a general signal affecting multiple cellular mechanisms (e.g. the abundance of ammonium affects nitrogen metabolism and thus nitrogen related transporters, Apostolaki et al., 2009; Gournas et al., 2010; Springael and André, 1998; Susarla and Robinson, 2008; Valdez-Taubas et al., 2004). Work on UapA, Gap1 and Can1 has showed that substrate transport is critical for endocytosis; binding of substrates that are not transported or mutant versions of the transporter that cannot perform transport function do not undergo endocytosis (Cain and Kaiser, 2011; Ghaddar et al., 2014; Gournas et al., 2010; Van Zeebroeck et al., 2014). Interestingly, mutant versions of UapA and Can1 with novel specificities can undergo endocytosis

in the presence of their new substrates (Ghaddar et al., 2014; Gournas et al., 2010). All data concord in that binding of a substrate in the binding site is necessary but not sufficient to trigger endocytosis. The substrate has to be transported, or at least promote an occluded conformation of the transporter, in order to elicit the endocytosis-relevant conformation state. In agreement with this hypothesis is also a study on Fur4 from Keener and Babst (2013) that showed that substrate binding modifies the interactions of the N-tail region with specific cytoplasmic loops, which in turn affects endocytosis.

The current model predicts that transporter endocytosis is regulated by the conformational state of the protein (Diallinas, 2014; Ghaddar et al., 2014; Gournas et al., 2010; Keener and Babst, 2013; Van Zeebroeck et al., 2014). The relation of endocytosis with transport function and not solely substrate binding indicates that the intermediate occluded conformation might be the most probable to induce endocytosis. Recognition of a specific conformation implies that the protein remains in this state long enough to be recognized by the endocytic machinery (Diallinas, 2014). It is apparent that the structure of a transporter, as well as, the conformational changes it undertakes are critical not only for the function of the protein, but also for interaction with other proteins and regulation of cellular mechanisms.

1.5 Gating mechanism and specificity in transporters

Each transporter is substrate specific and its transport function occurs only when the appropriate ligand is bound. Each ligand binds at a specific binding site and then transport catalysis occurs. The exact mechanism of how specificity of a transporter is achieved or why ligands with similar structure can bind, but are not transported, or why compounds similar or smaller than the physiological substrates are not recognized is quite unclear.

An emerging theory is that transport cannot be performed if a substrate/ligand inhibits one of the alternative conformations described above. Particularly, in the case of LeuT, a crystal structure was obtained with Trp bound in an open outward conformation (Singh et al., 2008). Trp is a competitive inhibitor of LeuT that binds with great affinity but cannot be transported. In the presence of all amino acids transported by LeuT the crystals obtained were in an occluded outward conformation, but only Trp resulted in outward open conformation (Singh et al., 2008). Additionally, in the open conformation a second Trp molecule was bound ~4 Å above the first Trp, closer to the extracellular vestibule (Singh et al., 2008), suggesting a secondary binding site of lower affinity.

The work on the human glucose facilitator (GLUT1) by R. Naftalin and co-workers (Cunningham and Naftalin, 2013; Cunningham et al., 2006; Naftalin, 2010) suggests the existence of multiple binding sites. In more detail, computational analyses predict that glucose might bind in multiple binding sites with different affinities, sliding in the carrier cavity until it is released in the cytoplasm or the

periplasm, using in fact a channel-like mechanism. Yet, this mechanism is not supported by experimental data.

A similar to the LeuT-Trp bound structure was found for Mhp1, when the crystal structure of Mhp1 complexed with a bulky hydantoin analogue (NMH) was captured in the open outward conformation (Simmons et al., 2014). Comparing different conformations of Mhp1, rigid body movements occur after substrate binding and TMS10 moves until covering the binding cavity facing the periplasm and leading from the outward open to the occluded conformation. Compared with other analogues that could be transported and were captured bound in occluded conformations, NMH was a competitive inhibitor that was predicted to clash with a Leu residue of TMS10 inhibiting closure. Substitution of the specific Leu with an Ala led to transport of NMH (Simmons et al., 2014). Another instance of failure to transport because of inability to reach a closed conformation was observed for a double mutant of LacY, where replacement of two Gly residues with Trp at the termini of TMS2 and TMS8 led to abolishment of transport, but retained substrate binding (Smirnova et al., 2013). A crystal structure of this mutant revealed an almost occluded structure, which is partially open to the periplasmic side and tightly sealed to the cytoplasmic side (Kumar et al., 2014). The Trp substitutions are predicted to inhibit closure and thus the transporter cannot switch between conformations and complete a transport cycle (Kumar et al., 2014).

Computational and structural studies mentioned above, but also on other transporters, e.g. the dopamine transporter from *Drosophila melanogaster* (Wang et al., 2015) or the human serotonin transporter SERT (Gabrielsen et al., 2012), all indicate that a necessary step in order transport to be successful is the closure of the extracellular gate and occlusion of the binding site from both sites. A model of transport mechanism with outer and inner gates is shown in **Figure 13**. The question arising is whether gates or neighbouring residues have, in addition to their gating mechanism, properties similar to the selectivity filters of channels. Are these gates able to recognize their physiological substrates among all other compounds? And if yes, is this correlated with the existence of a secondary binding site?

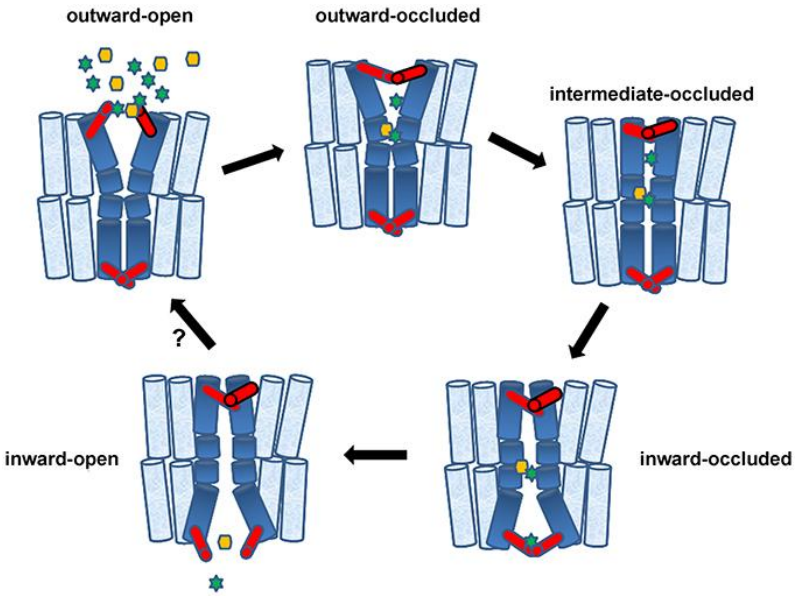


Figure 13. A model of transporter conformational changes that includes gates. Five conformational states of a symporter are shown: outward open, outward occluded, intermediate occluded, inward occluded and inward open. A similar model can be applied to antiporters and facilitators. Inner and outer gates are shown in red. The Figure is adapted from (Diallinas, 2014).

1.6 Study of purine/pyrimidine transporters of the model fungus *A. nidulans*

As mentioned above, the subject of the current Ph.D. Thesis is the study of purine/pyrimidine transporters of the model fungus *A. nidulans*. This fungus has been extensively used in our lab in order to study the gene regulation and expression, the structure-function relationships and the trafficking of a model protein, the uric acid permease UapA. Our lab has also contributed in the study of transporters with the characterization of several other transport proteins. In the current thesis, the proteins of interest belong to the NCS1 and NAT/NCS2 families. Here follows a brief presentation of *A. nidulans* and its biology, and the relevant purine/pyrimidine transporters.

1.6.1 The genus *Aspergillus* & the model system *A. nidulans*

The first scientific recognition of a member of the Aspergilli was reported by Pier A. Micheli (1679-1737), an Italian botanist, who observed microscopically a fungus with a characteristic structure and named it after the similar shaped “aspergillum” (an instrument used by the Roman Catholic Church to sprinkle holy water). Since then, the genus Aspergilli has come to include 344 identified species (Frisvad, 2014); among which less than ten are used systematically by hymans or are implicated in aspects of human life. Particularly, *A. oryzae* has been used in food fermentation, *A.*

niger in industry and production of citric acid, amylase, invertase and other enzymes, *A. fumigatus* is a pathogen for humans and constitutes an important hazard for immunosuppressed patients, *A. flavus* is responsible for the production of a toxic chemical compound (aflatoxin) that can contaminate grain before harvest or after storage and *A. nidulans* is one of the most studied microorganism in biology research. The genus Aspergilli belongs to the phylum of Ascomycetes and, especially, to the subphylum of Pezizomycotina and the class of Eurotiomycetes. At the moment (July 2015) 23 Aspergillus species' genomes have been sequenced as referred in *http://jgi.doe.gov/*, expanding the research field and the data available related to this fungal group. (Goldman and Osmani, 2008)

A. nidulans has been established since the 1940s by G. Pontecorvo, as a model system for the study of genetics. To date, this fungus has been used extensively for studying genetic mechanisms, metabolic regulation, and differentiation, cell cycle controls, transporters, and DNA, RNA and protein regulation. *A. nidulans* characteristics include properties such as: the existence of sexual, asexual and parasexual cycle, rapid growth rate, simple nutritional requirements, transformability (DNA insertion occurs by heterologous or homologous recombination in *ΔnkuA* mutant strains), distinct phenotypes, haploid genome, while under specific conditions diploid cells can also be constructed; all these have rendered it as one of the most studied organisms.

In our lab, *A. nidulans* has been used for about 20 years for the study of purine and pyrimidine transporters. This work has its origins in the study of purine metabolism and the creation of the first mutants related to purine metabolism by C. Scazzocchio in mid '60s (Darlington and Scazzocchio, 1967; Darlington et al., 1965). The great advantage of *A. nidulans* concerning purine/pyrimidine transport and metabolism is its distinct and reproducible growth phenotypes on media containing purines that can be used as nitrogen sources, or toxicity phenotypes on purine and pyrimidine analogues. Thus, mutations affecting purine transporters in a suitable genetic background lead to discernible phenotypes. In addition, the long-term research in this field has introduced simple techniques that allow the kinetic characterization of a transporter by direct uptake assays, as well as null mutant strains of characterized transporters (more than 20 transporters characterized so far) that can be combined through genetic crossing in order to construct strains lacking multiple selected transporters. Different genetic screens or selective approaches can also be applied to obtain mutations in purine or pyrimidine transporters. These mutations can be functionally analyzed and be distinguished into those affecting a transporters topology or function (e.g. transport capacity, substrate specificity, substrate affinity).

1.6.2 The life cycle of *A. nidulans*

As mentioned previously, *A. nidulans* can reproduce through sexual, asexual or parasexual cycles; a diagram of these cycles is shown in **Figure 14**. During the asexual cycle (blue shadowed part of **Figure 14**), a conidiospore (single cell, uninucleate) germinates under favourable conditions producing multicellular hyphae. Under appropriate for development environmental conditions (e.g. air, light) some hyphal cells differentiate into foot cells from which long stalks (70 μM) emerge. At the end of the stalk a vesicle is formed from which a layer of 50-70 metulae is produced through budding. Each metula produces 2-3 phialides, which in turn mitotically produce long chains of uninucleate conidiospores. The conidiospores are dispersed from the conidiophores, initiating another asexual cycle. (Todd et al. 2007, Bayram et al. 2010; and references therein)

When conditions are restraining for vegetative growth, a hypha can undergo a sexual cycle (pink shadowed part of **Figure 14**), where two individual hyphae may fuse to form a heterokaryon (contains nuclei from both parental hyphae). The heterokaryon can be alternatively transformed to a diploid or a homokaryon (parasexual cycle; yellow shadowed part of **Figure 14**). Sexual or asexual cycles can initiate from hetero- or homokaryotic hypha. During the sexual cycle, the end of an ascogenous heterokaryon comprises a final (uninucleate), a penultimate (binucleate) and a basal (uninucleate) cell. The penultimate cell differentiates and results in its enlargement, fusion of its two nuclei and meiosis of the diploid nucleus creating an ascus. The four nuclei produced by meiosis, divide mitotically to generate eight nuclei, included in eight ascospores in the ascus. The nucleus in each ascospore undergoes mitosis resulting in mature binucleate ascospores. Then, the nursing Hülle cells are produced and surround the closed fruiting bodies (cleistothecia, diameter ~200 mm) that contain the sexual spores (ascospores) arranged in nonlinear asci. The Hülle cells are responsible to provide nutrients for the developing cleistothecia. Each cleistothecium contains more than 10,000 binucleate ascospores that were produced by a single ascogenous hypha. (Todd et al. 2007, Bayram et al. 2010, and references therein)

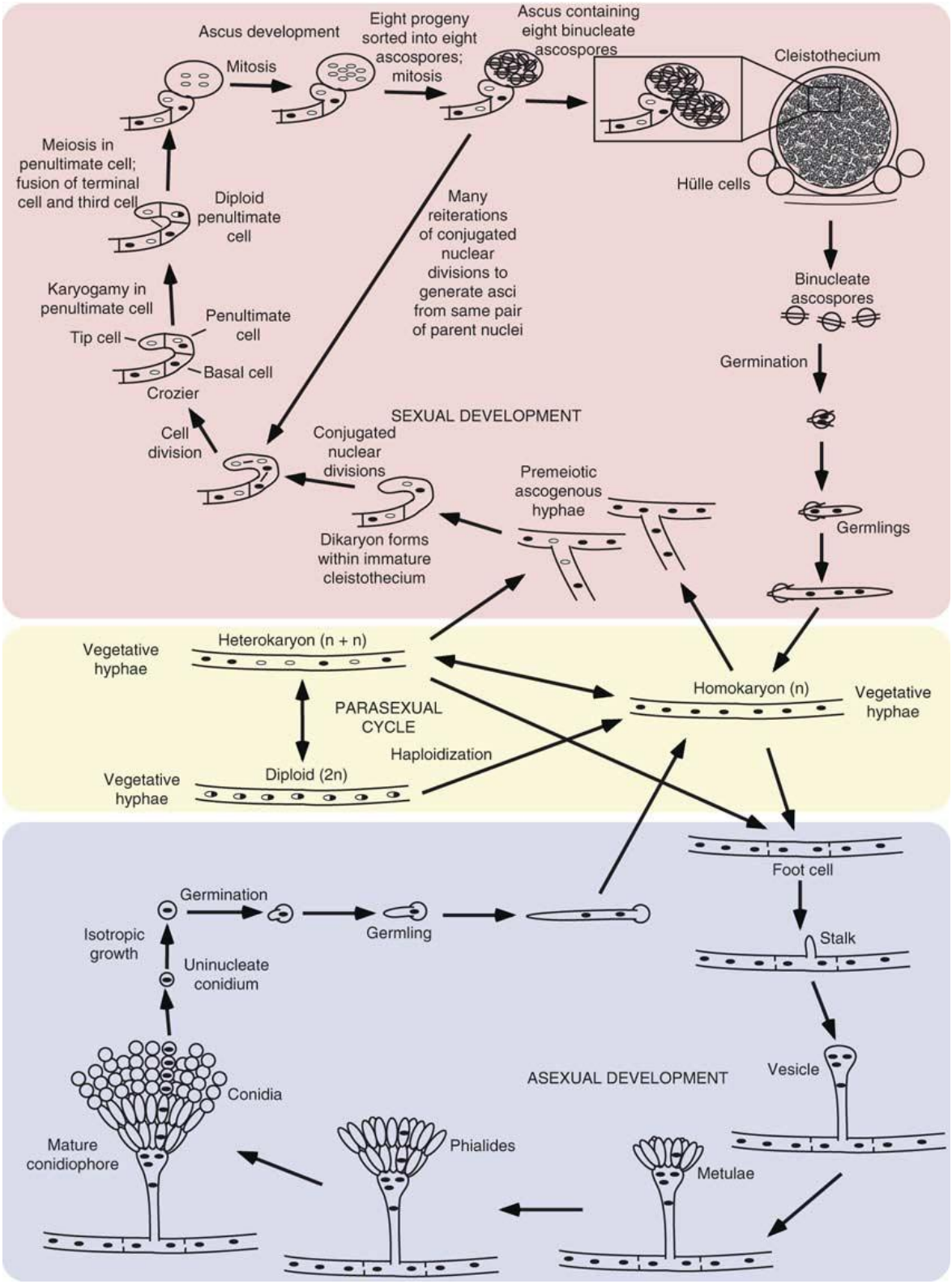


Figure 14. Diagrammatic representation of *A. nidulans* life cycle. Three different cycles exist: Sexual (highlighted in pink), where closed fruiting bodies (cleistothecia) are produced and contain ~10,000 binucleate ascospores, result of meiotic events from a single ascogenous hypha; Asexual cycle (highlighted in blue), where uninucleate conidiospores are produced by specific structures (conidiophores) that emerge under favourable conditions; Parasexual cycle (highlighted in yellow), where two hyphae can merge and create a heterokaryon (two populations of nuclei) or a diploid (one population of merged diploid nuclei). The figure is from Todd et al. (2007).

The morphology of *A. nidulans* in different developmental stages microscopically and macroscopically is shown in **Figure 15**. In particular, a germinated conidiospore in two stages is shown in **Figure 15A**, a stereoscopic observation of mycelia and conidiophores in **Figure 15B** and a macroscopic view of multiple colonies of *A. nidulans* strains of different colours, grown on solid medium in **Figure 15C**.

Figure 15. **Morphology of *A. nidulans* observed microscopically and macroscopically. (A)** Microscopy of germinated conidiospores in two different growth stages. **(B)** Stereoscopy of hyphae and conidiophores after pseudocoloring. **(C)** Grown colonies of *A. nidulans* strains of different colours on a solid medium. The Figure is adapted from *http://www.bartnieuwenhuis.eu/* and *http://www.microbeworld.org*.

The transition from the one developmental program to the other is coordinated with the control of secondary metabolism. Light is one of the several environmental factors regulating the secondary metabolism and hence the development of *A. nidulans*. Most strains used in laboratories contain a single mutation in the *velvet* gene (*veA1*), first reported by (Käfer, 1965) leading to increased conidiation. It was not until many years later that the mechanism underlying this phenotype was discovered. A single mutation in *veA* led to a shorter version of the protein resulting in cytoplasmic localization, unlike the nuclear or nuclear and cytoplasmic localization of the wt protein regulated by the light (Kim et al., 2002; Stinnett et al., 2007). The VeA protein can actually activate sexual cycle in the absence of light, inhibiting asexual cycle and hence reducing conidiation (Kim et al., 2002). The presence of *veA1* genetic background allows the abundant conidiation of strains under all light conditions, easing experimental procedures.

1.6.3 NCS1 transporters

The *A. nidulans* genome includes twelve proteins that belong to the NCS1 family. Based on their homology to the characterized homologues from *S. cerevisiae* (Fcy2p and Fur4p; Jund et al. 1988, Weber et al. 1990), they could be divided in two subfamilies, the Fcy-like and Fur-like subfamily (de Koning and Diallinas, 2000). The Fcy2p protein from *S. cerevisiae* was characterized as a high-affinity, high-capacity adenine-hypoxanthine-guanine-cytosine/H^+ symporter (Brèthes et al., 1992), while the

Fur4p as a high affinity, high capacity uracil/H^+ symporter (Wagner et al., 1998). Based on sequence alignments 5 homologues of Fcy2p were identified in *A. nidulans* (Fcy-like proteins) together with 7 homologues of Fur4p (Fur-like proteins). Even though, the Fur4p and Fcy2p share an overall sequence identity of ~18%, specific motifs were identified in both protein sequences indicating their phylogenetic relation, classifying them in the NCS1 family (de Koning and Diallinas, 2000).

At the beginning of the present work, FcyB was the only characterized member of the Fcy-like subfamily in *A. nidulans* (Vlanti and Diallinas, 2008), but currently preliminary data concerning the function of the other Fcy-like proteins are available (G. Sioupouli and G. Diallinas, unpublished results). FcyB is a high-affinity, low-capacity transporter specific for adenine, guanine, hypoxanthine, cytosine and the toxic analogue 5-fluorocytosine (**Figure 16**) (Vlanti and Diallinas, 2008). FcyB binds adenine with K_i 7 μM, hypoxanthine and cytosine with $K_{m/i}$ 20 μM and guanine with K_i 11.5 μM. FcyB contributes only ~15% in the uptake of adenine, hypoxanthine and guanine, since the major transporter for the uptake of these purines is another transporter, namely AzgA (see below). FcyB however is the only known transporter in *A. nidulans* specific for cytosine. FcyB shares 43.1% amino acid sequence identity with its orthologue in *S. cerevisiae* (Fcy2p) and 31-34% identity with the rest Fcy-like proteins of *A. nidulans* of unknown function (Vlanti and Diallinas, 2008).

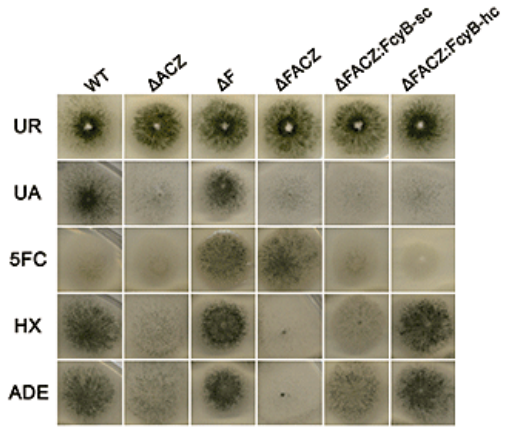


Figure 16. Functional characterization of FcyB. Growth test analysis of selected strains revealing the function and specificity of FcyB. Strains included are: wild type (wt), a triple deleted *ΔuapA ΔuapC ΔazgA* mutant (ΔACZ) and single copy (sc) and multicopy (hc) ΔACZ transformants expressing FcyB. Growth was tested on MM, at 37°C, at pH 6.8, supplemented with urea (UR), uric acid (UA), 5-fluorocytosine plus urea (5-FC), hypoxanthine (HX) or adenine (ADE). FcyB function confers as leaky growth on hypoxanthine and adenine and sensitivity to 5-fluorocytosine under its native promoter. Multiple copies of FcyB increase the growth phenotype on hypoxanthine and adenine. The Figure is adapted from (Vlanti and Diallinas, 2008).

Northern analysis revealed that *fcyB* is highly transcribed 3 h after conidiospore germination at 37 ^oC in the presence of a neutral nitrogen source (e.g. urea). Lower amounts of mRNA were detected at the onset of mycelium development (4-6 h) under the same conditions (Vlanti and Diallinas, 2008).

Similarly to other transporters studied, FcyB tagged with the green fluorescent protein (GFP) is mostly localized in the plasma membrane and occasionally in internal structures identified as the endoplasmic reticulum and the vacuoles (Vlanti

and Diallinas, 2008). Finally, conditions triggering substrate endocytosis (see later) revealed that FcyB-GFP was sensitive to endocytosis in the presence of ammonium, or excess of its substrates, adenine and cytosine (Vlanti and Diallinas, 2008).

The Fur-like subfamily consists of seven members (Fur-A, -B, -C, -D, -E, -F and -G) out of which two have been functionally characterized as uracil (FurD) or allantoin (FurA) transporters (Amillis et al., 2007; Hamari et al., 2009). FurA shares a 37% sequence identity with its homologue in *S. cerevisiae* (Fur4p) and 25-28% identity with the other *A. nidulans* Fur-like proteins. All seven *fur* null mutant strains were constructed and tested on various purines and pyrimidine-analogues (**Figure 17**A) (Hamari et al., 2009). This analysis revealed that FurA was specific for allantoin and FurD for uracil and the toxic analogue 5-fluorouracil, while no obvious phenotype was shown for the deletion of other paralogues under conditions tested (Hamari et al., 2009).

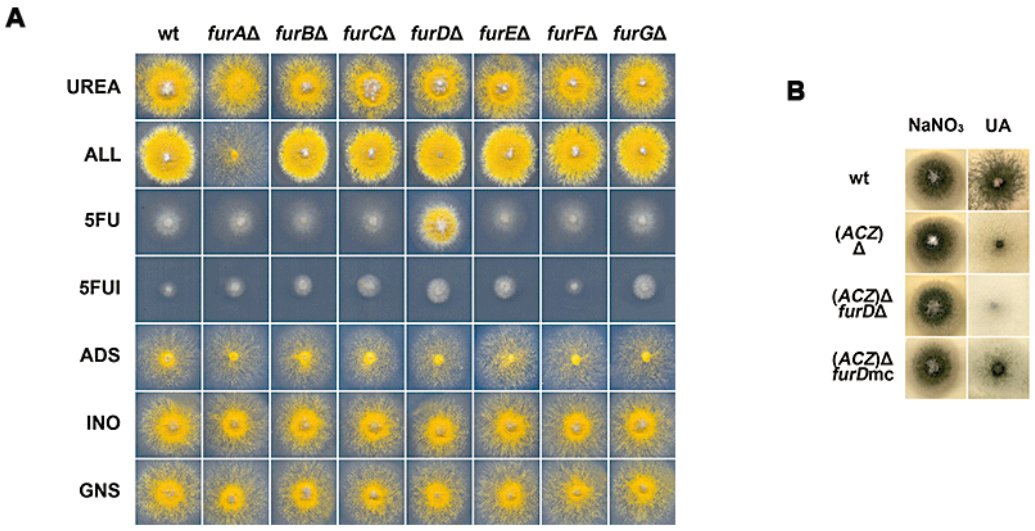


Figure 17. Functional characterization of *fur* null mutant strains. (A) Growth test analysis of all *fur* null mutant strains and a wild type (wt) control strain on urea, allantoin (ALL), adenosine (ADS), inosine (INO) and guanosine (GNS) as sole nitrogen sources and on the toxic analogues 5-fluorouracil (5FU) or 5-fluorouridine (5-FUI) on $NaNO_3$ at 37^{o}C pH 6.8. FurA function is revealed as lack of growth on allantoin in the absence of FurA, and FurD function as resistance on the toxic analogue 5-fluorouracil in the absence of FurD. **(B)** Growth test analysis on uric acid (UA) of multiply deleted mutant strains: *ΔuapA ΔuapC ΔazgA* (ΔACZ), *ΔuapA ΔuapC ΔazgA ΔfurD* (ΔACZ *furDΔ*), *ΔuapA ΔuapC ΔazgA* multiple integrations of *furD* (ΔACZ *furD*mc). The Figure is adapted from (Hamari et al., 2009).

Northern analysis (Hamari et al., 2009) showed that *furD* was highly expressed at 3 h after conidiospore germination at 37 ^{o}C, and remained relatively high after 4 and 5 h. RNA levels for *furA*, *furC* and *furE* were moderately detectable at 3 h after conidiospore germination at 37 ^{o}C and remained detectable at 4-5 h, while *furF*- and *furG*-RNA were marginally detected at 3h after conidiospore germination at 37 ^{o}C and almost undetectable at 4-5 h. *furB*-RNA was not detected under any of the conditions tested. Probably the inability to detect any function for FurB, FurC, FurE,

FurF and FurG could be related to the very low or null expression of these genes, at least under the conditions tested.

FurD is the only Fur protein extensively characterized by biochemical assays. Characterization of FurA is not possible since there is not commercially available radiolabeled allantoin to carry out uptake studies. FurD is a high capacity uracil/H^+ symporter that binds with high affinity uracil and the antifungal 5-fluorouracil ($K_{m/i}$ 0.45 μM), and with medium affinity uric acid and xanthine (K_i 100 μM) (Amillis et al., 2007). The minor contribution of FurD in growth on uric acid or xanthine is apparent only in the absence of the two major transporters for these purines, namely UapA and UapC (Hamari et al., 2009). This is reflected in growth tests, where the triple mutant lacking functional *furD*, *uapA* and *uapC* genes exhibits virtually no growth on uric acid or xanthine, whereas the double mutant Δ*uapA* Δ*uapC* still shows a degree of leaky growth, due to the activity of FurD. Overexpression of *furD*, via plasmid multiple copies integration, in a *ΔuapA ΔuapC* genetic background can confer full growth on uric acid or xanthine (**Figure 17B**) (Hamari et al., 2009).

Evolutionary relationships in the NCS1 family

All characterized fungal Fcy-like proteins are specific for adenine, hypoxanthine, guanine and cytosine (FcyB from *A. nidulans*, Fcy2p from *S. cerevisiae*, Fcy21p from *Candida albicans*; Weber et al. 1990, Goudela et al. 2006, Vlanti and Diallinas 2008) and all characterized fungal Fur-like proteins are specific for uracil, allantoin or uridine (FurD and FurA from *A. nidulans*, Fur4p, Dal4p and Fui1p from *S. cerevisiae*, Fur4 and Dal4 from *Candida lusitaniae*, Fur4 from *Schizosaccharomyces pombe*; Gabriel et al., 2014; Jund et al., 1988; de Montigny et al., 1998; Yoo et al., 1992). Interestingly, plant NCS1 homologues combine the specificities of both subfamilies, Fcy and Fur, particularly AtNCS1/PLUTO from *A. thaliana* is specific for adenine, guanine and uracil (Mourad et al., 2012; Witz et al., 2012) and CrNCS1 from *Chlamydomonas reinhardtii* is specific for adenine, guanine, uracil and allantoin (Schein et al., 2013).

Phylogenetic analysis suggests that the NCS1 family originates from prokaryotes and through horizontal gene transfer was established in fungi (Krypotou et al., 2015a). The two subfamilies, Fur and Fcy, appear to have emerged from independent gene transfers, while the plant NCS1 homologues are more related to the bacterial sequences in-between the Fur and Fcy subfamilies (**Figure 18**) (Krypotou et al., 2015a). The *fur* and *fcy* genes are predicted to have arisen through multiple gene duplications. The fungal *fcys* could probably be true orthologues, while the fungal *furs* are more distant and their common specificities might have evolved independently more than once through convergent evolution (Hamari et al., 2009; Krypotou et al., 2015a).

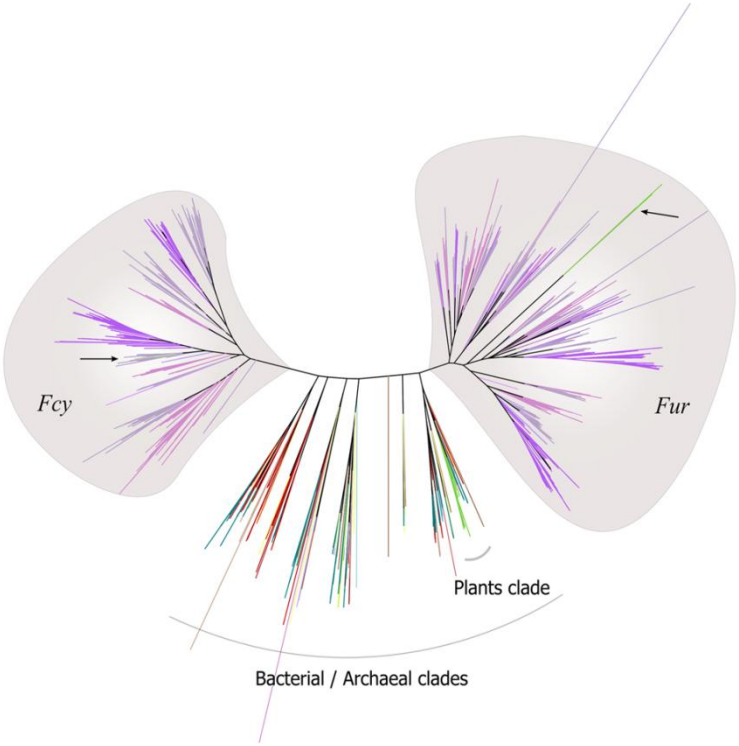


Figure 18. Phylogeny of NCS1 proteins. Maximum Likelihood phylogenetic tree of all NCS1 proteins from all domains of life. Fcy and Fur fungal proteins correspond to purple lines. Plant NCS1 proteins are in green, Actinobacteria, Proteobacteria and Firmicutes are in red, blue and yellow respectively. Cases of recent horizontal gene transfers from fungi to other groups are indicated with arrows. The Figure is adapted from (Krypotou et al., 2015a).

1.6.4 NAT/NCS2 transporters

A. nidulans includes two NAT members, the uric acid-xanthine/H^+ symporters UapA and UapC. UapA is one of the most extensively studied eukaryotic transporters with respect to regulation of expression, structure-function relationships and trafficking. It was identified in the mid 60s (Darlington and Scazzocchio, 1967), but cloned and characterized a decade later (Diallinas and Scazzocchio, 1989; Gorfinkiel et al., 1993). UapA is a high-affinity, high-capacity transporter with 7-8 μM binding affinity for uric acid and xanthine (**Figure 19A**). *uapA* transcription is developmentally induced after 5h of conidiospore germination at 37 ^{o}C in the presence of a neutral nitrogen source (e.g. urea) and remains at basal levels during vegetative growth (Amillis et al. 2004). In mycelia, *uapA* expression is induced by uric acid or repressed by ammonium (Amillis et al., 2004; Diallinas and Scazzocchio, 1989; Gorfinkiel et al., 1993), is also transiently induced under nitrogen starvation conditions (Amillis et al., 2004).

A UapA C-terminally tagged with GFP proved fully functional and allowed extensive studies on the subcellular expression, trafficking and turnover of UapA (Gournas et al., 2010; Karachaliou et al., 2013; Martzoukou et al., 2015; Pantazopoulou et al., 2007). In brief, UapA-GFP is localised principally in the plasma

membrane and occasionally in the vacuoles, as a result of low level constitutive turnover. Prominent vacuolar degradation is elicited in the presence of ammonium (i.e. a primary nitrogen source), excess substrates or stress (high temperature, antifungal drugs, oxidizing or reducing agents; Gournas et al. 2010, Karachaliou et al. 2013; G. Diallinas unpublished). An extensive study on the properties and mechanism underlying UapA-GFP vacuolar turnover, showed that C-terminal ubiquitination (probably de-ubiquitination) is the principal molecular modification necessary for initiation of internalization through endocytosis, sorting to early endosomes, MVB (multivesicular body) and eventual degradation in the vacuoles (Gournas et al., 2010; Karachaliou et al., 2013).

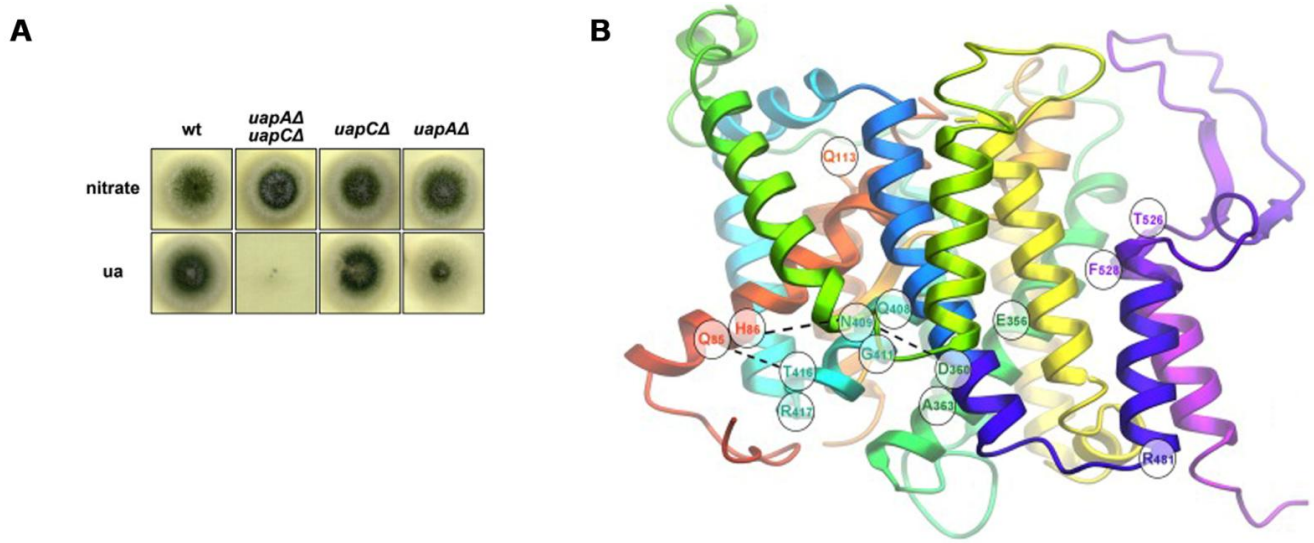


Figure 19. UapA function and putative structure. (A) Growth test analysis of strains expressing only *uapA* (*uapCΔ*) or *uapC* (*uapAΔ*) or neither of the two (*uapAΔ uapCΔ*) or both of them (wt) on uric acid (ua) and nitrate. **(B)** Putative structural model based on homology modeling on UraA. Residues identified through genetic and mutational approaches are highlighted in circles. The Figure is adapted from (Diallinas, 2013; Kosti et al., 2012).

To date more than 200 UapA mutants and chimeric molecules (see below) have been constructed and functionally characterized, revealing specific amino acids involved directly in substrate binding (e.g. Glu356 and Gln408) or in substrate translocation (Gln85, His86, Asn409, Gly411, T417) (Diallinas 2014; and references therein). A theoretical structural model has also been proposed based on homology modelling on the crystal structure of UraA (**Figure 19B**) (Amillis et al., 2011; Kosti et al., 2012). The model predicts that the major substrate (xanthine) binding site is formed by four residues located in TMS3 (Phe155), TMS8 (Glu356) and TMS10 (Ala407 and Gln408). Interestingly, mutants of UapA were isolated through genetic screening and systematic mutational analysis that had acquired the ability to transport non-physiological substrates, such as adenine and hypoxanthine, without significantly affecting the binding affinity and transport of physiological substrates (Papageorgiou et al. 2008, Kosti et al. 2010). The mutations that alter the specificity of UapA were located in TMSs 11, 12, 13, 14 and the loop connecting TMSs 1 and 2, briefly,

regions unrelated to the TMSs forming the binding site (TMSs 3, 8 and 10). The observation that residues outside the binding site can affect the specificity of a transporter led to the hypothesis that UapA might contain channel-like gating domains or filters that also affect substrate selection. Indeed, under the light of the modelled structure of UapA, the relevant residues affecting UapA specificity are positioned at putative outer and inner *gates,* at the beginning and the end of a substrate translocation trajectory, detected through Molecular Dynamic approaches (Kosti et al., 2010, 2012; Papageorgiou et al., 2008). The existence of channel-like gates in transporters is further discussed in the course of the present thesis.

UapC is a moderate capacity transporter, exhibiting high-affinity binding for xanthine (K_m 3.3 μM) but only moderate-affinity binding for uric acid (K_i 135 μM). UapC can also bind with very low affinity other purines or pyrimidines (adenine, hypoxanthine, guanine and uracil) (Diallinas et al., 1995; Krypotou and Diallinas, 2014; Papageorgiou et al., 2008). UapC is responsible for the residual growth on uric acid and xanthine in the absence of UapA (**Figure 19A**). The UapC transcription profile is very similar to that of UapA and most transporters involved in the uptake of nitrogenous compounds. Transcription initiates in early stages (2h) of conidiospore germination and remains at low constitutive levels in mycelium, unless is induced by substrates (uric acid). *uapC* transcription is also repressed by ammonium, but not as tightly as that of *uapA* (Amillis et al., 2007; Diallinas et al., 1995). Functional UapC-GFP, similarly to UapA-GFP, is mainly localised in the plasma membrane and occasionally in the vacuoles in the absence of ammonium or substrates (Valdez-Taubas et al., 2000). The UapC orthologue of *A. fumigatus* has also been characterized and shown to have a similar kinetic and specificity profile (Goudela et al., 2008). Finally, functional chimeric UapC-UapA molecules have been instructive in the identification of the binding site of these transporters (Diallinas et al., 1998).

Evolutionary relationships in the NAT family

NAT members of the UapA subfamily have been identified in all domains of life. Most functionally characterized members are H^+ (bacteria, fungi, plants) or Na^+ (mammals) coupled symporters specific for the uptake of xanthine, uric acid, uracil and related analogues. However, NAT members in primates are solely specific for L-ascorbic acid (Faaland et al., 1998; Tsukaguchi et al., 1999).

Phylogenetic analysis of fungal species shows that NATs are ubiquitous in dikarya but show a patchy distribution in more primitive fungi (**Figure 20**) (Krypotou et al., 2015b). UapC-like proteins are present in almost all fungi with identities 74-91% among them. UapA-like proteins are present in many fungi, but in only three Aspergilli (*A. nidulans*, *A. zonatus* and *A. ochraceoroseus*). Until recently only two groups of fungal NATs were known, the UapA- and the UapC-like proteins (or "canonical" NATs). However, the current abundance of genome sequences available led to the identification of a third group of homologues, named UapD-like, with

characteristics that diversify them from the rest NATs (Krypotou et al., 2015b). UapDs lack significant sequence motifs present in all “canonical” NATs, while sharing a sequence identity < 55% with them. UapD proteins (45-48% sequence identity among themselves) show patchy conservation in ascomycetes and are absent from basidiomycetes, suggestive of significant evolutionary losses. The UapD clade is predicted to be ancestral, since it includes homologues in divergent fungi as Aspergilli and *Tuber melanosporum*, a basal species of Pezizomycotina. UapD homologues are most abundant in dothideomycetes and eurotiomycetes. Interestingly, a UapD homologue is also present in a single member of the Saccharomycotina, *Kuraishia capsulata*.

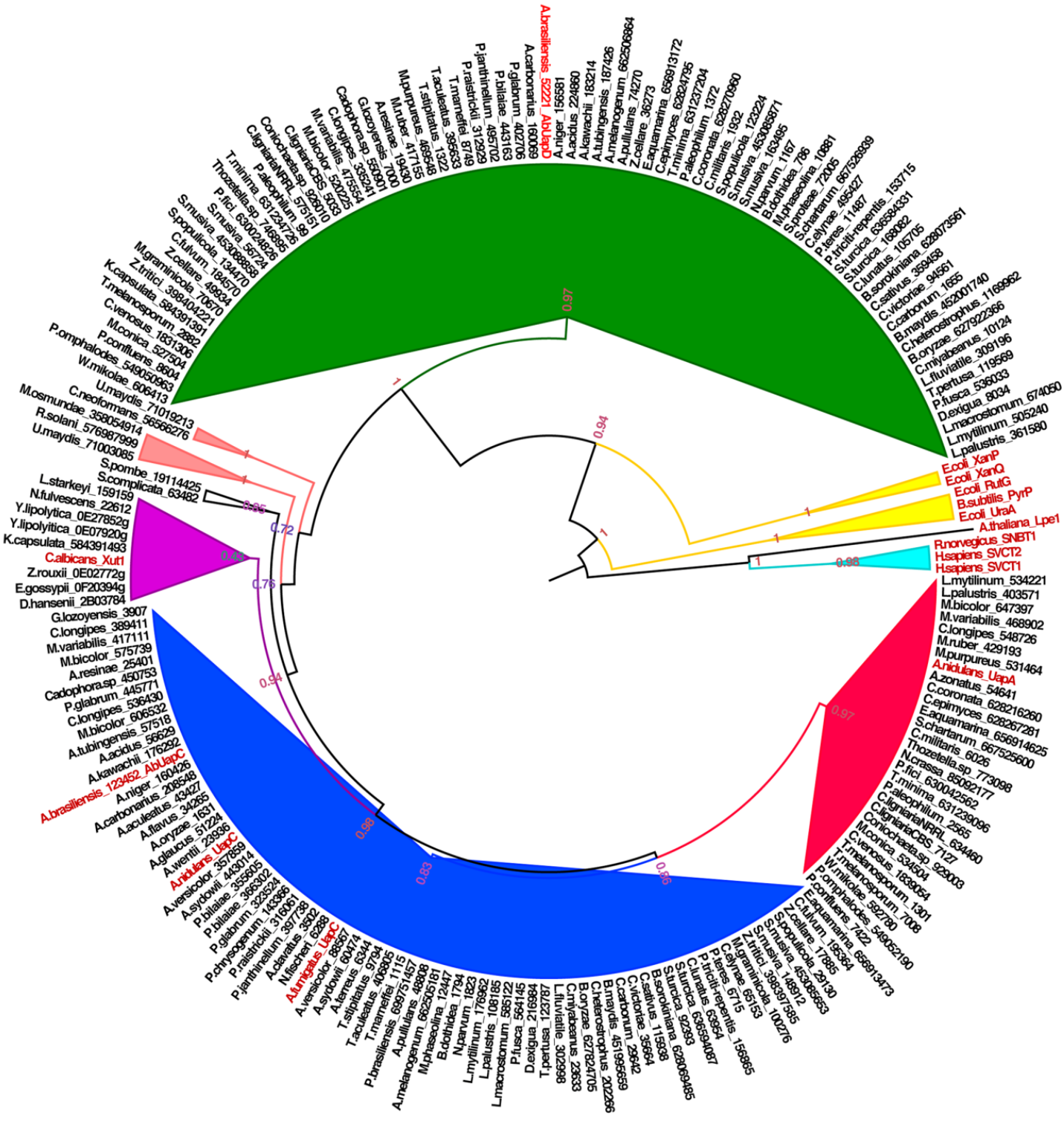


Figure 20. **Phylogenetic analysis of fungal NAT proteins.** Maximum Likelihood Phylogenetic rooted circular tree of NAT proteins of dikaryotic fungi in cartoon form. The cartoons comprising major clades are color coded as: Pezizomycotina: UapC-like, blue; UapA-like, red; UapD-like, green; Saccharomycotina: purple; Basidiomycetes: pink; Bacterial NATs: yellow; Mammalian NATs: cyan. The branches of the members included from the Taphrinomycotina (*Schizosaccharomyces pombe* and *Saitoella complicata*) and Plants (*A. thaliana*) respectively are in black. The Figure is adapted from (Krypotou et al., 2015b).

1.6.5 AzgA, the founding member of a NAT subfamily

AzgA, the major adenine, guanine and hypoxanthine transporter in *A. nidulans*, contributing at a level of 85% in the uptake of these purines (see **Figure 16**), was identified through toxic analogue resistance screens on 8-azaguanine (Darlington and Scazzocchio, 1967). AzgA functions as H^+ symporter and shows high-affinity for hypoxanthine (K_m 1.4 μM), adenine (K_i 3 μM) and guanine (K_i 3 μM) (Cecchetto et al., 2004; Goudela et al., 2006). The transcription profile of *azgA* is similar to that of *uapA*, *uapC* and other nitrogenous compound transporters, being developmentally induced during germination and partially controlled by substrate induction and ammonium repression in mycelia (Amillis et al., 2004; Cecchetto et al., 2004). Functional AzgA-GFP is localised in the plasma membrane in the presence of a neutral nitrogen source (e.g. urea), but it is significantly degraded in the vacuoles after endocytosis in the presence of substrate (Karachaliou et al., 2013).

AzgA-like proteins are present in bacteria, archaea, fungi and plants, but absent from metazoa. Fungal and plant homologues share an identity of 35%, but are more distant to the bacterial and archaeal homologues with an identity of 25-30% (Krypotou et al., 2014). A single *azgA* gene is present in all Aspergilli with an identity of 87-94% rendering them orthologues. Identification of AzgA defined a new purine transporter family, called the AzgA-like family, appearing phylogenetically distinct from the other transporter families (Cecchetto et al., 2004). However, recent data showed that the AzgA-family appears structurally related to the NAT family, as its topology can be modelled using the UraA structure (see Results in the present thesis). Thus, the AzgA family can also be considered as an evolutionary distant member of the NAT family (**Figure 21**) (Cecchetto et al., 2004; Krypotou et al., 2014).

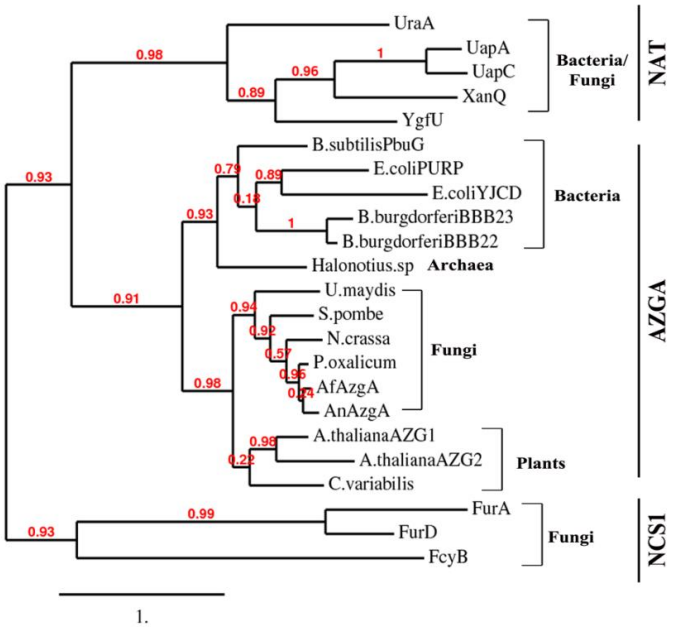


Figure 21. Phylogeny of AzgA-like proteins in respect to NAT and NCS1 families. Maximum Likelihood phylogenetic tree that includes AzgA-like sequences of known function, plus some selected representatives as out-groups, and selected NAT and NCS1 proteins of known function and substrate specificity. The Figure is adapted from (Krypotou et al., 2014).

1.7 Goals and purposes of this study

The purpose of the current study is to establish through structure-function approaches the amino acid residues critical for substrate binding, specificity and transport of *A. nidulans* purine/pyrimidine transporters. The study of various transporters originating from different organisms is necessary, because, even though many crystal structures - representatives of different transporter families- are available, the exact structure or transport mechanism does not apply in all homologous proteins.

Only few crystal structures are available for transmembrane proteins compared with a plethora for cytoplasmic proteins. This is due to the fragile nature of the structure of a transmembrane protein during *in vitro* procedures. As a result, the majority of crystal structures available today, belong to functionally uncharacterized proteins of thermophilic bacteria. The choice of these proteins is due to their endogenous physical properties to withstand the high temperatures used during crystallization and X-ray diffraction protocols. Thus, building structural models of well-characterized proteins through homology modelling allows the study and visualisation of the molecular mechanism of transport and substrate specificity. In addition, the rational design of drugs targeting specific proteins to enter the cells of a pathogenic organism is more and more compelling, since pathogens have become resistant to many drugs used nowadays. A successful drug that binds solely to its target proteins can be developed only under the prism of the knowledge of the transport mechanism.

The use of *A. nidulans* in the studies of transporters has been established over fifty years ago with the isolation of the first resistant strains to purine toxic analogues. Since then, over 20 different transporters responsible for the uptake of important nitrogenous and carbon compounds have been characterized. Furthermore, the metabolic pathways of purines and pyrimidines have been studied. All the above combined with the excellent knowledge concerning the physiology and cellular mechanisms of *A. nidulans*, the over the years development of different experimental protocols and the easy handling of this organism, qualify *A. nidulans* as one of the best model organisms for the study of transporters.

The current thesis deals with transporters from the three major purine/pyrimidine transport families NCS1, NAT/NCS2 and AzgA-like. More specifically, we extensively studied the major adenine/guanine/hypoxanthine transporter AzgA, the secondary adenine/guanine/ hypoxanthine/cytosine transporter FcyB and the uracil transporter FurD. In addition, we attempted to characterize a member of the newly identified NAT subgroup UapD and to investigate the probable function of orphan Furs through overexpression.

The results of this study are divided in four distinct sections:

I. ***Structure-function analysis of the purine transporter AzgA*****:** The major purine transporter of *A. nidulans*, AzgA, was functionally analyzed through mutational analysis in order to determine residues participating in substrate binding

and/or transport. Uncovering important residues for the AzgA function provided information concerning structure, transport mechanism and substrate specificity, as well as the probable structural relation between AzgA and the NAT family.

II. ***Identification of the function of the non-canonical NATs*****:** The discovery of a new subgroup of the NAT proteins, UapDs, allows the characterization of another NAT protein group that was so far unknown. The fact that these proteins lack critical for function residues present in "canonical" NATs and appear phylogenetically closer to a common ancestral protein raises questions concerning their functionality and specificity profile. We thus selected the UapD-like protein of *A. brasiliensis*, in order to heterologously express and characterize it in *A. nidulans*. The UapC-like protein from *A. brasiliensis* was used as a control for proper heterologous expression.

III. ***Structure function analysis of the purine/cytosine transporter FcyB*****:** The crystal structures of Mhp1 provide information concerning the residues participating in substrate binding in the NCS1 family. An extensive mutational analysis on FcyB, an NCS1 member, revealed residues critical for specificity and/or function. We also attempted to change the specificity profile of FcyB by specific amino acid substitutions that mimic NCS1 homologues with different specificities.

IV. ***Identification of the function of the orphan Furs and structure-function analysis of FurD*****:** We used over-expression in order to seek for a function relating to the Fur proteins of *A. nidulans* that did not show any obvious phenotype when the relevant null mutants were constructed (Fur-B, -C, -E, -F and -G). The interesting phylogenetic profile of the NCS1 family and especially the Fur subfamily could be re-evaluated by the discovery of the function of each Fur protein. In addition, data from the Mhp1 structure and the FcyB mutational analysis indicated residues critical for the function of the Fur proteins. We selected FurD, the sole characterized Fur protein, to perform a mutational analysis in order to identify the residues participating in substrate specificity and transport. In addition, we have attempted to alter the specificity profile of FurD by rationally designed mutations that mimic FurA or other NCS1 homologues.

An approach that relates function to structure has been used widely in the study of transporters. The crystal structures and the computational analyses are not sufficient to reveal and describe the transport mechanism of each protein. Multiple data arising from genetics, molecular biology, crystallography, biochemistry, molecular dynamics, docking, homology modelling, must be combined in order to understand transporter function. We hope that the data arising from this study will support, expand and reinforce the current knowledge concerning transporters in general and not only in *A. nidulans*.

2 Materials and Methods

2.1 Strains

The strains used in this study, either constructed herein or previously published, are included in **Table 2**.

Table 2. Strains used in this study. Genotype and reference for every strain is included in the table. Strain names are as referred in this study.

Strain name	Genotype	Reference
wt	*pabaA1*	FGSC
wt	*pyroA4*	G. Diallinas, unpublished
wt pyrG89	*nkuAΔ::argB pyroA4 pyrG89*	Nayak et al., 2006
ΔACZ	*ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1*	Pantazopoulou et al., 2007

ΔACZ argB2	*ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2*	Vlanti et al., 2006
ΔΔ	*ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
Δ7	*ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pantoB100 pabaA1*	This study.
FcyB-Δ7	*pBS-uapA$_p$-FcyB-pabaA ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FcyB/V83N-Δ7	*pBS-uapA$_p$-fcyB/V83N-pabaA ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FcyB/N163Q-Δ7	*pBS-uapA$_p$-fcyB/N163Q-pabaA ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FcyB/N163L-Δ7	*pBS-uapA$_p$-fcyB/N163L-pabaA ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FcyB/S261A-Δ7	*pBS-uapA$_p$-fcyB/S261A pabaA ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FcyB-ΔΔ	*pGEM-uapA$_p$-fcyB AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N81A ΔΔ	*pGEM-uapA$_p$-fcyB/N81A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N81S ΔΔ	*pGEM-uapA$_p$-fcyB/N81S AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/S85A ΔΔ	*pGEM-uapA$_p$-fcyB/S85A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/W159A ΔΔ	*pGEM-uapA$_p$-fcyB/W159A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/T191A ΔΔ	*pGEM-uapA$_p$-fcyB/T191A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/E206A ΔΔ	*pGEM-uapA$_p$-fcyB/E206A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/W259A ΔΔ	*pGEM-uapA$_p$-fcyB/W259A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N350A ΔΔ	*pGEM-uapA$_p$-fcyB/N350A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N350Q ΔΔ	*pGEM-uapA$_p$-fcyB/N350Q AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N350D ΔΔ	*pGEM-uapA$_p$-fcyB/N350D AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N351A ΔΔ	*pGEM-uapA$_p$-fcyB/N351A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/P353A ΔΔ	*pGEM-uapA$_p$-fcyB/P353A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N354A ΔΔ	*pGEM-uapA$_p$-fcyB/N354A AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N354Q	*pGEM-uapA$_p$-fcyB/N354Q AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA*	This study.

ΔΔ	*ΔfcyB::argB pabaA1 riboB2*	
FcyB/N354D ΔΔ	*pGEM-uapA$_p$-fcyB/N354D AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB-GFP ΔΔ	*pGEM-uapA$_p$-fcyB-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N81A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N81A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N81S-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N81S-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/V83N-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/V83N-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/S85A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/S85A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/W159A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/W159A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N163Q-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N163Q-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N163L-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N163L-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/T191A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/T191A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/E206A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/E206A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/W259A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/W259A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/S261A-GFP-ΔΔ	*pGEM-uapA$_p$-fcyB/S261A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N350A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N350A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N350Q-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N350Q-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N350D-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N350D-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N351A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N351A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/P353A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/P353A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N354A-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N354A-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N354Q-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N354Q-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
FcyB/N354D-GFP ΔΔ	*pGEM-uapA$_p$-fcyB/N354D-gfp AfriboB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB pabaA1 riboB2*	This study.
AzgA-GFP ΔACZ	*pBS-azgA-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	(Karachaliou et al., 2013)
AzgA/T49A-GFP ΔACZ	*pBS-azgA/T49A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/T49S-	*pBS-azgA/T49S-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.

GFP ΔACZ		
AzgA/G129A -GFP-ΔACZ	*pBS-azgA/G129A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/G129F-GFP ΔACZ	*pBS-azgA/G129F-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/N131A -GFP ΔACZ	*pBS-azgA/N131A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/N131D -GFP ΔACZ	*pBS-azgA/N131D-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/F335A-GFP ΔACZ	*pBS-azgA/F335A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/D339A -GFP ΔACZ	*pBS-azgA/D339A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/D339N -GFP ΔACZ	*pBS-azgA/D339N-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/D339E-GFP ΔACZ	*pBS-azgA/D339E-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/D342A -GFP ΔACZ	*pBS-azgA/D342A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/D342N -GFP ΔACZ	*pBS-azgA/D342N-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/D342E-GFP ΔACZ	*pBS-azgA/D342E-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/T346A-GFP ΔACZ	*pBS-azgA/T346A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/T390A-GFP ΔACZ	*pBS-azgA/T390A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/E394A-GFP ΔACZ	*pBS-azgA/E394A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/E394Q-GFP ΔACZ	*pBS-azgA/E394Q-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/E394D-GFP ΔACZ	*pBS-azgA/E394D-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/S395A-GFP ΔACZ	*pBS-azgA/S395A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/S395N-GFP ΔACZ	*pBS-azgA/S395N-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/G398A -GFP ΔACZ	*pBS-azgA/G398A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/G402A -GFP ΔACZ	*pBS-azgA/G402A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/K404R-GFP ΔACZ	*pBS-azgA/K404R-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
AzgA/D478A -GFP ΔACZ	*pBS-azgA/D478A-gfp argB ΔuapA ΔuapC::AfpyrG ΔazgA argB2 pabaA1*	This study.
UapA-GFP ΔACZ	*uapA-gfp::argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1*	(Pantazopoulou et al., 2007)
UapC-GFP-	*uapC-gfp-argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2*	(Pantazopoulou

ΔACZ		et al., 2007)
AbUapD ΔACZ	*pBS-uapA$_p$-AbuapD-argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2*	This study.
AbUapC-GFP ΔACZ	*pBS-uapA$_p$-AbuapC-gfp-argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2*	This study.
AbUapD-GFP ΔACZ	*pBS-uapA$_p$-AbuapD-gfp-argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2*	This study.
GFP-AbUapD ΔACZ	*pBS-uapA$_p$-gfp-AbuapD-argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2*	This study.
AbUapD-A-GFP ΔACZ	*pBS-uapA$_p$-AbuapD-A-gfp-argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2*	This study.
ADA-GFP ΔACZ	*pBS-uapA$_p$-ADA-gfp-argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2*	This study.
ΔFurA	*ΔfurA::riboB riboB2 pantoB100 biA1*	S. Amillis, unpublished
ΔFurC	*ΔfurC::riboB riboB2 yA2*	(Hamari et al., 2009)
ΔFurD	*ΔfurD::riboB riboB2 pantoB100*	(Hamari et al., 2009)
ΔFurF	*ΔfurF::riboB riboB2 pantoB100 pyroA4 biA1*	S. Amillis, unpublished
ΔFurG	*ΔfurG::riboB riboB2 pyroA4 biA1*	S. Amillis, unpublished
ΔFurA pyrG89	*ΔfurA::riboB pyrG89 riboB2 pyroA4 biA1*	(Hamari et al., 2009)
ΔFurC pyrG89	*ΔfurC::riboB pyrG89 riboB2 pyroA4 biA1*	(Hamari et al., 2009)
ΔFurD pyrG89	*ΔfurD::riboB pyrG89 riboB2 pyroA4 biA*	(Hamari et al., 2009)
ΔFurF pyrG89	*ΔfurF::riboB pyrG89 riboB2 pyroA4 biA1*	(Hamari et al., 2009)
ΔFurG pyrG89	*ΔfurG::riboB pyrG89 riboB2 pyroA4 biA1*	(Hamari et al., 2009)
ΔFurA ΔFurD	*ΔfurD::riboB ΔfurA::riboB riboB2 pantoB100 biA1*	S. Amillis, unpublished
ΔFurD ΔFurA pyrG89	*ΔfurD::riboB ΔfurA::riboB pyrG89 riboB2 pyroA4 pantoB100 biA*	This study.
ΔFurD ΔFurC	*ΔfurD::riboB ΔfurC::riboB riboB2 pantoB100*	This study.
ΔFurD ΔFurC pyrG89	*ΔfurD::riboB ΔfurC::riboB pyrG89 riboB2 pyroA4 pantoB100 biA1*	This study.
ΔFurD ΔFurE	*ΔfurD::riboB ΔfurE::riboB riboB2 biA1*	This study.
ΔFurD ΔFurE pyrG89	*ΔfurD::riboB ΔfurE::riboB pyrG89 riboB2 pyroA4 pantoB100 biA1*	This study.
ΔFurD ΔFurF	*ΔfurD::riboB ΔfurF::riboB riboB2 biA1*	This study.
ΔFurD ΔFurF pyrG89	*ΔfurD::riboB ΔfurF::riboB pyrG89 riboB2 pyroA4 yA2*	This study.
ΔFurD ΔFurG	*ΔfurD::riboB ΔfurG::riboB riboB2 pantoB100*	This study.
ΔFurD ΔFurG	*ΔfurD::riboB ΔfurG::riboB pyrG89 riboB2 pyroA4*	This study.

pyrG89		
ΔFurD ΔFurA ΔFurE	*ΔfurD::riboB ΔfurA::riboB ΔfurE::riboB riboB2 pantoB100 biA1*	This study.
ΔFurD ΔFurA ΔFurE pyrG89	*ΔfurD::riboB ΔfurA::riboB ΔfurE::riboB riboB2 pyroA4 biA1pyrG89*	This study.
ΔFurA ΔFurE	*ΔfurA::riboB ΔfurE::riboB riboB2 pantoB100 biA1*	This study.
ΔFurA ΔFurE pyrG89	*ΔfurA::riboB ΔfurE::riboB pyrG89 riboB2 pantoB100 biA1*	This study.
FurA-GFP Δ7	*pGEM-gpdA$_p$-furA-gfp-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurB Δ7	*pGEM-gpdA$_p$-furB-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurB-GFP Δ7	*pGEM-gpdA$_p$-furB-gfp-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurC Δ7	*pGEM-gpdA$_p$-furC-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurC-GFP Δ7	*pGEM-gpdA$_p$-furC-gfp-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD Δ7	*pGEM-gpdA$_p$-furD-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD-GFP Δ7	*pGEM-gpdA$_p$-furD-gfp-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurE Δ7	*pGEM-gpdA$_p$-furE-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurE-GFP Δ7	*pGEM-gpdA$_p$-furE-gfp-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurF Δ7	*pGEM-gpdA$_p$-furF-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurF-GFP-Δ7	*pGEM-gpdA$_p$-furF-gfp-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurG-GFP Δ7	*pGEM-gpdA$_p$-furG-gfp-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/N54A Δ7	*pGEM-gpdA$_p$-furD/N54A-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/T57A Δ7	*pGEM-gpdA$_p$-furD/T57A-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/W130A Δ7	*pGEM-gpdA$_p$-furD/W130A-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1*	This study.

	pantoB100	
FurD/Q134A Δ7	*pGEM-gpdA$_p$-furD/Q134A-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/Q134E Δ7	*pGEM-gpdA$_p$-furD/Q134E-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/N249A Δ7	*pGEM-gpdA$_p$-furD/N249A-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/Y250A Δ7	*pGEM-gpdA$_p$-furD/Y250A-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/Y250K Δ7	*pGEM-gpdA$_p$-furD/Y250K-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/Y250F Δ7	*pGEM-gpdA$_p$-furD/Y250F-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/Y250W Δ7	*pGEM-gpdA$_p$-furD/Y250W-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/N341A Δ7	*pGEM-gpdA$_p$-furD/N341A-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/L386A Δ7	*pGEM-gpdA$_p$-furD/L386A-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/LEN Δ7	*pGEM-gpdA$_p$-furD/L386-E-N387-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/M389A Δ7	*pGEM-gpdA$_p$-furD/M389A-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/L386S Δ7	*pGEM-gpdA$_p$-furD/L386S-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurD/M389L Δ7	*pGEM-gpdA$_p$-furD/M389L-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurE/K252F Δ7	*pGEM-gpdA$_p$-furE/K252F-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurE/P251N Δ7	*pGEM-gpdA$_p$-furE/P251N-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurE/P251N/ K252F Δ7	*pGEM-gpdA$_p$-furE/P251N/K252F-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.

FurE/S54N/S56A/M57T Δ7	*pGEM-gpdA$_p$-furE/S54N/S56A/M57T-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
FurE/S54N/S56N/M57T Δ7	*pGEM-gpdA$_p$-furE/S54N/S56N/M57T-pantoB ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*	This study.
Aspergillus brasiliensis	CBS101740 Fungal Biodiversity Centre Institute of the Royal Netherlands Academy of Arts and Sciences (KNAW)	
DH5α *E.coli*	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	New England Biolabs

2.2 Growth media and growth conditions

The *A. nidulans* and *A. brasiliensis* growth media for cultures used during the current study are described in **Table 3** and **Table 4**. Media and supplemented vitamin concentrations are as in *www.fgsc.net*.

Table 3. Growth media for *A. nidulans* and *A. brasiliensis.*

	Complete Medium (CM)	Minimal Medium (MM)	Sucrose Minimal Medium (SM)
Salt solution I (50x)	20 mL	20 mL	20 mL
Vitamin solution (100x)	10 mL	-	-
D-glucose	10 g	10 g	10 g
Peptone	2 g	-	-
Yeast extract	1 g	-	-
Cas-amino acids	1 g	-	-
Sucrose	-	-	342.3 g
Distilled H_2O	up to 1 L	up to 1 L	up to 1 L
pH (addition of NaOH)	6.8	6.8	6.8

*For cultures on solid media 1-2% (w/v) Agar was added.

Table 4. Recipes for salt, vitamin and trace elements solutions.

Salt solution (50x)		Vitamin solution (100x)		Trace elements (20x)	
KCl	26 g	**p-aminobenzoic acid**	20 mg	**$Na_2B_4O_7 \cdot 10\ H_2O$**	40 mg
$MgSO_4 \cdot 7H_2O$	26 g	**Biotin**	1 mg	**$CuSO_4 \cdot 5H_2O$**	400 mg
KH_2PO_4	76 g	**Calcium-D-pantothenate**	50 mg	**$FeO_4P \cdot 4H_2O$**	714 mg
				$MnSO_4 \cdot 1H_2O$	728 mg
Chloroform	2 mL	**Riboflavin**	50 mg	**$Na_2MoO_4 \cdot 2H_2O$**	800 mg
Trace elements (20x)	50 mL	**Pyridoxine**	50 mg	**$ZnSO_4 \cdot 7H_2O$**	8 mg
Distilled H_2O	up to 1L	**Distilled H_2O**	up to 1 L	**Distilled H_2O**	up to 1 L

Minimal medium was used as a selective growth medium for detection of distinct phenotypes in the presence of various purines (usually at a final concentration of ~500 μM) as sole nitrogen sources or sensitivity/resistance on different purine/ pyrimidine/nucleoside analogues. Urea (5 mM), $NaNO_3$ (10mM) or NH_4-L-tartrate (10 mM) were used as neutral nitrogen sources. Growth tests on toxic analogues were performed in the presence of urea or $NaNO_3$ as a nitrogen source. Vitamin auxotrophies were supplemented in each occasion as required.

For bacterial cultures Luria Bertani (LB) medium (Bacto Tryptone 1% w/v, NaCl 1% w/v, BactoYeast Extract 0.5% w/v, pH 7.0) was used. For solid cultures 1.5% w/v Agar was added as well. Ampicillin was added at a final concentration 100 μg/mL.

All media were sterilized by autoclaving or filtration before use.

The growth conditions were usually 25^{o}C or 37^{o}C for *A. nidulans* and *A. brasiliensis* and 37^{o}C for *E. coli*.

2.3 Genetic crossing of *A. nidulans* strains

Genetic crossing of strains was performed based on the protocol published by Todd et al., 2007. In more detail, two strains were selected in each case with at least one non-mutual auxotrophy and were inoculated on a CM-agar plate in close proximity, and then incubated for 2-3 days at 37^{o}C or 25^{o}C or until the two colonies met. An agar piece containing the mixed hyphae was transferred on a new plate (30 mm diameter) with MM-agar supplemented with $NaNO_3$ and any mutual vitamin auxotrophy, again, incubation followed for 2-3 days at 37^{o}C or 25^{o}C or until the heterocaryotic hyphae appeared. Then, the plate was tightly sealed and incubated for at least 2 weeks at 37^{o}C or 25^{o}C. The fruiting bodies, cleistothecia, appeared as black spheres and were identified with a stereoscope and isolated. Hülle cells, aerial hyphae and any conidiophores were removed from the cleistothecia by rolling on an agar plate. The "clean" cleistothecium was disrupted against the walls of a centrifuge tube containing 1 mL sterile dH_2O, in order for the asci to be released. The asci were disrupted by vortexing to generate an ascospore suspension; a small volume from the suspension was spread on a MM-agar plate containing only the mutual auxotrophies in order to separate the selfed cleistothecia from the crossed ones. Those proven crossed were spread on CM or MM media and the colonies arisen were used for further analysis by growth testing and PCR, when necessary, to identify the progeny of interest in each case.

2.4 DNA extraction from *A. nidulans* and *A. brasiliensis*

DNA extraction was performed as described in *www.fgsc.net*. In more detail, the strain of interest was inoculated on a CM-agar plate and incubated at 37^{o}C for 3-4

days. Conidiospores from a colony sized one quarter of the plate were transferred in an Erlenmeyer flask containing liquid MM supplemented with NH_4^+ or $NaNO_3$, and the required auxotrophies, the culture was incubated overnight at 25°C or 37°C in a shaking incubator at 140 rpm, until full growth and formation of mycelia.

The mycelia were collected by a Blutex filter, dried with paper towels and frozen in liquid nitrogen. Mortar and pestle were used in constant presence of liquid nitrogen to convert the frozen mycelia to a fine powder. 100-200 mg of the powder were then transferred to a centrifuge tube and 800 μL extraction buffer (Tris-HCl 0.2 M pH 8, SDS 1% w/v, EDTA 1 mM pH 8) were added. The mixture was vortexed until homogenization occurred and followed by 20 min incubation on ice. Then, 800 μL of basic phenol were added and thorough vortexing was performed for protein denaturation to occur, and centrifuged at 13,000 g for 5 min. After the centrifugation, two liquid phases were formed, one organic containing phenol, cell wall debris and proteins and one aqueous containing buffer, DNA, RNA and some proteins. The aqueous phase was transferred carefully in a new centrifuge tube avoiding any phenol traces. Then, equal volume of chloroform was added; thorough vortexing was performed again and the sample was centrifuged at 13,000 g for 5 min. As previously, two phases were formed after the centrifugation, one organic containing the chloroform and traces of phenol and one aqueous containing DNA, RNA and some proteins. The aqueous phase (V μL) was transferred in a new centrifuge tube.

Then, the DNA was precipitated by ethanol. In more detail, V μL isopropanol and 1/10 V μL CH_3COONa were added, followed by mild agitation and centrifugation at 13,000 g for 5 min. The supernatant was discarded and 200 μL of 70% v/v EtOH were added to wash the pellet from salts, centrifugation at 13,000 g for 2 min followed. The supernatant again was discarded and the pellet was airdried at RT or at 50°C for 10 min. The pellet was resuspended in 50 μL of dH_2O containing RNase (100 μg/mL) and incubated for 30 min at 37°C. The quality and concentration of DNA were determined by agarose gel electrophoresis or spectrophotometry at 260 nm and 280 nm (OD_{260}/OD_{280}= ~1.8 corresponds to pure DNA).

2.5 Plasmid DNA extraction from *E. coli*

Plasmid DNA preparation from *E. coli* was performed with a plasmid DNA extraction kit by Macherey-Nagel (NucleoSpin® Plasmid) according to the manufacturer's instructions or with a modified protocol based on Chapters 1.34-1.37 in (Sambrook et al., 1989) described below.

In more detail, 1 mL of saturated liquid bacterial culture was collected in a centrifuge tube and centrifuged at 13,000 g for 1 min. After the supernatant was discarded, 200 μL of resuspension solution (Tris-HCl 50 mM pH 8, EDTA 10 mM) were added, followed by vortexing until the pellet was completely resuspended. 200 μL of lysis solution (NaOH 200 mM, SDS 1% w/v) were added and mildly agitated. In this step the cell wall and membrane were disrupted and because of the basic pH the DNA (both genomic and plasmid) was denaturated. The nucleases function was

inhibited by the lack of free cations (Ca^{2+}, Mg^{2+}) in the solution because of the chelating reagent EDTA. The addition of 200 μL neutralizing solution (CH_3COONa 3 M pH 5.5) followed by mild agitation and led to protein denaturation, as well as, abrupt change of the pH leading to forcible formation of double stranded regions in DNA. At this step only small length DNAs (e.g. plasmids) are renatured properly. Centrifugation at 13,000 g for 5 min was then performed.

The supernatant was transferred in a new centrifuge tube and 500 μL of isopropanol were added, mild agitation and centrifugation at 13,000 g for 5 min followed. The plasmid DNA precipitated and the supernatant was removed. The pellet was washed with 200 μL 70% v/v EtOH and centrifuged at 13,000 g for 2 min. Finally, the supernatant was removed and the pellet was left to dry. The pellet was resuspended in 50 μL of dH_2O containing RNase 100 μg/mL and incubated for 30 min at 37^{o}C. The quality and concentration of DNA were determined by agarose gel electrophoresis or spectrophotometry at 260 nm and 280 nm (OD_{260}/OD_{280}= ~1.8 corresponds to pure DNA).

2.6 PCR, Site Directed Mutagenesis and Oligonucleotides

PCR was performed in order to amplify specific regions of DNA for diagnostic or cloning purposes. Proof reading polymerase KAPA HiFi (Kapa Biosystems) was used for reactions needing accuracy and KAPA Taq (Kapa Biosystems) for regular reactions. The reactions were performed according to the manufacturer's instructions, the T_m of the primers was calculated according to the formula T_m= 69.3 + 0.41 (*GC**100/*L*) – 650/*L*, where *GC* corresponds to the number of GC oligonucleotides that anneal to the target DNA and *L* the oligonucleotide length.

Specific amino acid substitutions on proteins were performed by PCR site directed mutagenesis on the corresponding ORF. KAPA Hifi was used for the mutagenesis and the protocol followed was according to the instructions of the manufacturer. Here follows a brief description of the protocol.

For every mutation, two complementary oligonucleotides that anneal at the region of interest of the template were designed. The required mutation was introduced in the middle of both oligonucleotides. To calculate the T_m, the previous formula is transformed to T_m= [69.3 + 0.41 (GC*100/L) – 650/L] – 5 – X, where X is every mismatch included in the oligonucleotide. PCR amplification was carried in 50 μL reaction with 1 ng of template DNA according to the instructions of the manufacturer.

One μL of the PCR product was run in an agarose gel next to the same amount of template used in the PCR reaction. The amplification was considered successful when the band corresponding to the PCR product was of higher intensity. Next step was a DpnI digestion of the PCR product, in order to digest the parental DNA strands (DpnI recognizes a methylated 4-nucleotide sequence target). Then, DpnI was deactivated by heating at 80^{o}C for 20 min. Part of DpnI-treated product was used for transformation of *E. coli* competent cells. Single colonies arisen from the

transformation where cultivated in liquid cultures, the plasmids were isolated and the mutation was verified by sequencing. Oligonucleotides used in this study are included in

Table 5.

Table 5. Forward primers used for site directed mutagenesis (reverse primers were complementary to the forward primers) and primers for cloning purposes. The nucleobase substitutions of the wild-type codons and restriction enzyme linkers are shown in bold. F stands for forward and R for reverse.

	Oligonucleotide name	Sequence 5' – 3'
1	5' uapA AatII F	CGCG**GACGTC**GGATCTGCGAGATCCTGGGGTC
2	3' uapA NotI R	CGCG**GCGGCCGC**CGCCCAATAATCTTCCCCCTCAG
3	fcyB ORF BamHI F	CG**GGATCC**CGCATTCGATTTCGACCTG
4	fcyB ORF XbaI R	CG**TCTAGA**TTACCTGCCAAAAAACTTGATC
5	fcyB ORF NS XbaI R	CGCG**TCTAGA**CCTGCCAAAAAACTTGATCTCAAAAGGCC
6	GFP XbaI F	CG**TCTAGA**ATGGTGAGCAAGGGCGAG
7	GFP SpeI R	CG**ACTAGT**TTACTTGTACAGCTCGTCCATG
8	GFP BamHI F	GCGC**GGATCC**ATGGTGAGCAAGGGCGAGG
9	GFP NS BamHI R	GCGC**GGATCC**CTTGTACAGCTCGTCCATGCC
10	AFriboB NdeI F	CGCG**CATATG**AAGCTTGATATCACAATCAGC
11	AFriboB NdeI R	CGCG**CATATG**CCCGGGCTGCAGGAATTCGATAAG
12	FcyB/N81A F	CAGCATGTGGCTGGCGGCC**GC**CATGGTCGTCAGTTCCTTTGCC
13	FcyB/N81S F	CAGCATGTGGCTGGCGGCC**AG**CATGGTCGTCAGTTCCTTTGCC
14	FcyB/V83N F	GGCTGGCGGCCAACATG**AA**CGT**T**AGTTCCTTTGCC
15	FcyB/S85A F	CGGCCAACATGGTCGTC**GC**TTCCTTTGCCATCGGTG
16	FcyB/W159A F	CTCAATATTCTTGCATGCTTGGG**CGCC**TCTGCTGCCAACGCCATCG
17	FcyB/N163Q F	GGGTTGGTCTGCTGC**GC**A**G**GCCATCGTAGGCGCTC
18	FcyB/N163L F	GGGTTGGTCTGCTGCC**CT**CGC**G**ATCGTAGGCGCTC
19	FcyB/T191A F	CGATCTTGATCATTTCCATTTGC**G**C**A**CTTTTGGTCACATTTGCGGG
20	FcyB/E206A F	GTGGTCCATTTGTATG**C**ATACTGGAGTTGGATTCCC
21	FcyB/W259A F	CGGCTTCGCTACGGGC**GCC**ACTAGTTACGCAGCCG
22	FcyB/S261A F	CGCTACGGGCTGGACT**GC**TTACGCAGCCGATTACAC
23	FcyB/N350Q F	GCCCTCTCAATCGTTGC**GC**A**G**AACTGCCCGAATTTCTAC
24	FcyB/N350A F	GCCCTCTCAATCGTTGC**GGCG**AACTGCCCGAATTTCTAC
25	FcyB/N350D F	GCCCTCTCAATCGTTGC**GG**ACAACTGCCCGAATTTCTAC
26	FcyB/N351A F	CCTCTCAATCGTTGCAAAC**GCA**TGCCCGAATTTCTACTCC
27	FcyB/P353A F	CGTTGCAAACAACTGC**G**C**A**AATTTCTACTCCGTGGCTC
28	FcyB/N354D F	CGTTGCAAACAACTGCCCG**G**ATTTCTACTCCGTGGCTCTC
29	FcyB/N354A F	CGTTGCAAACAACTGCCCG**GC**TTTCTACTCCGTGGCTCTC
30	FcyB/N354Q F	CGTTGCAAACAACTGCCCG**C**A**G**TTCTACTCCGTGGCTCTC
31	argB SalI FOR	CCG**GTCGAC**TGGGGTAGTCATCTAATG
32	argB SalI REV	CCG**GTCGAC**GTCGACCTACAGCCATTG
33	azgA BamHI F	CGG**GGATCC**GCATGCGCCTCGATCGTTG
34	azgA BamHI R	CGG**GGATCC**GTGATAATTCTTTATATTTGCAGAG
35	azgA stop destr XbaI F	GGAACAGGTCACGGAGAAATA**TCTAGA**GTGACCGGTATAGTATAGACAG
36	azgA stop destr XbaI R	CTGTCTATACTATACCGGTCAC**TCTTAGA**TATTTCTCCGTGACCTGTTCC
37	pBS XbaI destr F	GCTCCACCGCGGTGGCGGCCGCACTAGAACTAGTGGATC
38	pBS XbaI destr F	GATCCACTAGTTCTAGTGCGGCCGCCACCGCGGTGGAGC
39	gfp XbaI F	CG**TCTAGA**CATGGTGAGCAAGGGCGAG
40	gfp XbaI R	CG**TCTAGA**TTACTTGTACAGCTCGTCC

41	AzgA/T49A F	TGCCGGCCTTGCC**G**CCTTCTTCGCTATG
42	AzgA/T49S F	TGCCGGCCTTGCC**T**CCTTCTTCGCTATG
43	AzgA/G129A F	GCCCCGGGAATGG**C**TCTCAACGCGTACTTTG
44	AzgA/G129F F	CTGGCCCCGGGAATG**TT**TCTCAACGCGTACTTTGC
45	AzgA/N131A F	CGGGAATGGGTCTC**GC**CGCGTACTTTGCCTATA
46	AzgA/N131D F	CGGGAATGGGTCTC**G**ACGCGTACTTTGCCTATA
47	AzgA/F335A F	CTCGCATTGATCACG**GC**CTTGGTATGCTATCTAGC
48	AzgA/D339A F	GGATAGTACGTCG**C**CATTCTCGACGCTACG
49	AzgA/D339N F	GGATAGTACGTC**A**ACATTCTCGACGCTACG
50	AzgA/D339E F	GGATAGTACGTCGA**G**ATTCTCGACGCTACG
51	AzgA/D342A F	CGTCGACATTCTCG**C**CGCTACGGGTACAT
52	AzgA/D342E F	CGTCGACATTCTCGA**G**GCTACGGGTACAT
53	AzgA/D342N F	CGTCGACATTCTC**A**ACGCTACGGGTACAT
54	AzgA/T346A F	CGACGCTACGGGT**G**CATTATACTCAATGGCC
55	AzgA/T390A F	GGTTCTCCGCCTGTT**G**CAGCATTCGTCGAGAG
56	AzgA/E394A F	GTTACAGCATTCGTCG**CT**AGCGGTGCTGGTAT
57	AzgA/E394Q F	GTTACAGCATTCGTC**C**AGAGCGGTGCTGGTAT
58	AzgA/E394D F	CTGTTACAGCATTCGTCGA**C**AGCGGTGCTGGTATTTC
59	AzgA/S395A F	CAGCATTCGTCGAG**GC**CGGTGCTGGTATTTC
60	AzgA/S395N F	CAGCATTCGTCGAGA**A**CGGTGCTGGTATTTC
61	AzgA/G398A F	CGTCGAGAGCGGTGCTG**C**TATTTCGGAAGGTGG
62	AzgA/G402A F	GCTGGTATTTCGGAAG**C**TGGAAAGACCGGTCTGAC
63	AzgA/K404R F	GTATTTCGGAAGGTGGAA**G**GACCGGTCTGACATCATG
64	AzgA/D478A F	CTACAGCATTGCCG**C**CGGCCTGATCGCCG
65	pantoB NdeI F	GCG**CATATG**GGATTGGGAGGTGACACAGAACTTCC
66	pantoB NdeI R	GCG**CATATG**CACGCGAGGATTAACATGCGTGG
67	pGPDA AatII F	GCG**GACGTC**GGTTGACCGGTGCCTGGATC
68	pGPDA SpeI R	GCG**ACTAGT**TGTGATGTCTGCTCAAGCGGGG
69	trpC NotI F	GCG**GCGGCCGC**GATCCACTTAACGTTACTGAAATC
70	trpC PstI R	GCG**CTGCAG**CCTCTAAACAAGTGTACCTGTGC
71	furA SpeI F	GCG**ACTAGT**ATGTCAGCTATTAAACGATGGATC
72	furB SpeI F	GCG**ACTAGT**ATGAAAATCACCACAGCACTAAAGAAG
73	furB NotI R	GCG**GCGGCCGC**TCATGCTCCTTTGGCTGAATATAC
74	furC SpeI F	GCG**ACTAGT**ATGGACCGCCTCTCCATCAG
75	furC NotI R	GCG**GCGGCCGC**TCAAACCTCCCTAACTCCAACC
76	furD SpeI F	GCG**ACTAGT**ATGCGTTTCGGTCGCTTTCACC
77	furD NotI R	GCG**GCGGCCGC**TCAGTAAACAGCAAAACCCTTCTCC
78	furE SpeI F	GCG**ACTAGT**ATGGGACTACGAGAAAGACTCC
79	furE NotI R	GCG**GCGGCCGC**CTATGCAGAGACAGCCTCCTTC
80	furF SpeI F	GCG**ACTAGT**ATGGCACTCCGGCGTTTTATAC
81	furF NotI R	GCG**GCGGCCGC**TCAAACACTGATCTTATCGCCTTTC
82	furG SpeI F	GCG**ACTAGT**ATGAAGACTTTGTCGTGGCCC
83	furA NotI NS R	GCG**GCGGCCGC**TACACTAGCGGATCCAGCC
84	furB NotI NS R	GCG**GCGGCCGC**TGCTCCTTTGGCTGAATATACC
85	furC NotI NS R	GCG**GCGGCCGC**AACCTCCCTAACTCCAACCTC
86	furE NotI NS R	GCG**GCGGCCGC**TGCAGAGACAGCCTCCTTC
87	furF NotI NS R	GCG**GCGGCCGC**AACACTGATCTTATCGCCTTTCTC
88	furG NotI NS R	GCG**GCGGCCGC**GACGACGGAAACCTCTGCC
89	AbfurB2 SpeI F	GCG**ACTAGT**ATGACATTACTTCAATGGCTCGAGG
90	AbfurB2 NotI R	GCG**GCGGCCGC**GTCATCCTCCACACCCTTTGG
91	GFP NotI F	GCG**GCGGCCGCC**ATGGTGAGCAAGGGCGAGG

92	GFP NotI R	CGC**GCGGCCGC**TTACTTGTACAGCTCGTCC
93	FurD N54A F	CTCTGACGCCTTC**GC**CGCTGCGACCTGGC
94	FurD T57A F	GCCTTCAACGCTGCG**G**CCTGGCAATTCGCC
95	FurD W130A F	CCTGGCCATTTTC**GC**GTTTGCGATCCAGAATG
96	FurD Q134A F	CTGGTTTGCGATC**GC**GAATGTCAACGGTGC
97	FurD Q134E F	CTGGTTTGCGATC**G**AGAATGTCAACGGTGC
98	FurD N249A F	GAGTGTGCTGGGG**GC**TTATGCCACGCTTAGTG
99	FurD Y250A F	GTGTGCTGGGGAAT**GC**TGCCACGCTTAGTGTTAAC
100	FurD Y250K F	GTGTGCTGGGGAAT**A**AA**A**GCCACGCTTAGTGTTAAC
101	FurD Y250W F	GTGTGCTGGGGAATT**GG**GCCACGCTTAGTGTTAAC
102	FurD Y250F F	GTGTGCTGGGGAATT**T**TGCCACGCTTAGTGTTAAC
103	FurD N341A F	GTGAACATCAGCGCG**GC**CTCGATTTCCGCTGC
104	FurD L386A F	GAGTGCCAGCAATTTC**GCC**AACTTCATGTCTGC
105	FurD L386S F	GTGCCAGCAATTTC**TC**GAACTTCATGTCTGC
106	FurD M389A F	GCAATTTCCTGAACTTC**GCC**TCTGCATATGCC
107	FurD M389L F	GCAATTTCCTGAACTTC**C**TGTCTGCATATGCC
108	FurD L^{386}EN^{387} F	GAGTGCCAGCAATTTCCTG**GAG**AACTTCATGTCTGC
109	FurE P251N F	CTCGGCCATCGGG**AAC**AAGGCGACGCTGGCC
110	FurE K252F F	GGCCATCGGGCCG**TTC**GCGACGCTGGCC
111	FurE P251K252/NF F	GTTACCTCGGCCATCGGG**AACTTC**GCGACGCTGGCCCTGAAC
112	FurES54N/S56N/M57T F	CGCCGAGGCGTTTA**A**TATTA**A**CA**C**GTATCAAGGTACTGTCCGC
113	FurES54N/S56A/M57T F	CGCCGAGGCGTTTA**A**TATT**GC**CA**C**GTATCAAGGTACTGTCCGC
114	5’ FurC F	GTAGATGGCAGAGGGGTAAC
115	5’ FurE F	CCAACTGACGAACCTGCC
116	5’ FurF F	GCTACGACTGCAACGACTTC
117	5’ FurG F	CCACACCGAAACTAACTGCG
118	riboB R	CCGAATGGTGGATTTAGAGCC
119	AbuapC BglII F	GCG**AGATCT**ACCTGACCAGATAGGCCCG
120	AbuapC SpeI R	GCG**ACTAGT**GCACTCGAGGGCTCGC
121	5’UTR PstI R	GCG**CTGCAG**GCTGTCGTCACTCGAATAGG
122	AbuapD PstI F	GCG**CTGCAG**ATGGAGTCCAAGTCTAGCCC
123	AbuapD XbaI R	GCG**TCTAGA**TCAAGGCAACACTGTATCCACCTTC
124	AbuapD NS XbaI R	GCG**TCTAGA**GGCAACACTGTATCCACCTTC
125	AbuapD-NO-CT PstI R	CGC**CTGCAG**GCAACAACAAATTCAACAACACCCC
126	uapAct PstI F	GCG**CTGCAG**AAGTCGAAGAGATCGGTG
127	AbuapD ADA PstI F	GCGC**CTGCAG**CAACAGGCTACTGGTCTCGC
128	AbuapD ADA PstI R	GCGC**CTGCAG**CGAATTTACCGAAGATGCCC
130	GFP PstI F	CGC**CTGCAG**ATGGTGAGCAAGGGCGAGG
131	GFP NS PstI R	CGC**CTGCAG**CTTGTACAGCTCGTCCATGCC
132	UapA NS XbaI R	GC**TCTAGA**GCCTGCTTGCTCTGATACTC

2.7 Agarose gel electrophoresis

DNA was run in agarose gels in order to separate a mixed population of DNA after digestion or PCR. The fragments are separated by length by applying an electrical field to move the charged molecules through an agarose matrix. The negative charged biomolecules move towards the positively-charged anode during the electrophoresis.

In order to perform an agarose gel electrophoresis we used the following protocol, modified version of Sambrook et al. (1989), 0.8-2% w/v agarose was added in TAE 1x buffer (Tris-acetate 0.04 M, EDTA 0.001 M) and heated until completely dissolved. The mixture was left to cool down for a while and ethidium bromide (which intercalates into the major grooves of the DNA/RNA and fluoresces under UV light) was added to a final concentration of 0.5 μg/mL and finally it was poured in a cast. A comb was placed in the cast to create wells for sample loading and the gel was completely set before use. The concentration of agarose in the gel affects the resolution of DNA separation. For a standard agarose gel electrophoresis, a concentration of agarose 0.8% w/v gives good resolution of large 5–10 kb DNA fragments, while 2% w/v gel gives good resolution for small 0.2–1 kb fragments. 1% gels are common for many applications.

Once the gel had set, the comb was removed and the gel was placed in the electrophoresis tank and was filled up with TAE 1x buffer until the gel was covered, a representation of the apparatus is shown in **Figure 22**. The samples were prepared with the addition of a loading buffer that consisted of a colored dye such as xylene cyanol or bromophenol blue used to estimate the migration of DNA and glycerol to raise the density of the sample, so that it sank to the bottom of the well. An electrical field was applied on the gels; the voltage applied depended on the resolution needed. The visualization of DNA was performed under UV light.

Specific bands needed for cloning or other purposes where cut from the gel and the DNA was extracted with NucleoSpin® Gel and PCR Clean-up (Macherey Nagel) according to the instructions of the manufacturer.

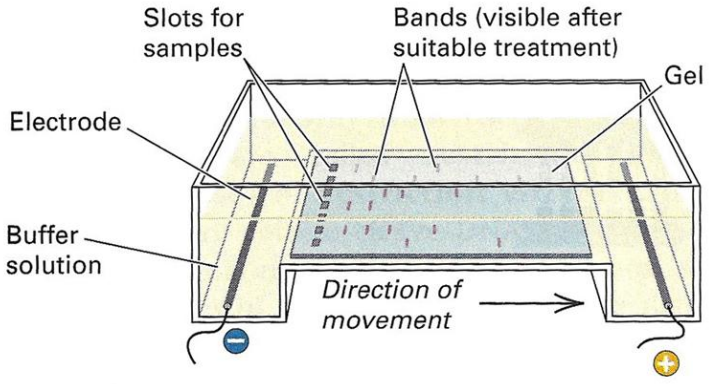


Figure 22. A schematic representation of DNA agarose gel electrophoresis. Gel is set in the middle of the electrophoresis tank and is filled up with TAE buffer. The samples are loaded in the wells and electrical field is applied at the system. The negatively charged DNA moves towards the positive pole. Picture was taken from *http://www.discoveryandinnovation.com/BIOL202/notes/lecture23.html*.

2.8 Plasmids and plasmid constructions

The plasmids used in this study are included in **Table 6**. Construction details and map representations follow for each plasmid constructed during this study. The plasmid maps were constructed using the online tool *http://www.bioinformatics.org/savvy/*.

Table 6. Plasmids used or constructed in this study.

Plasmid name	Reference
pBS-uapA$_p$-fcyB-pabaA	(Vlanti and Diallinas, 2008)
pGEM-uapA$_p$-fcyB AfriboB	(Krypotou et al., 2012)
pGEM-uapA$_p$-fcy-gfp AfriboB	(Krypotou et al., 2012)
pBS-azgA-gfp argB	(Krypotou et al., 2014)
pGEM-gpdA$_p$-furA-gfp	This study.
pGEM-gpdA$_p$-furB	This study.
pGEM-gpdA$_p$-furB-gfp	This study.
pGEM-gpdA$_p$-furC	This study.
pGEM-gpdA$_p$-furC-gfp	This study.
pGEM-gpdA$_p$-furD	This study.
pGEM-gpdA$_p$-furD-gfp	This study.
pGEM-gpdA$_p$-furE	This study.
pGEM-gpdA$_p$-furE-gfp	This study.
pGEM-gpdA$_p$-furF	This study.
pGEM-gpdA$_p$-furF-gfp	This study.
pGEM-gpdA$_p$-furG-gfp	This study.
pGEM-gpdA$_p$-AbfurB2-gfp	This study.
pBS-uapA$_p$-AbuapC-gfp-argB	This study.
pBS-uapA$_p$-AbuapD-argB	This study.
pBS-uapA$_p$-AbuapD-gfp-argB	This study.
pBS-uapA$_p$-gfp-AbuapD-argB	This study.
pBS-uapA$_p$-AbuapD-A-gfp-argB	This study.
pBS-uapA$_p$-ADA-gfp-argB	This study.

- *pGEM-uapA$_p$-fcyB AfriboB* and *pGEM-uapA$_p$-fcy-gfp AfriboB*

A modified pGEM vector was used for cloning of the *fcyB* ORF and the *5'* and *3' uapA* regulatory sequences. In more detail, the *pBS-uapA$_p$-FcyB-pabaA* plasmid (Vlanti and Diallinas, 2008) was used as a template for amplification of the *5' uapA-fcyB-3' uapA* with primers (1) and (2) and after digestion cloned in a pGEM vector. The *fcyB* ORF included *BamHI* and *XbaI* extensions. The *riboB* gene from *A.*

fumigatus genomic DNA was amplified with primers (10) and (11) with *NdeI* extensions and after digestion cloned in the same plasmid. In order to construct the *fcyB-gfp* allele, the *fcyB* ORF was digested out of the plasmid and the *fcyB*-allele without stop codon amplified with primers (3) and (5) was cloned in the same position of the plasmid. The *gfp* ORF was amplified with primers (6) and (7) and cloned in *XbaI* position of the plasmid. Detailed maps of the plasmids are shown in **Figure 23**. These plasmids were constructed by Sotiris Amillis in the J. Strauss's lab in BOKU, Vienna.

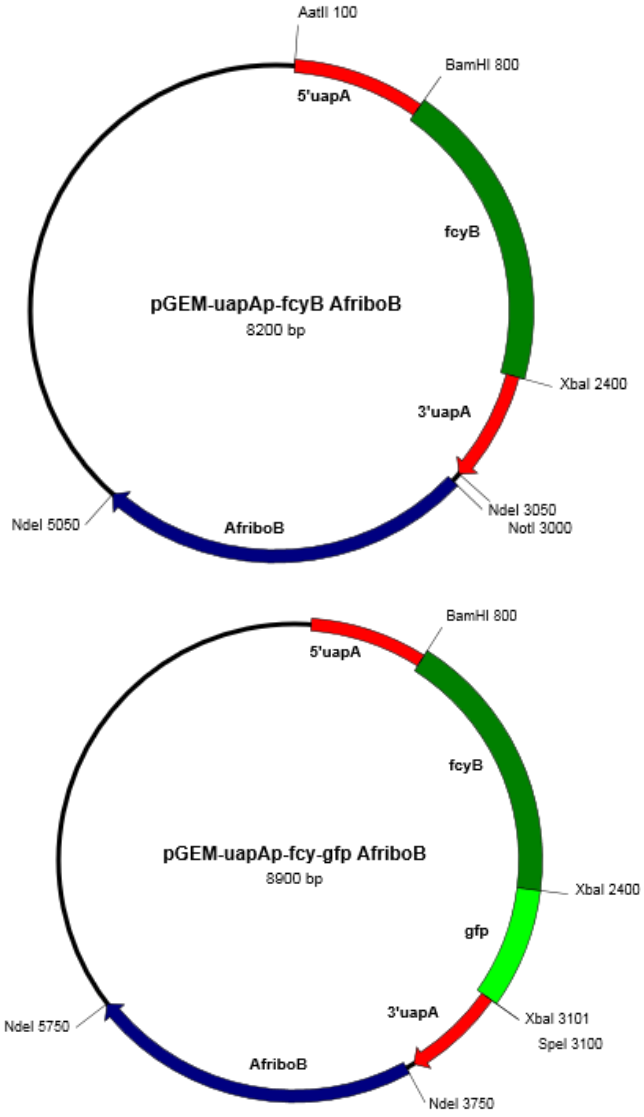


Figure 23. Plasmid maps of *pGEM-uapA$_p$-fcyB AfriboB* and *pGEM-uapA$_p$-fcyB-gfp AfriboB.* Different DNA fragments cloned in the plasmid are represented with different colours. Positions of restriction sites are also indicated.

- *pBS-azgA-gfp argB*

This plasmid is a modified *pBluescript KS*(+) vector (Stratagene) carrying the *azgA* gene fused C-terminally with the *gfp* gene, as well as, the *argB* gene as a selection

marker. The *argB* gene was cloned in the *SalI* site of the pBluescript multicloning site amplified with primers (31) and (32) and the *azgA* gene (ORF and regulatory sequences) was amplified from *A. nidulans* genome locus with primers (33) and (34) and cloned in the *BamHI* position of the multicloning site. PCR mutagenesis with primers (35) and (36) was used to destroy the *XbaI* site of the multi-cloning site and replace the *azgA* stop codon with a *XbaI* restriction site. Finally, *gfp* was amplified with *XbaI* extensions with primers (39) and (40) and cloned in frame at the end of the *azgA* ORF. A detailed map of the plasmid is shown in **Figure 24**. This plasmid was constructed by Sotiris Amillis in G. Diallinas' lab.

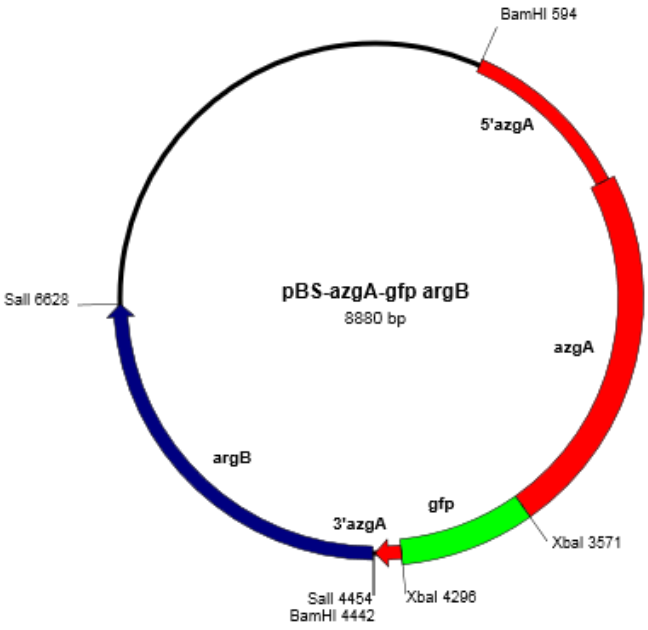


Figure 24. Plasmid map of *pBS-azgA-gfp argB.* Different DNA fragments cloned in the plasmid are represented with different colours. Positions of restriction sites are also indicated.

➢ *pGEM-gpdA$_p$-fur and pGEM-gpdA$_p$-fur-gfp*

This vector is a modified *pGEM-T-easy* vector carrying the *gpdA* promoter (~1000 bp) amplified with primers (67) and (68), cloned at *AatII* and *SpeI* restriction sites of the multicloning site (MCS) [1], the *trpC 3'* termination region amplified with primers (69) and (70), cloned at *NotI* and *PstI* restriction sites of the MCS, and the *panB* gene amplified with primers (65) and (66), cloned at the *NdeI* restriction site of the MCS. The *fur* ORFs were obtained from wild-type genomic DNA with PCR using primers (72)-(81) and cloned at the *SpeI-NotI* sites of the vector. For *fur-gfp* constructions, the *fur* ORFs lacking the translation stop codon were amplified using the reverse primers (83)-(90) and the corresponding forward primers. *gfp* ORF was amplified using

[1]Important Note: *After cloning of the gpdA promoter in pGEM vector a fraction of ~200 bp at the beginning of the promoter appears to have been lost probably due to DNA recombination events during ligation without, however, affecting the expression capacity of the promoter.*

primers (91)-(92) with *NotI* extensions and cloned in frame at the end of the *fur* ORFs. Detailed maps of the plasmids are shown in **Figure 25**.

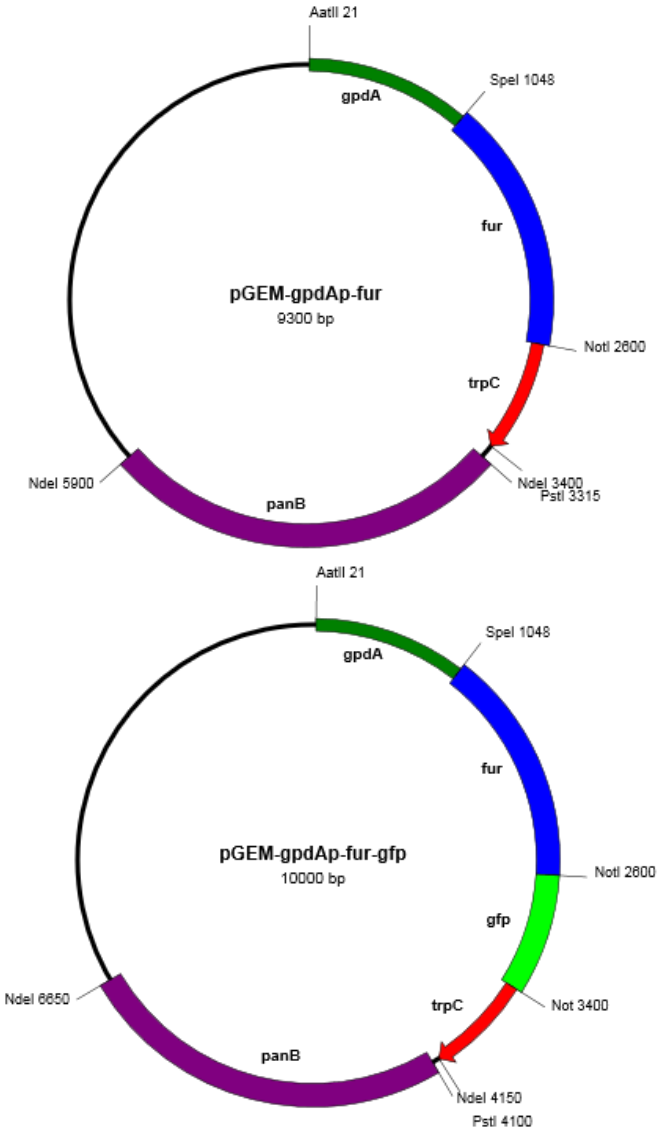


Figure 25. Plasmid maps of *pGEM-gpdA$_p$-fur and pGEM-gpdA$_p$-fur-gfp*. Different DNA fragments cloned in the plasmid are represented with different colours. Positions of restriction sites are also indicated.

- *pBS-uapA$_p$-AbuapC-gfp-argB*

This plasmid was constructed using as a template the *pAN510exp-GFP* plasmid from Papageorgiou et al. (2008), which includes the *5'* and *3' uapA* regulative sequences, the *uapA* ORF cloned with *BamHI* and *XbaI* extensions and the *argB* gene as a selective marker. The *gfp* is cloned in this vector with *XbaI-SpeI* extensions making the *3'* terminus fixed on the vector. To construct *pBS-uapA$_p$-AbuapC-gfp-argB*, *uapA* was removed after digestion with *BamHI* and *XbaI* and replaced with *AbuapC* amplified with primers (119) and (120) from *A. brasiliensis* genomic DNA with

BglII-SpeI extensions and in frame with *gfp*. Detailed map of the plasmid is shown in **Figure 26**.

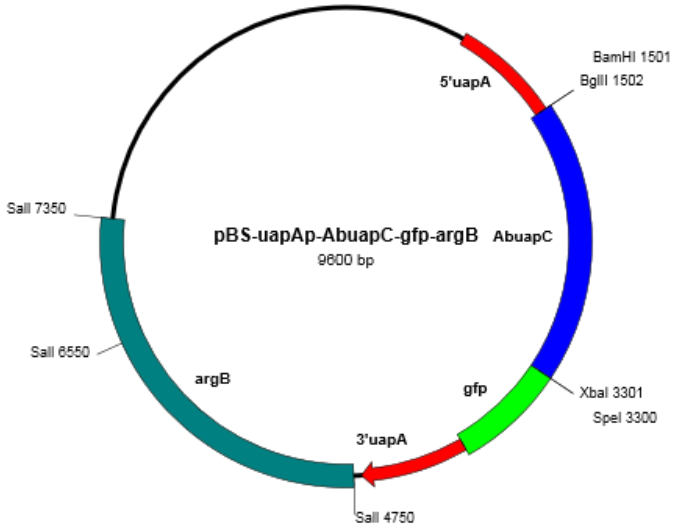


Figure 26. Plasmid map of *pBS-uapA$_p$-AbuapC-gfp-argB*. Different DNA fragments cloned in the plasmid are represented with different colours. Positions of restriction sites are also indicated.

- *pBS-uapA$_p$-AbuapD-gfp-argB, pBS-uapA$_p$-AbuapD-argB and pBS-uapA$_p$-gfp-AbuapD-argB*

In order to construct *pBS-uapA$_p$-AbuapD-gfp-argB* plasmid, the *pAN510exp-GFP* plasmid from Papageorgiou et al. (2008) was used as a template. The vector was constructed by removing the *5' uapA* regulatory sequence and the *uapA* ORF from the plasmid with *NotI-XbaI* digestion. The *5' uapA* was amplified from the *pAN510exp-GFP* plasmid with M13 forward universal primer (binds on the MCS next to *5'uapA*) and primer (121) with *PstI* extension and digested with *NotI-PstI*. The *AbuapD* ORF was amplified with primers (122) and (124) with *PstI-XbaI* extensions and without a stop codon. Both *5' uapA* and *AbuapD* were cloned in the vector with a single step ligation.

In the case of the *pBS-uapA$_p$-AbuapD-argB* plasmid, the pAN510exp (Papageorgiou et al., 2008) plasmid was used to remove with digestion the *5' uapA* regulatory sequence and the *uapA* ORF and replace it with an *5'-uapA$_p$-AbuapD* allele with a stop codon at the respective position, amplified with primers M13 forward and (123) from *pBS-uapA$_p$-AbuapD-gfp-argB*. Using the *pBS-uapA$_p$-AbuapD-argB* plasmid, *gfp* was cloned in frame at the N-terminus of *AbuapD* using primers (130) and (131) to add *PstI* extensions and remove the stop codon of *gfp* to construct the *pBS-uapA$_p$-gfp-AbuapD-argB* plasmid. Detailed maps of the plasmids are shown in **Figure 27**.

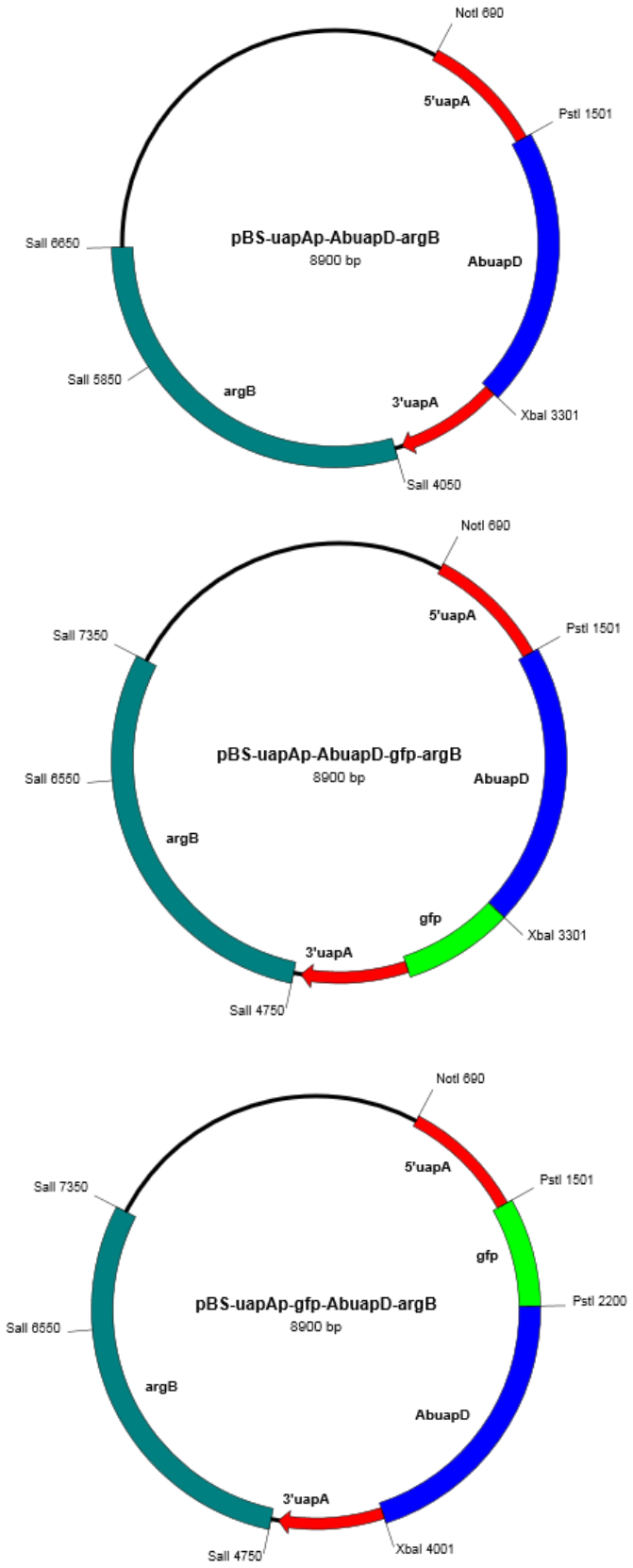
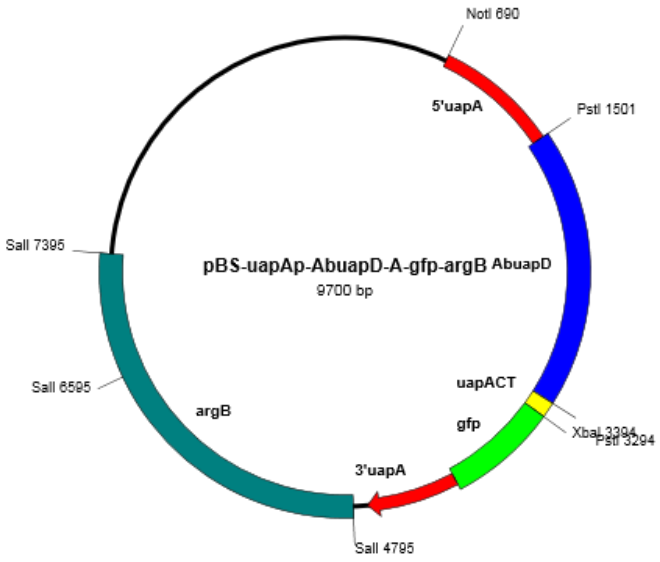


Figure 27. Plasmid maps of *pBS-uapA$_p$-AbuapD-arg*, *pBS-uapA$_p$-AbuapD-gfp-argB* and *pBS-uapA$_p$-gfp-AbuapD-argB*. Different DNA fragments cloned in the plasmid are represented with different colours. Positions of restriction sites are also indicated.

➢ *pBS-uapA$_p$-AbuapD-A-gfp-argB and pBS-uapA$_p$-ADA-gfp-argB*

Two chimeric plasmids were constructed combining parts of *AbuapD* and *uapA* ORFs. The first plasmid, *pBS-uapA$_p$-AbuapD-A-gfp-argB*, consisted of the *5'-3'* regulative sequences of the *uapA* gene and the *argB* gene as a selective marker, also included an *AbuapD* allele without the codons responsible for the last 11 predicted amino acids of AbUapD C-terminus, which were replaced by the last 32 residues of the UapA C-terminus. So, the *AbuapD* allele was digested from the *pBS-uapA$_p$-AbuapD-gfp-argB* plasmid with *PstI-XbaI*, and an *AbuapD* allele lacking its C-tail was amplified with primers (122) and (125) with *PstI* extensions at both sides, the *uapA-ctail* was amplified with primers (126) and (132) with *PstI-XbaI* extensions and both pieces were cloned with a single ligation step in the vector (orientation of *AbuapD* after ligation was checked by PCR).

The second chimeric plasmid constructed, *pBS-uapA$_p$-ADA-gfp-argB*, was a modified *pAN510exp-GFP* plasmid (Papageorgiou et al., 2008), where a part of *uapA* was replaced with a part of *AbuapD*. In particular, on protein level amino acid residues Ala^{311} to Ala^{441} of UapA were replaced with Ala^{319} to Gly^{449} from AbUapD (Gly^{449} in the new plasmid was changed to Ala). This construction was performed by removing with *PstI* digestion the respective *uapA* fragment of *pAN510exp-GFP* by taking advantage the natively existing *PstI* restriction sites in *uapA* ORF. The corresponding part of *AbuapD* was amplified with primers (127) and (128) adding *PstI* extensions and was cloned in frame with *uapA*. Detailed maps of the plasmids are shown in **Figure 28**.



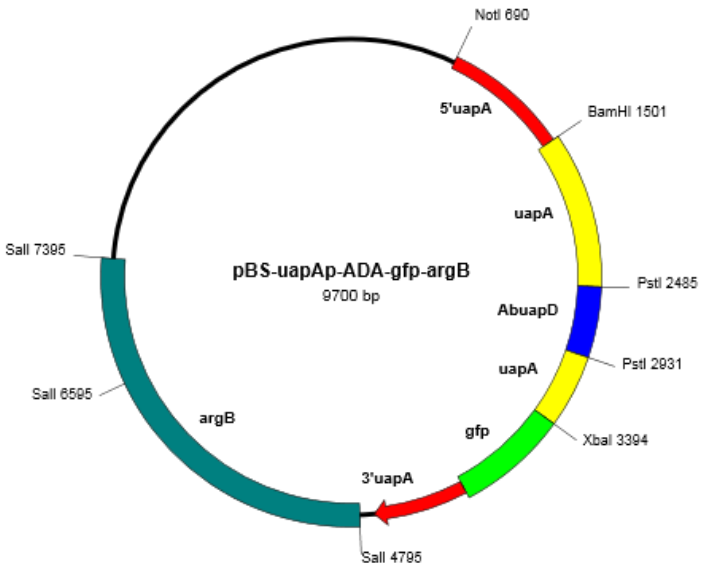


Figure 28. Plasmid maps of *pBS-uapA$_p$-AbuapD-A-gfp-argB* and *pBS-uapA$_p$-ADA-gfp-argB*. Different DNA fragments cloned in the plasmid are represented with different colours. Positions of restriction sites are also indicated.

2.9 Preparation of *E. coli* competent cells

The method used to prepare *E. coli* competent cells was a modified protocol of the one described in chapters 1.82-1.84 by Sambrook et al. (1989). In more detail, cells from a single colony of *E. coli* (DH5α) were used to inoculate 20 mL of liquid LB medium and the culture was incubated for 16 h at 37^{o}C and 220 rpm. When the culture was saturated, 2.5 mL of the culture were transferred in a 2 L Erlenmeyer flask containing 500 mL LB medium. Incubation at 37^{o}C and 220 rpm followed until the culture acquired optical density 0.45-0.55 at 600 nm.

Then, the cells were collected by centrifugation at 4,500 g for 5 min at 4^{o}C and resuspended in 100 mL ice cold solution TFB I (CH_3COOK 30 mM, $CaCl_2$ 10 mM, $MnCl_2$ 50 mM, $RbCl_2$ 100 mM, glycerol 15% v/v, pH was set at 5.8 with CH_3COOH 1 M). After incubation on ice for 5 min, the suspension was centrifuged at 4,500 g for 5 min at 4^{o}C. The supernatant was discarded and the cells were resuspended in 10 mL ice cold solution TFB II (MOPS 10 mM pH 6.5, $CaCl_2$ 75 mM, $RbCl_2$ 10 mM, glycerol 15% v/v, pH was set at 6.5 with KOH 1 M). The suspension was incubated on ice for 15-60 min. Finally, the suspension was aliquoted in centrifuge tubes and stored at -80^{o}C for at least 6 months.

It should be noted that at all steps the solutions and equipment used were sterile.

2.10 Transformation of *E. coli* competent cells

The method used to transform *E. coli* competent cells with plasmid DNA was a modified protocol based on the one described in chapters 1.74-1.81 by Sambrook et al. (1989). In more detail, in a centrifuge tube containing 200 μL competent cells, prepared as described above, 0.01-0.5 μg of plasmid DNA were added, mild agitation followed and incubation for 20-30 min on ice. Heat shock at 42^{o}C for 90 sec followed in order for the plasma membrane to become more fluid and permeable. Then, 1 mL liquid LB medium was added and after mild agitation the suspension was incubated at 37^{o}C for 1 h. This incubation period is necessary for the cells to recover from the heat shock, as well as, the bacteria to generate the antibiotic resistance proteins encoded on the plasmid backbone, in order to be able to grow once plated on the antibiotic containing LB-agar plate. The bacterial cells were collected with centrifugation and plated on LB agar containing plate supplemented with the appropriate antibiotic and incubated at 37^{o}C for 16 h.

2.11 Transformation of *A. nidulans*

Transformation of *A. nidulans* was performed using a modified protocol based on the one described by Koukaki et al. (2003). In more detail, the strain of interest was inoculated on a CM-agar plate and incubated at 37^{o}C for 3-4 days or until it reached full growth. Conidiospores were collected with a spatula in a centrifuge tube containing 5 mL Tween-80 0.01 % v/v. The suspension was well-vortexed in order for the conidiospores to get separated; then, the suspension was filtered through a Blutex into a 1 L Erlenmeyer flask containing 200 mL liquid MM supplemented with vitamins and a nitrogen source (e.g. urea) as required. The culture was incubated at 37^{o}C for about 4 h or until the conidiospores reached their isotropic phase of growth.

When the culture was ready, the conidiospores were collected by centrifugation at 3,000 g for 10 min. The cells were resuspended in solution I ($MgSO_4$ 1.2 M, Orthophosphate 10 mM pH 5.8) and 0.2 g of lysing enzymes from *Trichoderma harzianiae* were added. The lysing enzymes digest the cell wall and protoplasts are generated. The suspension was incubated at 30^{o}C for 2 h and 60 rpm. When the incubation period was complete, the cells were collected by centrifugation at 3,000 g for 10 min. Then, the supernatant was removed and the cells were resuspended in solution II (Sorbitol 1 M, $CaCl_2$ 10 mM, Tris-HCl 10 mM pH 7.5) and centrifuged again as before. The supernatant was removed and the protoplasts were resuspended in V (1-10 mL, depending on the density of the culture) solution II and 1/4V solution III (Polyethylene Glycol 4000 60% w/v, $CaCl_2$ 10 mM, Tris-HCl 10 mM pH 7.5) and mildly agitated until homogenization; 250 μL of the suspension were aliquoted in centrifuge tubes. At this step the protoplasts could be frozen at -80^{o}C after addition of 1% v/v DMSO.

At this step, 0.5-3 μg of DNA were added in the protoplasts and 15 min incubation on ice followed. Next, 500 μL of solution III were added, agitated mildly

until homogenization and incubated at RT for 15 min. Then, the protoplasts were collected by centrifugation at 5,000 g for 10 min, the supernatant was discarded and the pellet was washed (without resuspension) with 1 mL of solution II. Again, centrifugation at 5,000 g for 2 min was performed and the supernatant was discarded. Finally, the pellet was resuspended in 200 μL solution II and added in melted SM-Top agar 0.35% w/v and after mild agitation plated on an SM-agar plate. The SM-bottom was supplemented with vitamins or nitrogen sources as required, depending on the selection marker carried in the plasmid/DNA cassette; e.g. for an *argB2* auxotroph strain transformed with an *argB* carrying plasmid, the growth medium must not contain arginine. The SM plates were left to become solid and then incubated at 37^{o}C for 3-4 days or until the transformants appeared.

- *DNA molecules used for transformation and their incorporation into the protoplasts*

DNA used to transform protoplasts can vary from a linear DNA construct (cassette), to an integrative or an autonomous replicating plasmid. The autonomous replicating plasmid remains in the nucleus of the cell and is inherited in progenic cells. The inheritance of these plasmids depends on the growth conditions; if selective conditions are applied, the plasmid is inherited in all progeny and in some under non-selective conditions. Thus, this kind of plasmids is very unstable. An integrative plasmid can be integrated in the genomic DNA in non-homologous regions by one crossover junction. In the case of a cassette, the foreign DNA will be integrated in a chromosomal region of the host, where they share partial homology. In order for the cassette to be integrated, two crossover junctions are needed between the hosting DNA and the cassette, in the case of the plasmid only one is needed. The ability of the DNA homologous integration at specific genetic loci is used in targeted genetic replacements, e.g. targeting a specific locus to construct null mutants. This gene replacement technique is highly efficient by using: (i) a cassette that includes the respective flanking DNA regions (>1000 bp length of each flanking) of the target gene and (ii) *A. nidulans* strains with null *nkuA* genetic background, in which heterologous integration is genetically inactivated.

2.12 Southern blot of genomic DNA

Southern blot is used to detect specific sequences in genomic DNA. The protocol that was followed in this study was based on the one described by (Sambrook et al., 1989). This method was used to detect the plasmid copies integrated in the transformant strains constructed.

In more detail, genomic DNA of the *A. nidulans* strain of interest was isolated and quantified as described above. 3-5 μg of DNA were digested with a restriction enzyme for 12 h and after the addition of new enzyme for another 2 h. The choice of

the restriction enzyme is crucial, because there should not be a restriction site in the sequence that is going to be probed. When this is not possible the exact digestion site must be located and a proper radiolabeled probe must be constructed.

When the digestion is complete, the samples were run in a 0.8% w/v agarose gel applying 55 V electrical field. The electrophoresis period was complete when the dye band had run a distance of ¾ of the gel. Then, the gel was exposed under UV light in order to fragment the DNA.

Afterwards, the gel was immersed in denaturation solution (NaCl 1.5 M, NaOH 0.5 M) at room temperature (RT) for 30 min under mild agitation in order to denature the double-stranded DNA, and then it was immersed in neutralizing solution (NaCl 1.5 M, Tris-HCl 1 M pH 8) at RT for 30 min under mild agitation. Then, the DNA was transferred from the gel to a positively charged nylon membrane. The transferring procedure was performed by making a construction as shown in **Figure 29**. A sheet of membrane was placed on top of the gel. Pressure was applied evenly to the gel by placing a stack of paper towels and a weight on top of the membrane and gel, to ensure good and even contact between gel and membrane. The buffer used for the transfer was 20xSSC (NaCl 3 M, $Na_3C_6H_5O_7$ 0.3 M pH 7 with NaOH). Buffer transfer by capillary action from a region of high water potential to a region of low water potential (usually filter paper and paper tissues) was then used to move the DNA from the gel on to the membrane; ion exchange interactions bind the DNA to the membrane due to the negative charge of the DNA and positive charge of the membrane. DNA transfer was performed overnight and when complete the membrane was exposed under UV light for 5 min to immobilize DNA on the membrane.

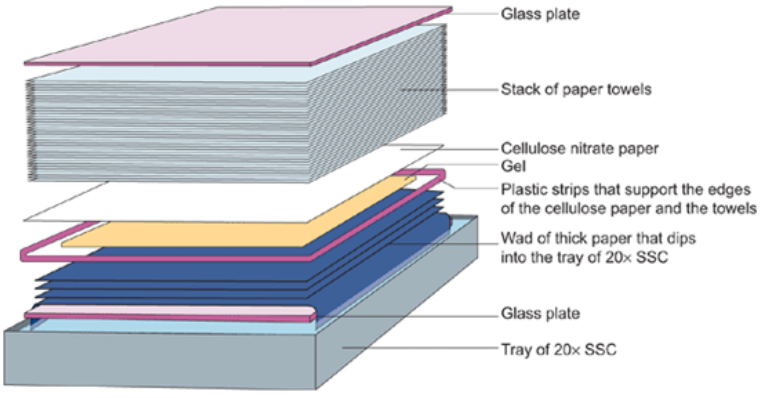


Figure 29. Southern blot DNA transfer from gel to membrane. The gel is set on top of Whatmann papers with edges soaked in 20xSSC buffer. The membrane is laid over the gel and a stack of paper towels is set on top. Finally a weight is put on the top to succeed equal distribution of the weight. (Southern, 2006)

The membrane was used for detection of specific DNA sequences with the appropriate probe in each case. The probe was constructed by synthesis of a DNA single stranded molecule containing [^{32}P] nucleotides. The probe template was usually

an ~1 kb length DNA acquired by PCR or plasmid digestion; 200-400 ng of the DNA were added in a centrifuge tube with 10 ng random hexamers and dH_2O (up to a final volume of 21 μL) and boiled for 2 min, in order the DNA fragments to become single-stranded. The sample was rapidly stored on ice to prevent formation of double stranded DNA. Then, the reaction mix was added in the sample. The reaction mix included: (*i*) all nucleotides (10 mM) (except the radiolabeled one that was the last component of the reaction added), (*ii*) the Klenow polymerase and Klenow buffer (concentrations according to the manufacturer's instructions), (*iii*) dH_2O up to 50 μL and (*iv*) the radiolabeled nucleotide. The sample was incubated at 37^{o}C for 1 h. The random hexamers bind on the single stranded DNA and the polymerase fills in the single stranded gaps with the complementary nucleotides, leading to DNA molecules that contain radiolabeled nucleotides.When the incubation was complete the mixture was added in a sepharose column (mini Quick Spin Oligo Columns, Roche) according to the manufacturer's instructions in order to get rid of the unincorporated radiolabeled nucleotides. Before use, the probe was denaturated by addition of NaOH.

Treatment of the membrane with the probe is called hybridization, during which the probe binds to complementary DNA strands bound on the membrane. Before hybridization the membrane was immersed in Church solution (NaH_2PO_4/NaH_2PO_4 0.5 M pH 7, BSA 1% w/v, SDS 7% w/v, EDTA 1 mM) and incubated at 65^{o}C for 2 h, under mild agitation. This step, also called pre-hybridization, is necessary to inhibit non-specific binding of the probe in areas without any DNA and also results for the double stranded DNA bound on the membrane to become single stranded. The temperature chosen for the hybridization step depended on the Tm of the probe; low hybridization temperature could lead to non-specific binding. In this study, 65^{o}C were used in all cases. After the pre-hybridization, the probe was added in the solution and the membrane was incubated at 65^{o}C for at least 12 h under mild agitation.

After hybridization the probe was collected in a centrifuge tube and stored at -20^{o}C. The probe could be reused after boiling for 10 min. Four washing stages of the membrane followed in order for the non-specifically bound probes to be removed. The first three washes were with 2xSSC solution and 0.1 w/v SDS for 30 min under the same conditions as the hybridization. The final wash was with 0.2xSSC and 0.1% w/v SDS for 30 min under the same conditions. Then, the membrane was wrapped with Saran wrap and set in an autoradiography cassette with an X-ray film on top. [^{32}P]-particles cause darkening of an X-ray film each time they move through it. An autoradiography cassette contains special screens on both sides impervious to [^{32}P]-particles, causing a constant periodic movement of the particles at the same point that results in band formation on the film. The cassette was incubated at -80^{o}C (low temperature increases the impermeability of the cassette screens) for period enough for the bands to appear during film development, that could vary from 20 min to 48 h, depending on the specific activity of the radiolabeled nucleotide and the probe synthesis efficiency.

2.13 Total RNA extraction from *A. nidulans*

A modified protocol for total RNA extraction was performed based on the protocol accompanying the reagent TRIzol (Invitrogen). It is crucial that during the RNA extraction procedure the samples are protected from RNases. So gloves are worn at all times, sterile equipment is used and every solution is prepared with DEPC (diethyl bicarbonate) water; DEPC is used to deactivate RNAses, it is added in water and incubated overnight at 37^{o}C and then deactivated by autoclaving.

In more detail, the strain of interest was inoculated in an Erlenmeyer flask containing liquid MM or CM supplemented as required and incubated at the appropriate temperature and period depending on the conditions tested. When the incubation was complete, the mycelium was collected and grinded with liquid nitrogen similarly to the DNA extraction protocol. When the RNA extraction was performed from conidiospores, they were collected by centrifugation and the pellet was frozen with liquid nitrogen. In a centrifuge tube, 50 μL of powder and equal volume of acid-washed glass-beads (particle-size 425-600 μm) were added and grinded with a blue tip to a fine powder. Then, 1 mL of TRIzol reagent was added and vortexed immediately for ~1 min and incubated at RT for 5 min. 200 μL of chloroform were added and after well-shaking the sample, it was incubated at RT for another 2-3 min before centrifuged at 12,000 g for 15 min at 4^{o}C. The upper aqueous phase was transferred to a new centrifuge tube and 600 μL isopropanol were added, the mixture was mixed by inversion and incubated for 10 min on ice. After centrifugation at 12,000 g for 10 min the supernatant was discarded and the pellet was washed with 1 mL 70% v/v EtOH. Then, the sample was vortexed, centrifuged at 7,500 g for 5 min at 4^{o}C, the supernatant was removed and the pellet was left to air dry for 15-20 min, but some moisture should be maintained in order not to decrease solubility. Finally, the pellet was resuspended in 50-100 μL DEPC-dH_2O and incubated at 65^{o}C for 5 min to completely dissolve RNA. The quality and concentration of the RNA were estimated by agarose gel electrophoresis and spectrophotometry at 260 nm and 280 nm (OD_{260}/OD_{280}= ~2 corresponds to pure RNA).

2.14 Northern Blot of total RNA

Total RNA was extracted as described above and analyzed by Northern Blot according to a modified protocol based on Sambrook et al. (1989). In more detail, the first step of the protocol was to denature the RNA; V μL RNA that corresponded to ~10 μg were mixed with 3V denaturation solution [DMSO (67.44% v/v), Orthophosphate 0.1 M pH 7 (13.19% v/v), glyoxal 6 M (19.35% v/v)] and incubated at 50^{o}C for 1 h, in order to eliminate the tertiary structures and the molecules to be separated in an agarose gel according to their length. Then the samples were stored on ice for at least 5 min. The RNA samples with the denaturation solution (final volume 4V μL) were mixed with 0.5V μL of loading buffer (Orthophosphate 10 mM pH 7,

Bromophenol-blue 0.2% w/v, glycerol 50% v/v). The samples were analyzed in a 1% w/v agarose gel made with orthophosphate buffer 10 mM and ran in the same buffer (TAE is not used in RNA electrophoresis). The electrophoresis was performed in an electrical field of 48 V corresponding to 48 mA. After a period of 5 min the homogeneity of the running buffer had to be reassured (because orthophosphate is dissociated during the electrophoresis) in order not to create a concentration gradient; for this reason a peristaltic pump with continuous flow was set in order to transfer solution (and apparently ions) from the one side of the gel to the other. The electrophoresis was considered complete when the dye of the loading buffer had traveled ~9 cm from the starting point.

RNAs were transferred to a positively charged nylon membrane identically as described in Southern blot protocol above. After the completion of the transfer, ribosomic RNAs (rRNAs) were detected on the membrane by immersion of the membrane in a dye (Methylene-blue 0.03% w/v, CH_3COONa 0.5 M pH 5.2) specifically binding on RNAs and discoloring with dH_2O. The abundant rRNA bands were coloured and could be used as an indication of equal loading and successful transfer. The total discoloring of the membrane was performed with SDS 1% w/v and 0.2xSSC.

Prehybridization, probe preparation and hybridization steps were performed as described for Southern Blot above. After hybridization the membrane was washed twice in 2xSSC and SDS 0.1% w/v for 10 and 20 min, and once in 1xSSC and SDS 0.1% w/v for 15 min, under the same temperature as hybridization. The membrane was exposed to an X-ray film as described for Southern blot above.

2.15 Protein extraction from *A. nidulans*

2.15.1 Total protein extraction

The method used in this study for total protein extraction was performed as described in Galanopoulou et al. (2014). In more detail, the strain of interest was grown in an Erlenmeyer flask containing liquid MM supplemented as required, at the selected conditions. The mycelium was harvested and grinded in fine powder as described in DNA extraction protocol above. The powder was collected in a centrifuge tube and 800 μL precipitation buffer (Tris 50 mM pH 8, NaCl 50 mM, TCA 12.5% v/v) was added and incubated on ice for 20 min after vortexing. The suspension was centrifuged for 10 min at 13,000 g at 4^{o}C. The pellet was washed once with 1 mL ice cold EtOH and twice with 1 mL ice cold acetone, air-dried and resuspended in 200-500 μL extraction buffer (Tris 100 mM pH 8, NaCl 50 mM, SDS 1% w/v, EDTA 1 mM) supplemented with protease inhibitor cocktail. After centrifugation at 13,000 g for 10 min at 4^{o}C the supernatant was collected. The protein concentration was estimated by the Bradford assay.

2.15.2 Membrane enriched protein extraction

The method used in this study for membrane enriched protein extraction was performed as described in Pantazopoulou et al. (2007). In more detail, mycelia grown in MM, supplemented as required, were collected as described before and grinded to fine powder. The powder was suspended in 1 mL cold extraction buffer (Tris 10 mM pH 7.5, $MgCl_2$ 5 mM, NaCl 100 mM, sorbitol 0.3 M, PMSF 1 mM) and protease inhibitor cocktail and centrifuged at 2,000 g for 10 min to remove unbroken cells and cell walls. The supernatant was collect to a new centrifuge tube and centrifuged at 15,000 g for 1 h to pellet membranes. The pellet was resuspended in 100-200 μL resuspension buffer (HEPES 10 mM, Sucrose 250 mM, PMSF 1mM, DTT 1 mM, protease inhibitor cocktail) and centrifuged again at 3,000 g for 10 min. The supernatant was collected in a new centrifuge tube. The protein concentration was estimated by the Bradford assay.

2.16 Protein electrophoresis and western blot

Protein electrophoresis and western blot are used to detect a specific protein of interest included in a mixture of proteins. Protein electrophoresis is usually performed in denaturating acrylamide gels containing SDS, a technique known as SDS-page, where the proteins are separated according to their electrophoretic mobility (determined by the length of a polypeptide chain and its charge) while in the denatured (unfolded) state. Then, the proteins may be transferred to a membrane and specific antibodies are used to detect the protein of interest, a technique known as Western blot. In this study, the SDS-page and Western blot were performed according to the following protocol.

In more detail, the gel used for the SDS-page was a polyacrylamide gel consisting of two gel parts of different acrylamide concentration; the stacking and separating or running gel (**Figure 30**). The stacking gel has low concentration of acrylamide, allowing the proteins to line up before entering the separating gel, which has a higher concentration capable of retarding the movement of the proteins. The density of the separating gel, meaning the concentration of the acrylamide, determines the migration pattern of the proteins based on their size. In this study, the polyacrylamide gels were prepared as indicated in **Table 7**. The gels were casted in casting glass plates provided by Biorad and combs were used to form wells. When the gels were ready, were placed in the apparatus Mini-PROTEAN Tetracell (Biorad) and the tank was filled up with Running buffer (Tris 25 mM, glycine 192 mM, SDS 0.1% w/v).

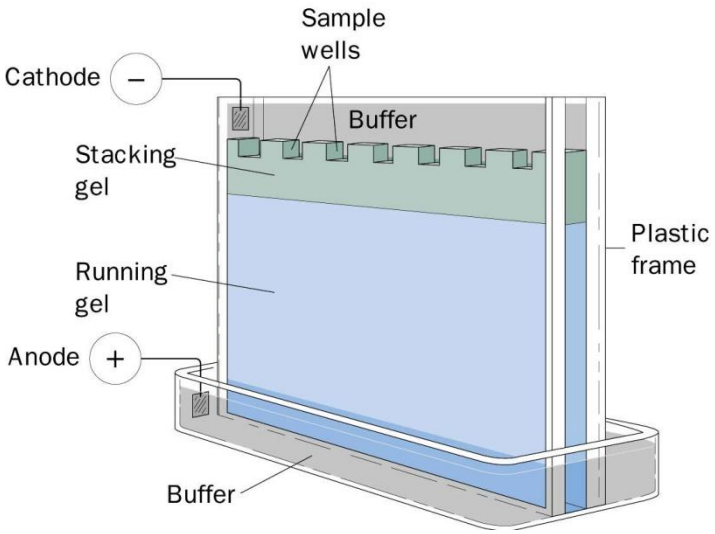


Figure 30. Representation of SDS page electrophoresis. Stacking gel is casted on top of the separating (running) gel. Proteins are loaded in the wells and migrate towards the anode. The picture was taken from *http://www.siumed.edu/~bbartholomew/ course_material/protein_methods.htm*.

Table 7. Composition of the SDS-acrylamide gels used in this study.

	Stacking gel 4 %	Separating gel 10 %
30 % Acrylamide/Bisacrylamide 37.5:1	1.3 mL	3.3 mL
Tris 1.5 M pH 8.8 (Lower Tris)	-	2.5 mL
Tris 0.5 M pH .8 (Upper Tris)	2.5 mL	-
SDS 20 % w/v	50 µL	50 µL
APS (Ammonium persulfate) 10 % w/v	50 µL	50 µL
TEMED	20 µL	10 µL
dH_2O	Up to 5 mL	Up to 10 mL

The protein samples were prepared with the addition of denaturating Laemli buffer (Tris 200 mM pH 6.8, SDS 8% w/v, bromophenol blue 0.2 % w/v, glycerol 40% v/v, β-mercaptoethanol 400 mM) and incubated for 30 min at 37^{o}C, in order to succeed complete denaturation of the proteins. Then, the samples were cooled down to RT and then loaded in the wells. Electrophoresis was performed by applying 80V electrical field until the dye reached the separating gel and then the electrical field was changed to 140V. When the electrophoresis was complete the gel was removed from the glass plates and prepared for transfer to a membrane or gel staining with Coomassie stain. Coomassie Brilliant Blue is a reagent that binds to proteins through ionic interactions between sulfonic acid groups of the dye and positive amine groups of the protein, as well as, through Van der Waals interactions and is often used to check the quality of the gel and proteins or as a second method to verify the equal loading of the gel after Bradford quantification.

In order to perform immunodetection of a specific protein, the separated proteins in the gel were transferred to a PVDF (Polyvinylidene fluoride) membrane with application of an electrical field in an apparatus similar to the one shown in

Figure 31. In more detail, the membrane was laid on the gel and Whatmann papers were laid on both sides of membrane/gel and all were set into a cassette specific for the apparatus used (Mini PROTEAN Trans-blot, Biorad). The tank containing the cassette was filled up with cold transfer buffer (Tris 25 mM, glycine 192 mM, methanol 20% v/v). The transfer was performed by applying 100V electrical field for 1-2 h.

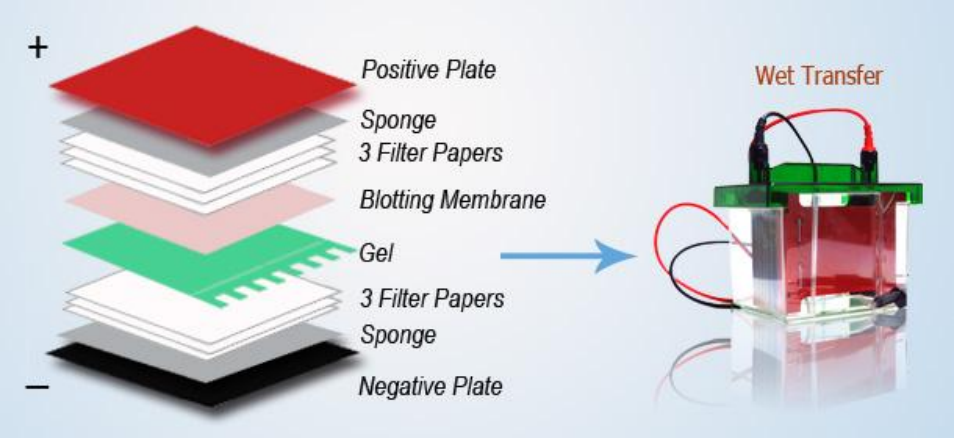


Figure 31. Representation of protein transfer from the gel to membrane. A membrane is laid on the gel and as shown in the picture, the gel and membrane are placed between filter papers and then in a transfer cassette. Proteins migrate with a direction from the negative pole towards the positive. The apparatus tank is filled with transfer buffer. The picture was taken from *http://www.western-blot.us/procedure-of-western-blot/western-blot-transfer*.

After the completion of the transfer, the membrane was removed and the efficiency of the transfer was checked by staining with Ponceau S, which specifically binds on the proteins. The membrane was discolored with dH_2O and was ready for immunodetection. The membrane can be stored for short periods at 4^oC in TBS-Tween solution (Tris-HCl 20 mM pH 7.5, NaCl 500 mM, Tween-20 0.1% v/v).

Western blot is used to detect particular proteins with specific antibodies. The antibodies used usually recognize the protein itself, or epitopes bound on the protein e.g. GFP (green fluorescent protein), His-tag (10 histidine residues). Since the membrane (nitrocellulose or PVDF) used has the ability to bind proteins, a preparation step is needed before adding the antibodies on the membrane to avoid non-specific binding. So, the membrane was first incubated with Blocking solution (non-fat dried milk 2% w/v in TBS-Tween solution) for 1 h at RT under mild agitation. In this study, two primary antibodies were used, the anti-GFP monoclonal antibody (Roche) and the anti-actin C4 antibody. The anti-GFP was used in dilution 1:2,500 in non-fat milk 2% w/v in TBS-Tween solution and the anti-actin in dilution 1:5,000 in BSA 3% w/v in TBS-Tween solution. In both cases the membrane was incubated with the antibodies for 2 h at RT under mild agitation. Then the antibodies were removed and the membrane was washed three times with TBS-Tween for 10

min at RT under mild agitation. Then, the membrane was incubated with the secondary antibody which was conjugated with the enzyme horseradish peroxidase. This enzyme in the presence of luminol emits low-intensity light at 428 nm; in the presence of certain chemicals, the light emitted is enhanced up to 1000-fold, making the light easier to detect and increasing the sensitivity of the reaction, a technique called enhanced chemiluminescence (ECL)

The secondary antibody was used in dilution 1:3,000-5,000 and incubated for 1 h at RT under mild agitation. Then, the membrane was washed as before three times with TBS-Tween. Afterwards, the excess solution was removed from the membrane with paper towels and the ECL substrate was applied on the membrane and incubated for 2 min at RT, afterwards the excess ECL substrate was removed and the membrane was wrapped in Saran wrap and put in a cassette with a film on top for 20 sec-10 min depending on the intensity of the signal. Finally, the film was developed and the bands corresponding to GFP-tagged proteins or actin appeared; actin is used as an internal marker for equal loading. The intensity of the bands was quantified by *Image J* with densitometry analysis. The relative density of the GFP specific bands were quantified and normalized by dividing with the respective actin-band density.

2.17 Uptake assays

Uptake assays were performed with germinating conidiospores at a morphological stage after polarity establishment and just prior to germ tube emergence. Approximately 10^8 conidiospores were used, germinating for 4-4.5 h at 37^{o}C, at 130-150 rpm. Empirically, approximately 10^8 conidiospores can be initially obtained, using a spatula, from a 4 day-grown culture, incubated on CM, at 37^{o}C. The inoculum, sufficient for 40-50 assays, was transferred in 5 mL of Tween-80 0.01 % v/v, briefly vortexed in order to separate the conidiospores from vegetative structures, filtered through a nylon net filter (Blutex, pore size ≥ 50 μm), and used to inoculate a 100 mL Erlenmeyer flask containing 25 mL MM supplemented with appropriate carbon and nitrogen sources and necessary vitamins.

A stock solution of the radiolabeled substrate of interest was prepared in water or MM, so that for each assay 25 μl of the stock would be used. The stock concentrations were adapted to the specific activity of the available radiolabeled substrate, and in order to ideally be 10- to 50-fold lower from the K_m value of the transporter studied. After reaching the optimal germination time, the conidiospores were collected by centrifugation at 3,000 g for 5 min at 15-25^{o}C, resuspended in 4 mL standard MM and distributed in 75 μl aliquots in 1.5 mL centrifuge tubes. Prior to uptake, aliquots were incubated in a heat block at 37°C (or 25^{o}C, if needed) for 5-10 min, before addition of radiolabeled substrate. A crucial parameter for estimating the kinetic parameters of solute uptake for any given transporter is the time of exposure to a relevant radiolabeled substrate. For measuring initial uptake rates, which are necessary for determining $K_{m/i}$ and apparent V_{max} values, the time of incubation was

defined through time-course experiments at different concentrations of labeled substrate. Most transporters show linearly increased activities for at least 1 min.

The transport reaction was stopped by adding an equal volume (100 μl) of ice-cold unlabeled substrate at 100- to 1000-fold excess concentration, brief vortexing, and direct transfer of the assay/eppendorf tube in an ice bucket. The conidiospores were then collected by centrifugation at 11,000 g for 3-5 min at 4°C, the supernatant was discarded through vacuum aspiration using a capillary Pasteur pipette, washed by vortexing in 1 mL ice-cold MM, followed by a second centrifugation at the same conditions as above. The pellet was resuspended in 0.5-1 mL of scintillation fluid (Toluene 66.6% v/v, Triton-X 33.3% v/v, PPO 0.266% w/v, POPOP 0.0066% w/v) and the centrifuge tubes were placed into scintillation vials for analyzing substrate accumulation in a scintillation counter.

For K_m determination of a transporter, different substrate concentrations should be used, for a fixed incubation time, previously determined to reflect initial uptake rates. In most cases this is 1 min. The range of concentrations used was determined at first empirically. In the final experiments, at least 3-4 concentration points below and above the apparent K_m value were used. For each concentration point measurements were performed in triplicate. The stock solutions were prepared using a mixture of fixed radiolabeled substrate and increasing concentrations of non-radiolabeled substrate. K_i measurements were determined by estimating IC_{50} values (concentration for obtaining 50% inhibition) of a given substrate/inhibitor, using the formula $K_i = IC_{50}/1+[S]/K_m$, where [S] is the fixed concentration of the radiolabeled substrate used. In this study, K_i values were equal to IC_{50} values since the [S] was very low (at least 10-fold lower than the K_m value).

K_i and V_{max} determination were carried out using standard Lineweaver-Burk or Eadie-Hofstee equations formulas, or a relevant software (e.g. GraphPad Prism). The analysis was performed through a non-linear regression curve fit and sigmoidal dose response analysis. Quality factors for the analysis result were: R^2 which should be > 0.99 and the Hill co-efficient (Weiss, 1997), which should be approximately -1 for a transporter with one binding site.

2.18 Microscopic analysis

Samples for fluorescence microscopy were prepared as previously described by Gournas et al. (2010). In brief, germlings incubated on coverslips or in 35 mm Petri dishes (WillCo-dish®, Series GWSt-3512) in liquid MM supplemented with $NaNO_3$ as the nitrogen source for 12–14 h at 25-32°C were observed on an Axioplan Zeiss phase-contrast or Zeiss Axio Observer Z1 inverted epifluorescent microscope, and the resulting images were acquired with a Zeiss-MRC5 digital camera using the AxioVs40 V4.40.0 software or with an AxioCamMR3 digital camera using the Zen 2012 software.

In cases that fluorescent dyes were used to recognize cellular structures e.g. vacuoles, the samples were incubated with the dye (at a final concentration described

by the manufacturer) for a short period, and then washed with MM and afterwards observed microscopically. Images were analyzed using Zen 2012 software, *ImageJ* or Adobe Photoshop CS5.

3 Results

3.1 Functional analysis of the purine transporter AzgA

In the current study, we performed a mutational analysis of AzgA in order to identify residues critical for transport activity and substrate specificity. AzgA is a high affinity and high capacity transporter, specific for adenine, hypoxanthine and guanine and one of the major purine transporters in *A. nidulans*, and as described in detail in Introduction, the 'founding' member of the AzgA-like subfamily of the NAT/NCS2 transporter family. The residues selected to be mutated were based on the model constructed by E. Mikros' team using homology modeling. The strains expressing the relevant mutants were functionally analyzed as will be described below.

3.1.1 Homology modelling, substrate docking and molecular dynamics of AzgA[2]

Homology modelling for AzgA was performed by E. Mikros' team (G. Lambrinidis and T. Evangelidis). The crystal structure of UraA, a uracil transporter from *E. coli* and a NAT/NCS2 family member, was used as a template. AzgA and UraA share a rather moderate sequence similarity (~20% identity in 418 amino acids overlap), which is however adequate to support a theoretical model. The AzgA model, shown in **Figure 32**, corresponds to a cytoplasmic-facing occluded conformation and

[2] Details concerning the methodology used in the computational analysis can be found in Krypotou et al., 2014, see Appendix.

consists of 14 TMSs, mostly helical connected with lengthy loops. Two characteristic beta strands are formed in the core of the protein disrupting the helical structure of TMS3 and TMS10. The model could be divided in two distinct domains, the core and the gate domain. The core domain consists of TMSs 1-4 and 8-11, whereas TMSs 5-7 and 12-14 form the gate domain. The predicted model is similar to the NAT models constructed with similar methods and already published (Amillis et al., 2011; Karena and Frillingos, 2011; Kosti et al., 2012). On the outer surface of the model a mild negative electrostatic potential is distributed, whereas the inner surface within a region below helices 13-14 is strongly positive. The intermediate region of the protein is neutral.

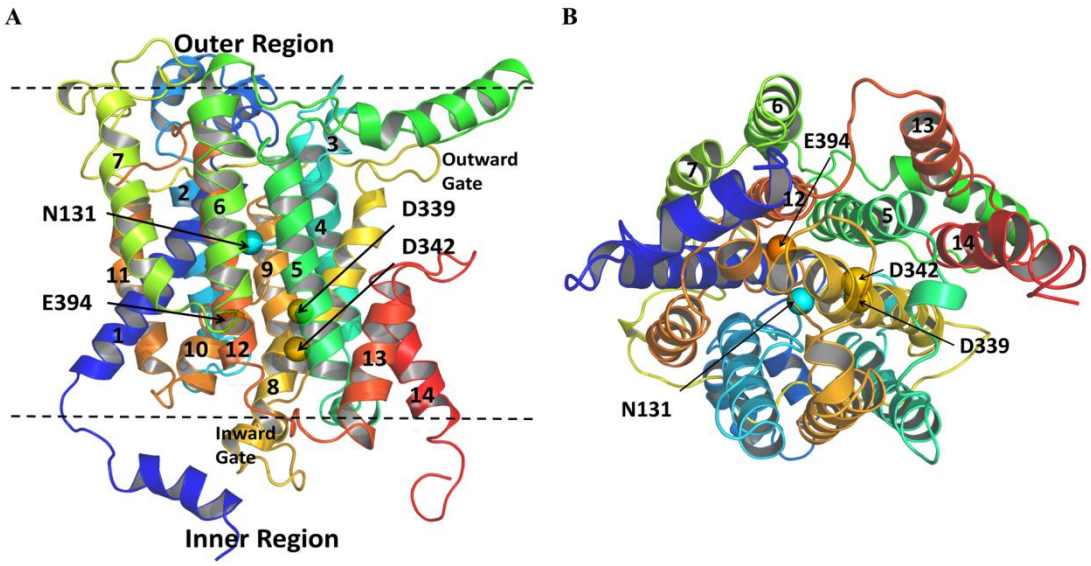


Figure 32. Theoretical model of AzgA. (A) Side-view of the model highlighting the topology of residues selected as crucial for the function. The transmembrane helices are designated with numbers (core domain: 1–4, 8–11; gate domain: 5–7, 12–14) and the planes of the membrane with dashed lines. **(B)** Bottom view of the same model. The substrate binding site is located in the space between the core and the gate domains. The topology of residues critical for the function of the substrate binding site is also shown. Homology modelling was performed by G. Lambrinidis and T. Evangelidis.

Induced Fit Docking (IFD) calculations were used to dock the physiological substrates (adenine, hypoxanthine or guanine) in the predicted binding site of AzgA. The poses found were validated in respect with the binding constants ($K_{m/i}$) of the ligands, namely K_m 2.8 μM for hypoxanthine, K_i 3 μM for adenine and K_i 1.3 μM for guanine. All three ligands were bound to the predicted binding site through H-bonds mediated by the polar side chains of residues Asn131, Asp339 and Glu394 and the amide group of the backbone of Val393. Guanine also interacted with Asn131 backbone and Thr390 side chain through a double hydrogen bond. Phe392 seemed to participate in a weak π-π stacking interaction with hypoxanthine and guanine, but not adenine. The results of the IFD, performed by G. Lambrinidis and T. Evangelidis, are shown in **Figure 33**.

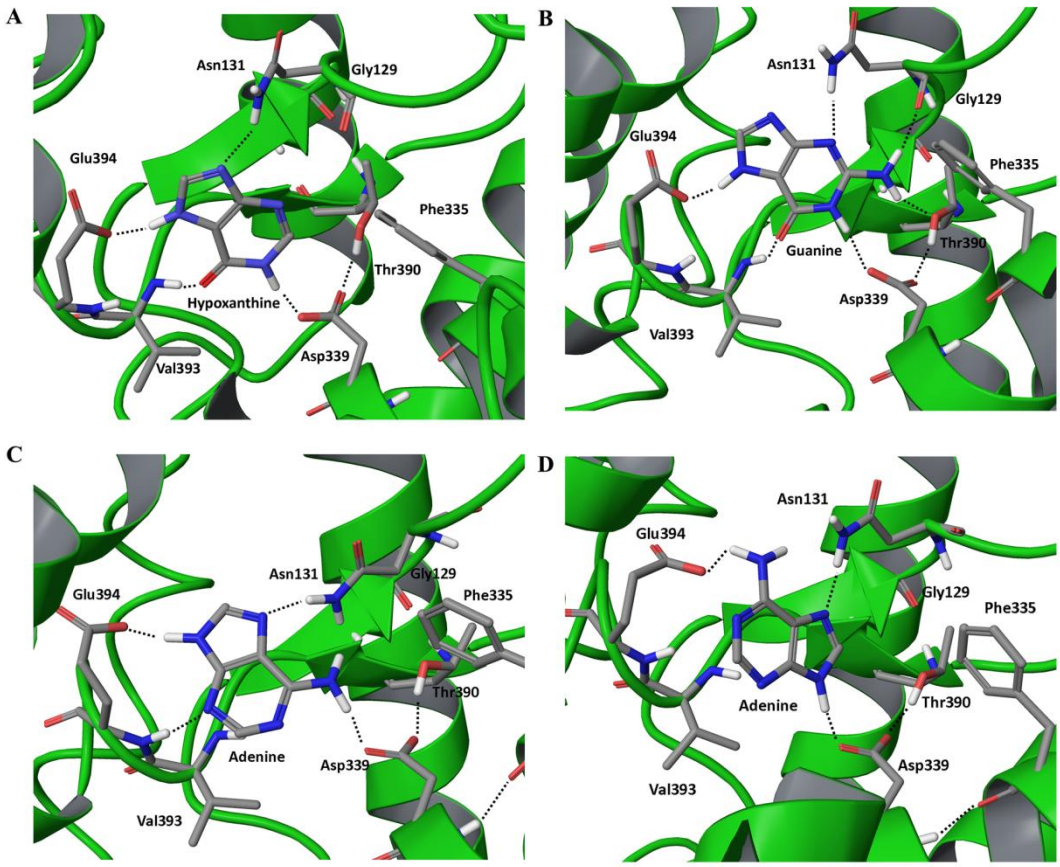


Figure 33. Substrate docking in AzgA. (A) Hypoxanthine, **(B)** guanine, **(C** and **D)** adenine. Amino acid residues implicated in substrate binding are shown. Hydrogen bonds are depicted with dashed lines. Substrate docking analysis was performed by G. Lambrinidis and T. Evangelidis.

Molecular dynamics (MD) analysis was performed by G. Lambrinidis and T. Evangelidis for the AzgA model in the presence of the three substrates used in IFD analysis. The MD proposed putative trajectories of the ligands and a transition of the inward-facing occluded state to an inward-facing open conformation. This transition was provoked by the upward movement of TMSs13-14 and the bending of TM helix 5, creating an inward-facing cavity below a putative inward-facing gate that is flooded with cytosolic water. During the simulation, hypoxanthine was constantly bound to Asn131, Asp339, Glu394 and Val393, whereas adenine and guanine were displaced from their initial position. Adenine moved from an initial position stabilized with H-bonds among residues Asn131, Ala132, Tyr133, Asp339, Thr390 and Glu394 towards TM helices 5, 8 and 14 and bound through H-bonding to Asp339, Glu394 and Asp478 side-chains. Guanine behaved similarly to adenine, initially was stabilized through H-bonding to residues Asn131, Asp339, Thr390 and Glu394 and then moved lower between TM helices 5, 8 and 14 and bound through H-bonding to Asp339, Asp342 and Phe335. Through water molecule mapping analysis Asp339, Glu394 and Asp342 were predicted to participate in a hydrogen bond network in the presence of hypoxanthine, putatively participating in H^+ trajectory.

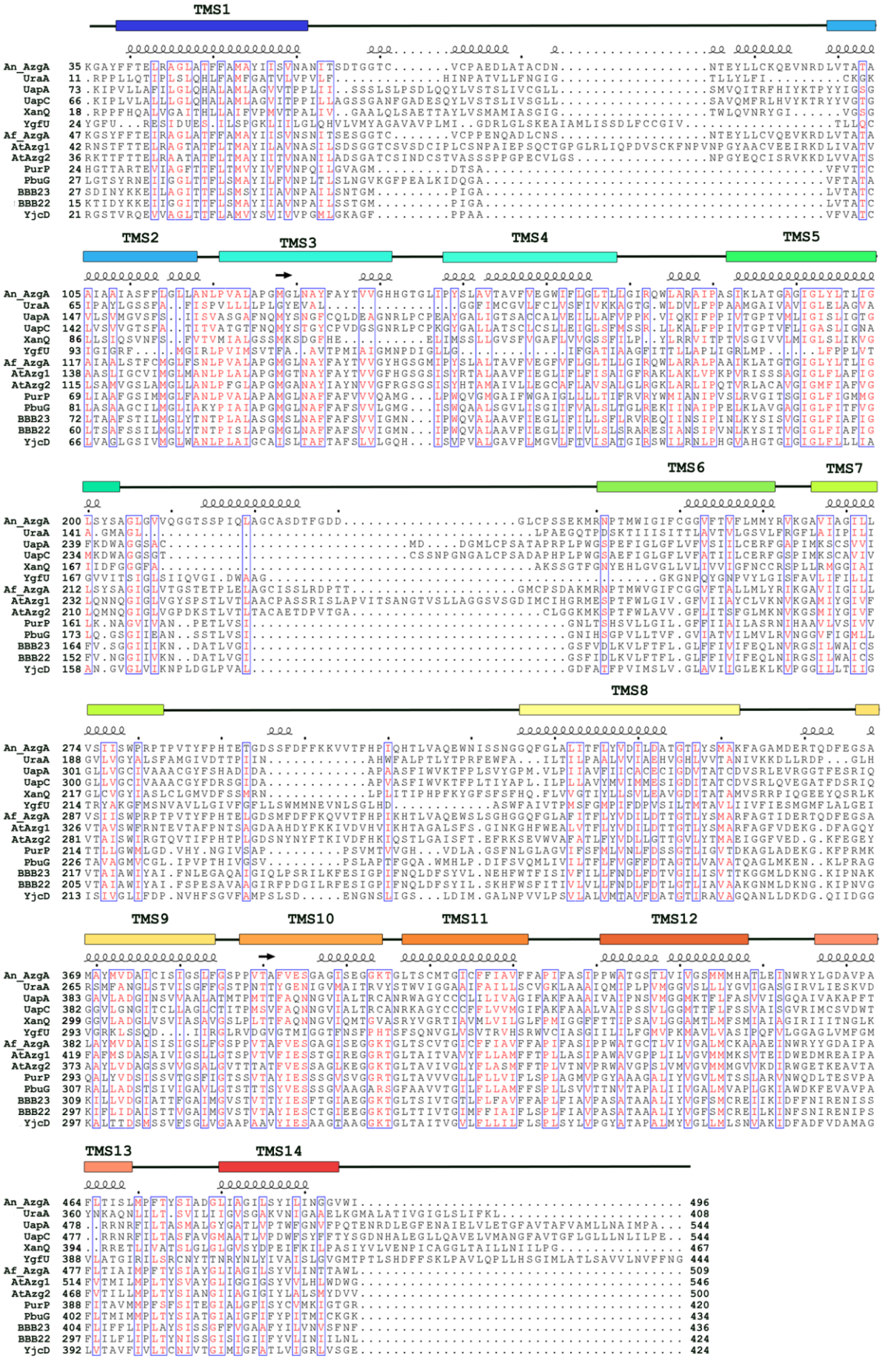
3.1.2 Mutational analysis of AzgA

From the computational analysis, few residues were predicted to be involved in substrate or/and H^+ binding. These residues were Asn131 (TMS3), Asp339 (TMS8), Asp342 (TMS8), Thr390 (TMS10) and Glu394 (TMS10). Protein sequence alignment between AzgA and characterized NAT homologues (**Figure 34**) was used to compare the residues identified by the computational analysis and those already known from previously published works as critical for function or/and specificity in NAT homologues.

Putatively critical residues emerging from the alignment are: (*i*) a highly conserved Thr in all AzgA-like proteins, corresponds to His86 (TMS1) in UapA that is proved important for transport capacity (Pantazopoulou and Diallinas, 2009). (*ii*) Asn131 (TMS3) in AzgA that is predicted to participate in direct interactions with all substrates, corresponds to a highly conserved Asn residue in AzgA-like proteins. (*iii*) Asp342 (TMS8) is highly conserved among AzgA-like proteins and corresponds to Glu356 of UapA characterized as one of the critical residues for UapA substrate binding (Papageorgiou et al., 2008). (*iv*) Glu394 (TMS10) is highly conserved in all AzgA homologues and corresponds to Gln408 in UapA, a residue important for substrate binding and specificity (Koukaki et al., 2005).

Based on the above, a series of amino acid substitutions were constructed to determine the role of the residues predicted by the computational analysis, as well as, those varying from respective important ones in other NAT homologues. Each substitution was designed to replace polar residues with Ala or isofunctional or isostructural residues. The mutants constructed were: T49A, T49S, G129A, G129F, N131A, N131D, F335A, D339A, D339N, D339E, D342A, D342N, D342E, T346A, T390A, E394A, E394D, E394Q, S395A, S395N, G398A, G402A, K404R and D478A. Each mutation was made in a vector carrying *azgA* cloned in frame with the green fluorescent protein *gfp* (*pBS-azgA-gfp argB*). The mutant alleles were introduced in a null mutant strain *ΔuapA ΔuapC ΔazgA argB2 pabaA1* (ΔACZ), lacking the three major endogenous purine transporters. Transformants were selected based on arginine auxotrophy complementation. Details about strains, plasmid construction and transformation are in the Materials and methods section.

Figure 34. Multiple sequence alignment of AzgA-like sequences of known function and characterized NAT homologues. Putative TMSs of AzgA are denoted in colored cylinders of the same color as that shown in Figure 32 depicting the 3D model. Invariant and highly conserved amino acids are shaded in red and blue-lined boxes, respectively. The sequences shown are: AzgA (CAE00849.1) and AzgA-like sequences from *A. fumigatus* (AFUA_5G09750), *E. coli* (PurP/P31466.2 and YcjD/ AGX33403.1), *Bacillus subtilis* (PbuG/O34987.1), *B. burgdorferi* (BBB22 NP_047008.1 and BBB23/NP_047009) and *A. thaliana* (NP_566384.1 and NP_199841.2), as well as NAT sequences from *A. nidulans* (UapA/CAA50681.2, UapC/CAA56190.1) and *E. coli* (UraA/YP_490725.1 XanQ/NP_417358.2, YgfU/YP_491089.1).


TMS1
TMS2
TMS3
TMS4
TMS5
TMS6
TMS7
TMS8
TMS9
TMS10
TMS11
TMS12
TMS13
TMS14

3.1.3 Phenotypic analysis of AzgA mutants

AzgA function can be tested by growth test assays in the presence of adenine or hypoxanthine as sole nitrogen sources, or on the toxic analogue 8-azaguanine. Functional AzgA confers growth on adenine and hypoxanthine and sensitivity to 8-azaguanine. Transformants obtained through single and multi-copy integration events, verified by Southern blot analysis, were subjected to growth tests on adenine, hypoxanthine, 8-azaguanine and uric acid. None of the mutations led to growth on uric acid (not shown). Growth tests of selected single copy transformants are shown in **Figure 35**. Single and multi-copy transformants had more or less similar phenotypes (not shown). Leaky growth on hypoxanthine and especially adenine of the *ΔazgA* negative control is due to the low capacity secondary purine transporter FcyB.

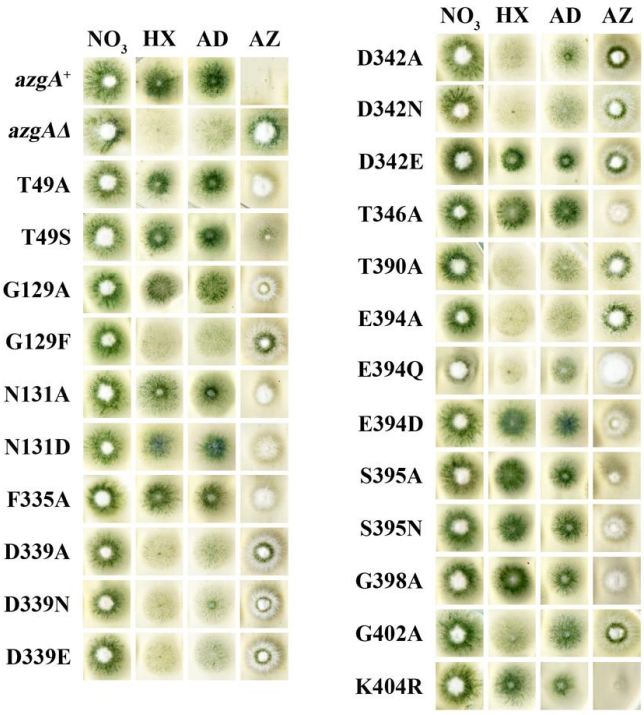


Figure 35. Phenotypic analysis of AzgA mutants. Growth tests on sodium nitrate (NO_3; 10 mM), hypoxanthine (HX; 0.1 mg/mL) or adenine (AD; 0.1 mg/mL) as sole nitrogen sources and resistance/sensitivity test on 8-azaguanine (AZ; 0.1 mg/mL) at 32°C. All AzgA mutant strains shown are single copy plasmid integrations. Positive (*azgA*$^+$) and negative (*azgAΔ*) isogenic strains are shown as isogenic controls.

In more detail, from the growth test analysis the mutants can be classified in two groups: (*i*) *wt-like*: correspond to full growth on hypoxanthine and adenine and sensitivity on 8-azaguanine (T49A, T49S, G129A, N131A, N131D, F335A, D342E, T346A, E394D, S395A, S395N, G398A, K404R and D487A); (*ii*) *loss-of-function*: correspond to total or partial absence of growth on hypoxanthine and/or adenine and/or resistance on 8-azaguanine (G129F, D339A, D339N, D339E, D342N, D342A, T390A, E394A, E394Q and G402A). As expected, all mutants showed normal growth on neutral nitrogen sources, e.g. nitrate (NO_3). Interestingly, two mutants grew better

on hypoxanthine than adenine (D342E, S395A) and two (E394Q, G402A) better on adenine than hypoxanthine.

3.1.4 AzgA localization and protein stability

We performed epifluorescence microscopy to detect the subcellular localization of AzgA-GFP for all mutant alleles (**Figure 36**). In all cases, except E394A, AzgA-GFP was localized in the plasma membrane. In some mutants, namely N131D, T390A, D342A, E394A, E394Q and G402A, a large amount of the protein was localized in cytoplasmic structures identified as the perinuclear endoplasmic reticulum (ER) (characteristic rings) or the vacuole. ER retention is usually associated with protein misfolding and high vacuolar localization with increased protein turnover.

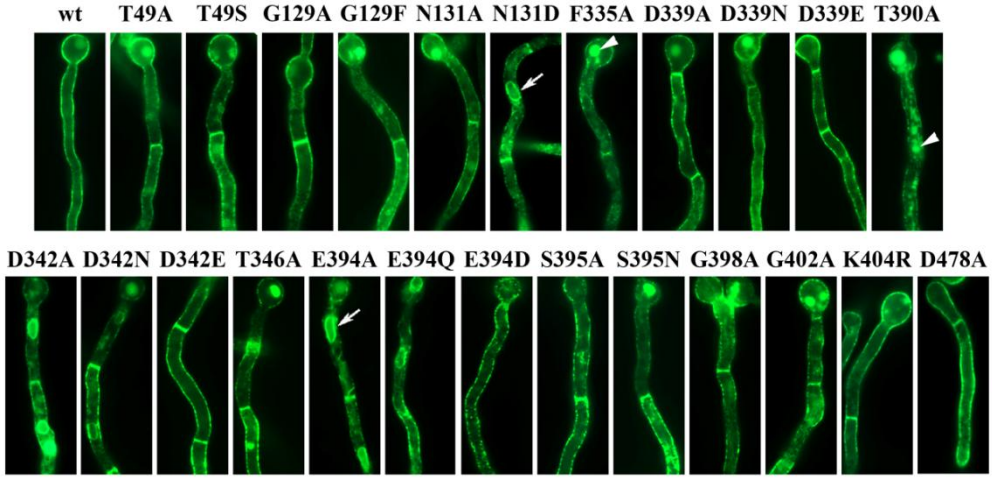


Figure 36. Epifluorescence microscopy showing *in vivo* subcellular expression of AzgA–GFP mutant alleles and a wt control. In selected samples, arrows and arrow heads depict perinuclear ER membrane and vacuoles respectively. All samples were prepared from fresh conidiospores inoculated in liquid MM supplemented with $NaNO_3$ (10 mM) and were incubated for 12-14h at 32^{o}C. Microscopic analysis was performed at 32^{o}C. The same strains were used in both the microscopic and growth test analysis.

Briefly, most mutations led to wt-like localization of the protein and some contributed to ER retention and increased turnover. Western blot analysis of total protein extracts was performed to further analyze protein stability and turnover of the mutants (**Figure 37**). Anti-GFP antibody was used to detect AzgA-GFP, as well as free-GFP that indicates the levels of AzgA vacuolar degradation. In six cases (T49A, T49S, G129A, N131A, T390A, F335A) extremely low amounts of AzgA-GFP were detected, accompanied with high amounts of free-GFP, indicating mutant proteins with increased turnover rate compared with the wild-type AzgA-GFP. Most mutations (N131D, D478A, D339E, D342A, D342N, D342E, T346A, E394A, S395N, G398A, G402A, K404R) led to wt-like protein stability. The few rest mutations (G129F, D339A, D339N, E394Q, E394D, S395A) led to proteins with increased stability compared with the wt protein.

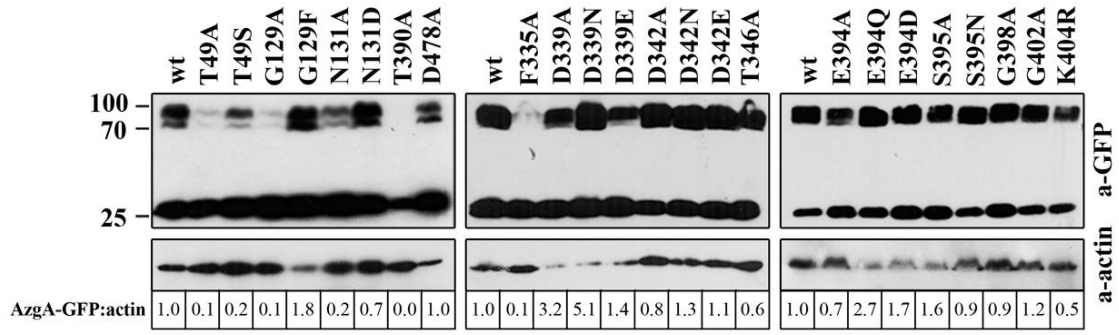


Figure 37. Western blot analysis of total protein extracts from AzgA–GFP mutants and a wt control detected with anti-GFP antibody. Antibody against actin was used as an internal marker for equal loading. Molecular weight markers in kDa are shown. Relative amounts of AzgA–GFP, normalized to actin protein levels, are shown below the blots. Cultures used for protein extraction were prepared from fresh conidiospores inoculated in liquid MM supplemented with $NaNO_3$ (10 mM) and were incubated for 12-14h at 32^{o}C and 120 rpm. Strains used for the western blot analysis correspond to the ones used for the previous analyses.

In summary, most mutations led to wt-like localization and protein levels. In this case, loss of purine uptake in the relevant mutants is very probably due to problematic substrate binding and/or transport. In contrast, loss of apparent transport activity in mutants with problematic protein localization and increased turnover, could simply, but not necessarily, be due to reduced protein steady state levels, rather than modified kinetic characteristic of AzgA. In this case, contribution of the relevant amino acid residues on the apparent rate of transport cannot be evaluated.

3.1.5 Kinetic analysis of AzgA mutants

Apparent initial uptake rates (*V*) for all mutants were measured by direct transport assays (See Materials and methods) using [^{3}H]-hypoxanthine. As shown in **Figure 38A**, only mutants K404R and D478A retained transport rates similar to the wt (80-90%), T346A and S395A had transport rates >50% compared with the wt, between 30-50% are the mutants: T49A, T49S, G129A, F335A, D342E, E349D, S359N and G398A, whereas the rest of the mutants showed dramatically lower transport rates, ranging from 1-20%, compared with the wt. Previous experience with other transporters studied in our lab has shown that a threshold of ~30% transport rate of that of the wt is critical for growth on the respective substrate. For AzgA, this threshold seems even lower, being ~20% of the wt (See **Figure 35** and **Figure 38**).

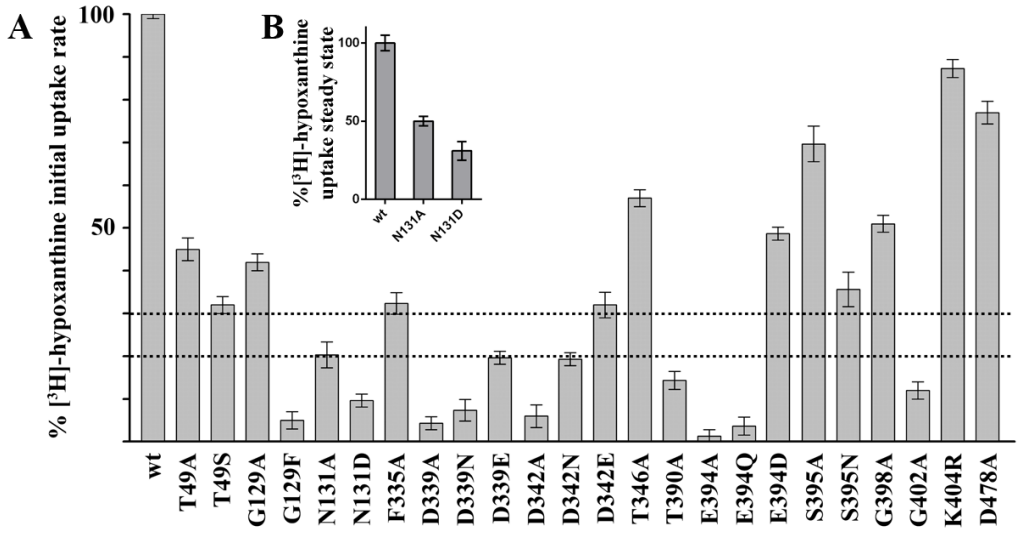


Figure 38. Apparent transport rates for AzgA mutants. (A) Comparative initial uptake rates of [^{3}H]-hypoxanthine for 1 min in AzgA mutant alleles and a wild type (wt) (azgA$^+$) control. 100% is the transport rate of the wt. All strains used are the same as the ones used in previous analyses and correspond to single copy plasmid integrations. **(B)** Comparative uptake of [^{3}H]-radiolabeled hypoxanthine for 5 min (steady state) for N131A and N131D and a wild type (wt) control. 100% is the transport rate in the wt.

In general the uptake assays were in good agreement with the growth tests on hypoxanthine. The only exceptions were mutants N131A and N131D, which grew similarly to the wt on hypoxanthine and had 20% and 8% *V* values, respectively, compared to the wt. This apparent disagreement could be explained by taking into account that growth tests reflect steady state substrate uptake, rather than uptake rates. Steady state uptake experiments were conducted for mutants N131A and N131D. AzgA steady state levels of transport correspond to the uptake rate plateau after 5 min incubation with the substrate (Cecchetto et al., 2004). Results showed that the N131A and N131D mutants reach a 5 min transport steady state level of 30-60% compared with the wt, in agreement with their growth on hypoxanthine (**Figure 38B**).

Binding and inhibition values ($K_{m/i}$) for hypoxanthine, adenine and guanine were determined and results are shown in **Table 8**. For mutants with low *V* (< 30%) we used multi-copy transformants (2-4 plasmid copies integration) that have higher apparent transport capacity (higher number of *azgA* copies results in more protein and, thus, higher substrate uptake, but no effect on binding and inhibition affinities). In detail, most mutations had minor effects on substrate binding, that corresponded to ~2-fold increase or decrease in K_m or/and K_i, e.g. T49A led to 1.5-fold decrease for hypoxanthine and adenine, but ~2-fold increase for guanine. Serine substitution of Thr49 led to variously increased affinities for all three substrates. Gly129 substitutions had more severe effects; when replaced with Ala, AzgA could bind with ~3-10-fold higher affinity all three substrates, but when replaced with Phe the [^{3}H]-hypoxanthine uptake could not be measured. Dramatic increase of affinity, up to a level of ~10-fold for all substrates, was observed in all substitutions of Asp339.

Asp342 substitutions had a more diverse result; D342A behaved similarly to wt, D342N had ~2-3.5-fold higher affinity for hypoxanthine and adenine and ~2-fold lower for guanine, whereas D342E led to decreased affinity for all substrates: 1.3-fold for hypoxanthine, ~4-fold for adenine and 23-fold for guanine, indicating a possible participation of Glu342 in substrate specificity and thus affected differentially the substrates tested. Similarly to D342E, E394D had a variable effect on substrate binding; hypoxanthine affinity increased ~2-fold, while adenine affinity decreased ~5-fold. Ala and Glu substitutions of Glu394 had no measurable uptake, probably led to protein misfolding as they were mainly localized in the ER (See **Figure 36**). Mutations of the rest TMS10 residues (S395A/N, G398A, G402A, K404R) resulted in minor changes in substrate affinities (See **Table 8**).

Table 8. Kinetic profile of AzgA mutants. $K_{m/i}$ values (μM) were determined as described in Materials and methods. Results are averages of at least three independent experiments in triplicates for each concentration point. Standard deviation was <20 %. (n.m.: non-measurable). HX, AD, GU stand for Hypoxanthine, Adenine and Guanine respectively. For mutants with low *V* (< 30%) multi-copy strains (2-4 plasmid integrations) were used.

	K_m (μM)	K_i (μM)	
	HX	**AD**	**GU**
Wt	2.8	3	1.3
T49A	1.8	2	2.5
T49S	1.6	0.7	0.4
G129A	0.3	0.3	0.4
G129F	n.m.	n.m.	n.m.
N131A	1	16	2
N131D	2	1	4
F335A	1.6	0.7	1.1
D339A	0.3	0.2	0.5
D339N	0.2	<0.2	<0.2
D339E	<0.2	0.3	0.2
D342A	3.1	3.9	1
D342N	0.8	1.5	3
D342E	3.7	16	30
T346A	2	3.2	1.3
T390A	3.7	1.3	2.8
E394A	n.m.	n.m.	n.m.
E394Q	n.m.	n.m.	n.m.
E394D	1.5	16	3.6
S395A	2.3	3.4	2.8
S395N	1	1	0.4
G398A	2.1	3.3	1.4
G402A	3.1	1.5	0.7
K404R	4.2	2.0	2.1
D478A	4	3.2	4.4

In summary, for mutations that led to problematic trafficking or high protein turnover we could not extract further conclusions from the initial uptake rate measurements. However, for all other mutations that do not alter the protein levels of AzgA expressed in the plasma membrane, we could directly draw conclusions on the contribution of different amino acids on substrate binding and transport.

Summarising:

(i) Mutations in Gly129 led to a protein with either increased substrate binding affinities (G129A) or no detectable function (G129F); these results indicate contribution of Gly129 in substrate binding.

(ii) Asn131 substitutions affected substrate binding; removal of the polar side chain and replacement with a neutral residue (N131A) led to an AzgA allele with lower transport capacity and decreased affinity for adenine, indicating interaction of Asn131 with substrates and especially adenine.

(iii) Replacements of the two aspartate residues (Asp339 and Asp342) led to diverse results, but in all cases to a less functional transporter. Specifically, replacement of Asp339 resulted in a high-affinity binding site for all substrates, but to a protein without transport capacity. This seems to be a case of a mutant that is able to "trap" and unable to release its substrate due to extremely high affinity. On the other hand, Asp342 replacements led to either misfolded partial loss-of-function proteins (D342A, D342N) or to a functional protein (D342E) with altered binding affinities and transport capacity. So, both Asp339 and Asp342 are predicted to be elements of the substrate binding pocket.

(iv) From the rest mutants, mutations in Glu394 affected grossly the protein folding and trafficking (E394A, E394Q) and affected differentially substrate binding affinities (E394D) when reducing the length of the side chain; thus Glu394 is also predicted to participate in substrate binding.

Briefly, the above observations suggest, independently from modelling and docking approaches, that residues Gly129 and Asn131 in TMS3, Asp339 and Asp342 in TMS8, and Glu394 in TMS10, are involved in substrate binding/or transport.

3.2 Characterization of two *A. brasiliensis* NAT homologues

The NAT family, as described in detail in the Introduction, includes a functionally uncharacterised group of proteins named UapD that shows a patchy conservation in ascomycetes. The black Aspergilli are the only members of the *Aspergillus* genus that are predicted to include in their genome *uapD*-like sequences. In this study, UapC- and UapD-like alleles from *A. brasiliensis* were selected to be characterized by heterologous expression in *A. nidulans.* Both genes were expressed in mutant strains lacking endogenous NAT transporters and the relevant strains were characterized through growth test and kinetic analysis.

3.2.1 Heterologous expression and functional characterization of AbUapC and AbUapD

The *A. brasiliensis* NAT homologues, from now on referred as AbUapC and AbUapD, were identified *in silico.* The *AbuapD* ORF was amplified from *A. brasiliensis* genomic DNA and cloned in a pBluescript vector carrying *uapA* up- and down-stream regulatory sequences and the *argB* gene as a selective marker (See Materials and methods). The selection of the *uapA* regulatory sequences was chosen given the extensive work already published and the knowledge of our lab concerning *uapA* expression profile. *AbuapC* ORF was also cloned in the same vector (See Materials and methods) as a control for testing whether 'canonical' NATs from *A. brasiliensis* are functional in *A. nidulans*. The vector used, also carried the *gfp* ORF so that both *AbuapC* and *AbuapD* genes were cloned in frame, at their C-termini, with it.

The plasmids were transformed into a null mutant strain lacking the NAT transporters (*ΔuapA ΔuapC ΔazgA* or ΔACZ) and transformants were selected based on complementation of the arginine auxotrophy. Selected transformants of single copy plasmid integration were tested by growth tests on uric acid, xanthine, hypoxanthine and the toxic analogue oxypurinol. The negative ΔACZ control could not grow on the purines tested and was resistant to oxypurinol. As shown in **Figure 39A**, strains expressing *AbuapC* grew on uric acid and xanthine and were sensitive to oxypurinol, whereas strains expressing *AbuapD* showed no distinct phenotype. A strain expressing *AbuapD* with *gfp* at the N-terminus (GFP-AbUapD) had similar phenotype to AbUapD-GFP, as shown in **Figure 39A**.

Microscopic analysis of the strains used in the growth tests was performed in order to detect their localization (**Figure 39B**). AbUapC-GFP is mainly localized in the plasma membrane and occasionally in vacuoles. On the other hand, AbUapD-GFP is mainly localized in internal cytoplasmic structures, which we were able to identify as vacuoles and perinuclear endoplasmic reticulum (ER) by staining (**Figure 39B**). Similar results were obtained for GFP-AbUapD. So, for AbUapD no plasma membrane localization was detected in all cases. Usually ER localization and increased vacuolar degradation is associated with protein misfolding, suggesting that AbUapD is probably intrinsically misfolded, at least when expressed in *A. nidulans*.

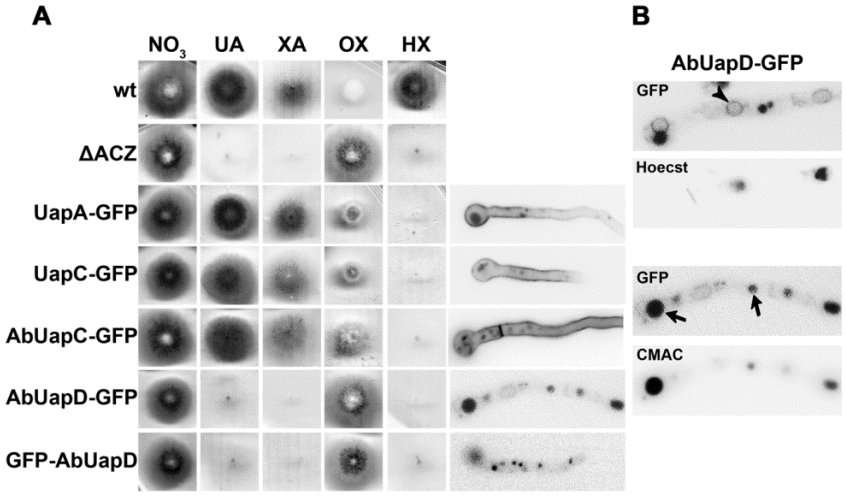


Figure 39. Characterization of the strains expressing *AbuapC* and *AbuapD* alleles. (A) Left panel: Growth test of selected transformants expressing *AbuapC-gfp* and *AbuapD-gfp*. A wild type strain (wt), an isogenic negative control (ΔACZ), and isogenic strains expressing *uapA-gfp* or *uapC-gfp* are also shown. Right panel: Microscopic analysis for GFP-detection in strains shown in the growth test. **(B)** Microscopic analysis of the strain expressing *AbuapD-gfp*. In the upper panel perinuclear ER rings (arrowhead) were identified by nuclear staining with Hoechst 33342 dye. In the lower panel vacuoles (arrows) were identified by staining with CMAC dye. Conidiospores were inoculated in MM and grown for 16 h at 25^{o}C and pH 6.8 (NO_3: sodium nitrate, 10 mM; UA: uric acid, 0.1 mg/mL; XA: xanthine, 0.1 mg/mL; HX: hypoxanthine, 0.1 mg/mL; OX: oxypurinol, 100 μM).

After having identified the specificity profile of AbUapC, we performed growth tests of *A. brasiliensis* wt strain on the same substrates (**Figure 40**). *A. brasiliensis* grew on xanthine and uric acid and was moderately sensitive to oxypurinol probably due to AbUapC substrate mediated transport.

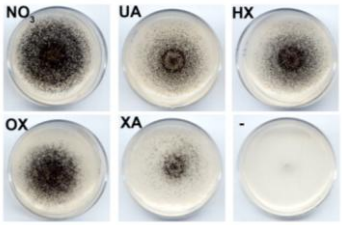


Figure 40. Growth test of *A. brasiliensis* at 37^{o}C. (NO_3: sodium nitrate, UA: uric acid, XA: xanthine, HX: hypoxanthine, OX: oxypurinol, – : lack of any nitrogen source; Concentrations are as in **Figure 39**).

3.2.2 Kinetic characterization of AbUapC

Radiolabeled xanthine uptake was used to characterize kinetically AbUapC (**Figure 41**). AbUapC xanthine mediated transport was linear for the first two minutes and then reached a plateau similarly to most transporters studied so far in *A. nidulans* (**Figure 41A**). A set of various purine and pyrimidine analogues were tested to determine the specificity profile of AbUapC (**Figure 41B**). AbUapC-mediated [^{3}H]-

xanthine uptake was reduced ~70–90% in the presence of excess (1000-fold) xanthine, uric acid, hypoxanthine, 1-, 3- or 8-methylxanthine and ~50–30% in the presence of guanine, adenine, 6-thioxanthine, 8-azaxanthine, uracil or cytosine.

Substrate binding affinities and inhibition constants ($K_{m/i}$) were determined for xanthine, uric acid and hypoxanthine (**Figure 41C**). AbUapC binds with high affinity xanthine (K_m 23 μM), medium affinity uric acid (K_i 160 μM) and low affinity hypoxanthine (K_i 700 μM). As shown in **Figure 41C**, AbUapC has similar affinities to UapC and thus properly was characterized as a UapC-like protein.

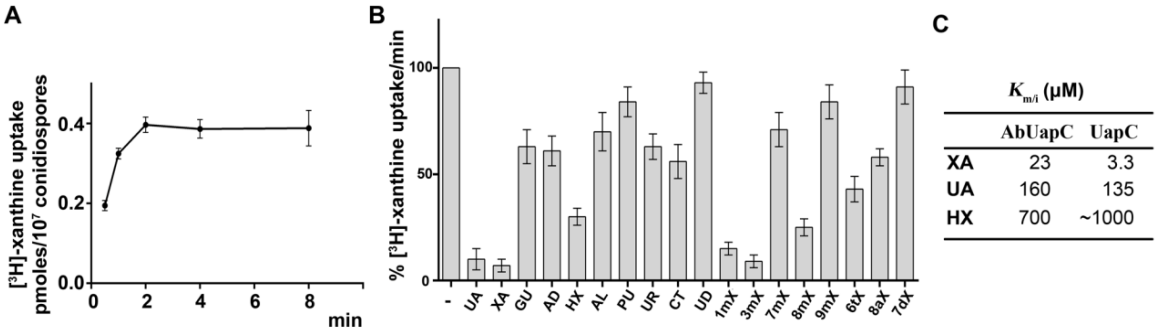
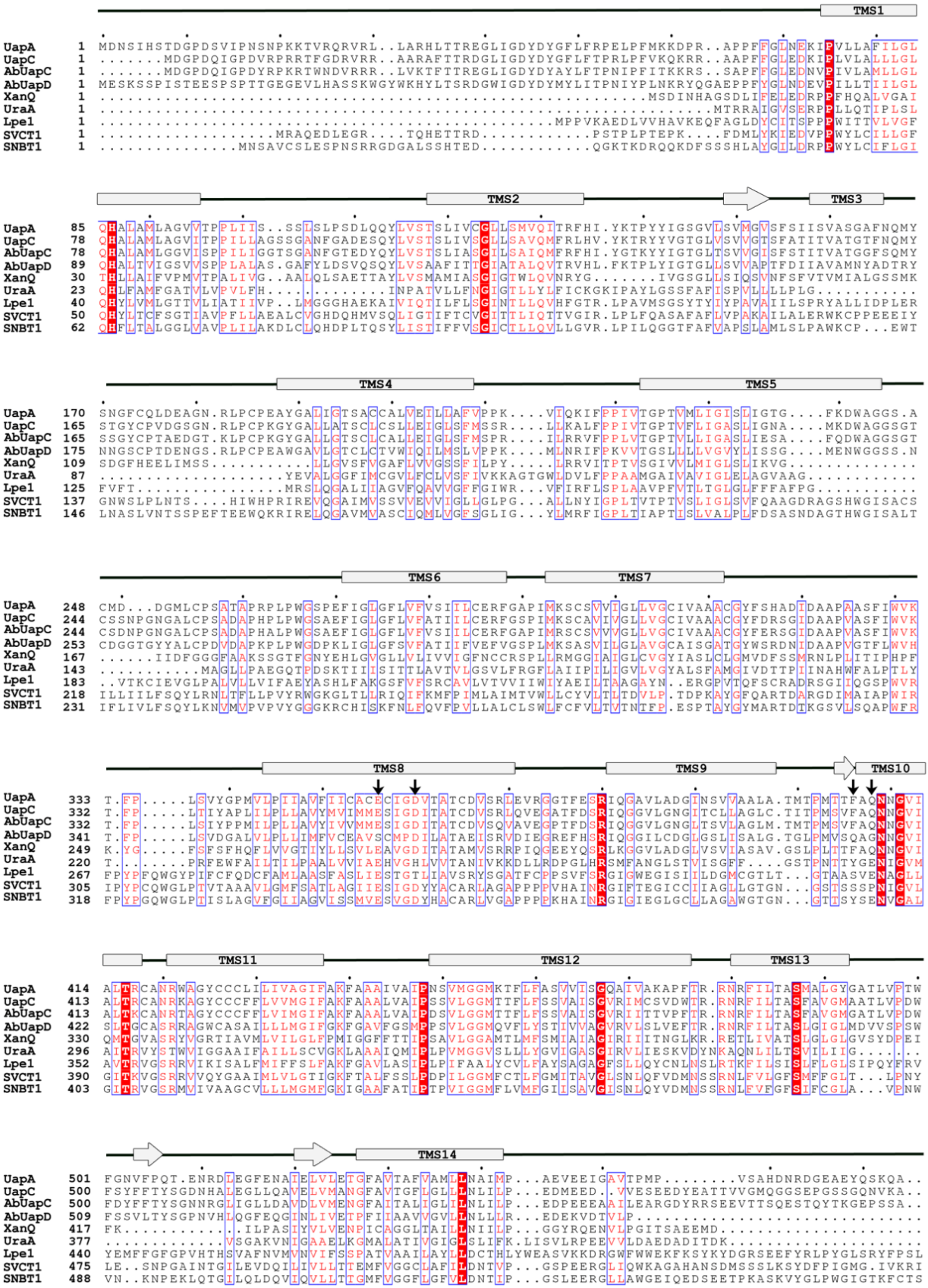


Figure 41. Kinetic profile of AbUapC. (A) Time course of AbUapC mediated [^{3}H]-xanthine. **(B)** AbUapC mediated [^{3}H]-xanthine uptake in the presence of excess (1000-fold) of different substrates/inhibitors. **(C)** Affinity and inhibition kinetic values $K_{m/i}$ (μM). Results are averages of at least three independent experiments in triplicate for each concentration point. Standard deviation was <20% (UA: uric acid, XA: xanthine, GU: guanine, AD: adenine, HX: hypoxanthine, AL: allantoin, PU: purine, UR: uracil, CT: cytosine, UD: uridine, 1mX: 1-methylxanthine, 3mX: 3-methylxanthine, 7mX: 7-methylxanthine, 8mX: 8-methylxanthine, 9mX: 9-methylxanthine, 6tX: 6-thioxanthine, 8aX: 8-azaxanthine, 7dX: 7-deazaxanthine).

3.2.3 Construction of chimeric AbUapD proteins

AbUapD shares 44.2% identity with UapA in 584 amino acids overlap and 43.7% identity with UapC in 593 amino acids overlap. Protein sequence alignment between AbUapD and NAT homologues is shown in **Figure 42**.

Figure 42. Sequence alignment of selected NAT homologues. The sequences shown are: AbUapC (JGI: 123452), AbUapD (JGI:52221; *A. brasiliensis*); UapA (UniprotID: Q07307), UapC (UniprotID: P48777; *A. nidulans*); Lpe1 (UniprotID: Q41760; *Arabidopsis thaliana*); UraA (UniprotID: P0AGM7), XanQ (UniprotID: P67444; *E. coli*); SVCT1 (UniprotID: Q9UHI7; *Homo sapiens*); SNBT1 (UniprotID: D2KX48; *Rattus norvegicus*). Predicted transmembrane segments (TMS) for UapA are denoted with rectangles and TMSs forming short b-sheets are shown with horizontal arrows (Kosti et al., 2012). Residues of the NAT signature motif varying in UapD-like proteins are shown with vertical arrows. Invariant and highly conserved amino acids are shaded in red and blue-lined boxes, respectively. Absolutely conserved amino acids are in red shaded boxes.


TMS1
TMS2
TMS3
TMS4
TMS5
TMS6
TMS7
TMS8
TMS9
TMS10
TMS11
TMS12
TMS13
TMS14
UapA
UapC
AbUapC
AbUapD
XanQ
UraA
Lpe1
SVCT1
SNBT1

AbUapD has a shorter C-terminus than UapA and UapC and thus, we made the hypothesis that a longer C-terminus might contribute to plasma membrane localization. We constructed an AbUapD chimaera in which the C-terminal region of AbUapD is replaced with that of UapA, tagged with GFP. The transformants constructed after single copy plasmid integration were analyzed by growth testing and epifluorescence microscopy (**Figure 43**). The chimeric protein, named AbUapD-A, had no distinct phenotype at growth tests and was localized in ER and vacuoles, similarly to AbUapD-GFP.

We observed that a functionally critical section of UapA, including major elements of the substrate-binding site of canonical NATs (TMS8–TMS10), is highly similar to that of AbUapD. As known from previous studies, replacement of this part of UapA with the respective one from UapC, resulted in a functional UapA molecule with a UapC-like specificity profile (Diallinas et al., 1998). We thus replaced the relative middle part of UapA (amino acids 311-443) with that of AbUapD (amino acids 319-451), and constructed a chimeric molecule, named UapADA. This chimera was introduced in the ΔACZ strain and single copy transformants were analyzed as described previously. No transport-related function was detected by growth testing in any of the transformants analyzed. In addition, in these transformants, UapADA-GFP was found to be localized in ER and vacuoles (**Figure 43**). Detailed descriptions of the plasmid constructions are included in Materials and methods.

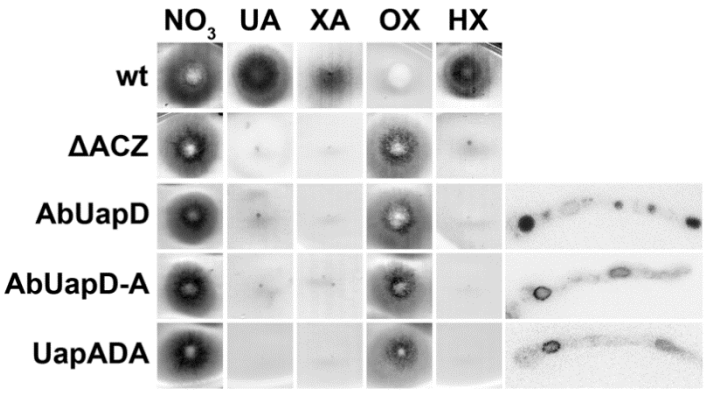


Figure 43. Growth tests and microscopic analysis of the AbUapD chimeric proteins (NO_3: sodium nitrate, UA: uric acid, XA: xanthine, HX: hypoxanthine, OX: oxypurinol; Concentrations are as in **Figure 39**).

In summary, AbUapC was found to be a high-affinity xanthine and uric acid transporter, which also contributes to low-affinity hypoxanthine binding, similar to the *A. nidulans* or *A. fumigatus* UapC orthologues (Diallinas et al., 1995; Goudela et al., 2008; Krypotou and Diallinas, 2014). In contrast, no distinct transport-related phenotype or plasma membrane localization was detected for AbUapD. Chimeric UapA/AbUapD proteins also led to apparently misfolded proteins and ER retention. Replacement of a critical ~110 amino acids part of UapA with the respective one from AbUapD, also led to a misfolded UapA allele, even though both proteins share high identity at this region and UapA mutations in this region never led to misfolded mutant proteins (Koukaki et al., 2005; Papageorgiou et al., 2008; Vlanti et al., 2006). One can assume that AbUapD includes intrinsic sequences that could lead a protein to misfolding and eventually degradation.

3.3 Functional analysis of the purine/cytosine transporter FcyB

In this chapter we describe a mutational analysis of the *A. nidulans* FcyB transporter and functionally analyzed the relevant mutants in order to identify residues critical for transport function or substrate specificity. FcyB, as described in the Introduction, was already characterized by Vlanti and Diallinas (2008) as a high-affinity, low-capacity, transporter, specific for hypoxanthine, adenine, guanine and cytosine. The crystal structure of Mhp1 (Weyand et al., 2008) allowed the construction of a theoretical model for FcyB, using homology modelling. Our work was supported by the computational analysis performed by E. Mikros' team, and reciprocally our data supported the theoretical model of FcyB.

3.3.1 Homology modelling and substrate docking in FcyB[3]

A theoretical structural model for FcyB was built by V. Kosti and V. Myrianthopoulos. The model was built by homology modeling using as a template Mhp1 (PDB ID: 2JLO). The FcyB and Mhp1 binary alignment showed that they share a rather moderate sequence similarity (18% identity), which is, however, adequate for sustaining a theoretical model of FcyB. The model constructed was validated through Molecular Dynamics analysis. The FcyB model corresponds to an outward-facing conformer that consists of 12 TMSs with mostly helical secondary structures (**Figure 44**). The model can be divided in two distinct domains, a compact core domain consisting of TMSs 1-10 connected with rather short loops and a C-terminal domain consisting of TMSs 11-12; the two domains are connected with a longer loop. The FcyB core domain is composed by two repeats made of TMSs 1-5 and TMSs 6-10. The helices of the two repeats are completely intertwined forming distinct subdomains made of TMSs 1, 2, 6 and 7 (referred also as bundle motif) and TMSs 3, 4, 8 and 9 (referred also as hash motif) linked with TMSs 5 and 10. The helical secondary structure of TMSs 1 and 6 is disrupted in the middle. The substrate binding site is located in the space between the central four-helix bundle and the outer hash motif (**Figure 44B**).

The Mhp1 structure used in homology modeling is occluded with a substrate and thus, residues participating in substrate binding could be identified (Weyand et al., 2008). Through superposition of the FcyB theoretical model constructed and the Mhp1 structural model we located the respective residues that could participate in substrate binding in FcyB. Residues probably participating in substrate binding are highlighted in **Figure 44**.

[3] Details concerning the methodology used in the computational analysis can be found in Krypotou et al. (2012), see Appendix.

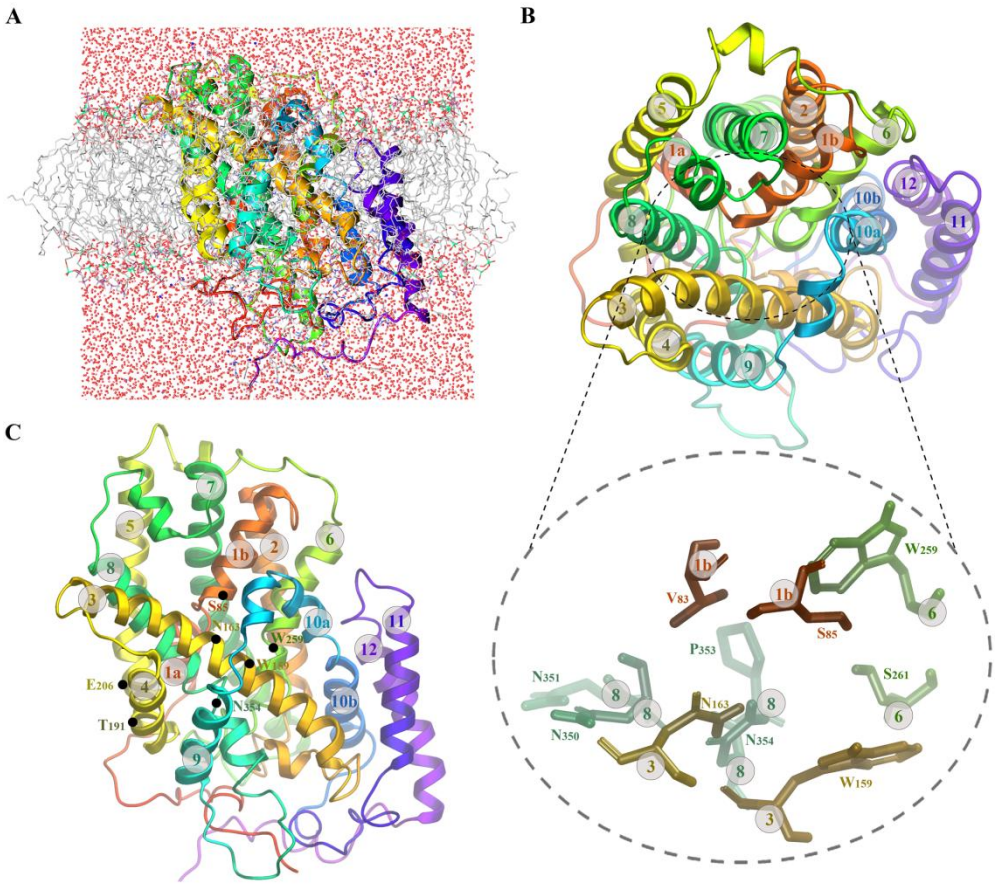
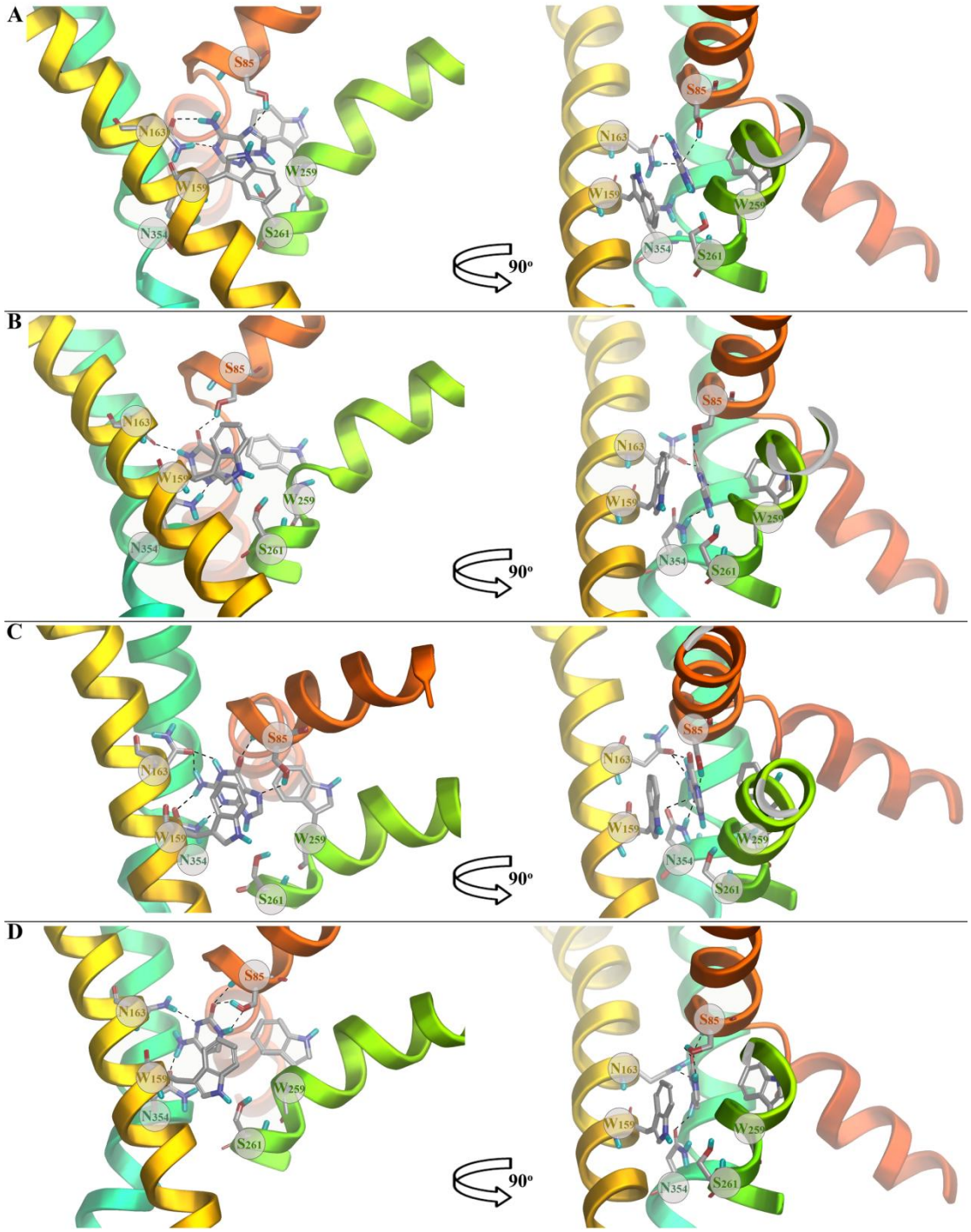


Figure 44. FcyB structural model. (**A**) The modeled three-dimensional structure of FcyB validated with molecular dynamics using the Desmond software. (**B)** Top view of the FcyB model where the core of the first 10 TMSs is clearly distinguished from the last two TMSs 11 and 12. The core can also be seen as two subdomains, the first made of TMSs 1, 2, 6 and 7 and the second of TMSs 3, 4, 8 and 9, linked with the flexible helices TMS5 and TMS10. The substrate binding site is located in the space between the two subdomains of the core. In the lower panel, the topology of probable residues participating in the substrate binding is shown in zoom-in. (**C)** Side view of the FcyB structure showing the topology of probably important for function residues are shown. The computational analysis was performed by V. Kosti and V. Myrianthopoulos.

Induced fit substrate docking (IFD) analysis was carried out by V. Kosti and V. Myrianthopoulos on the FcyB model. All the physiological substrates, namely hypoxanthine, adenine, guanine and cytosine, and the toxic analogue 5-fluorocytosine (5FC) were docked on the model. As shown in **Figure 45**, all substrates were located between two Trp residues, Trp159 and Trp259, which interact through π-π stacking with the purine or pyrimidine ring of each substrate. In addition to the π-π stacking interactions, adenine (K_i 7 μM) was stabilized in the binding site through a bidentate H-bond with the Asn163 side chain and by an H-bond with the Ser85 side chain; hypoxanthine (K_m 11 μM) and guanine (K_i 20 μM) interacted with FcyB through H-bonding with Ser85, Asn163 and Asn354 side chains; cytosine (K_i 17 μM) and 5FC (K_i 145 μM) formed H-bonds with the Ser85, Asn163 and Asn354 side chains. Although cytosine and 5FC were predicted to interact identically with the binding

residues, a lower affinity for 5FC could be due to a steric chlash of the bulky 5FC fluoride with the nearby residues.

In summary, Asn163 (TMS3), Trp159 (TMS3) and Trp259 (TMS6) interacted with all substrates docked in the binding site, Ser85 (TMS1) and Asn354 (TMS8) interactions occurred in all cases except for adenine. Ser261 (TMS6) was also located in the binding site and was in close proximity to the substrates. The computational analysis also predicted that Asn350 and Asn351 (TMS8) formed H-bonds with Asn354, probably providing stability to the Asn354 interactions with the substrates during transport cycle and contributing to the proper architecture of the binding site.



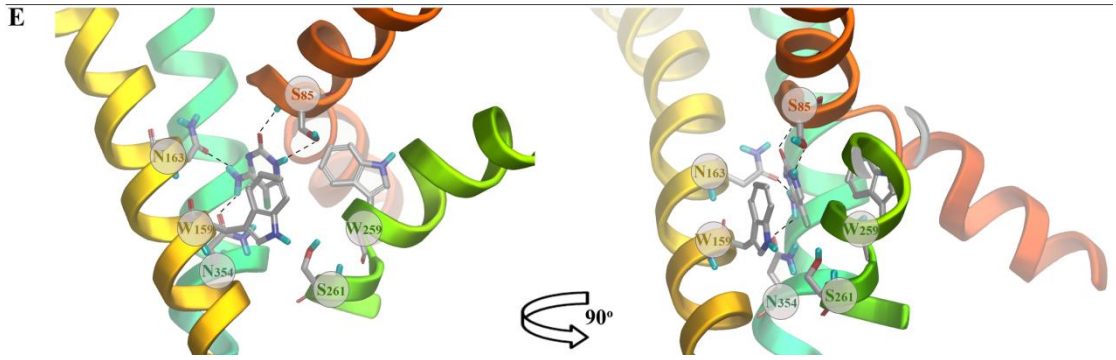


Figure 45. Substrate docking in FcyB. (A) Adenine. **(B)** Hypoxanthine. **(C)** Guanine. **(D)** Cytosine. **(E)** 5-FC. Hydrogen bonds are depicted with black dashed lines. Residues predicted to participate in substrate binding are shown. The computational analysis was performed by V. Kosti and V. Myrianthopoulos.

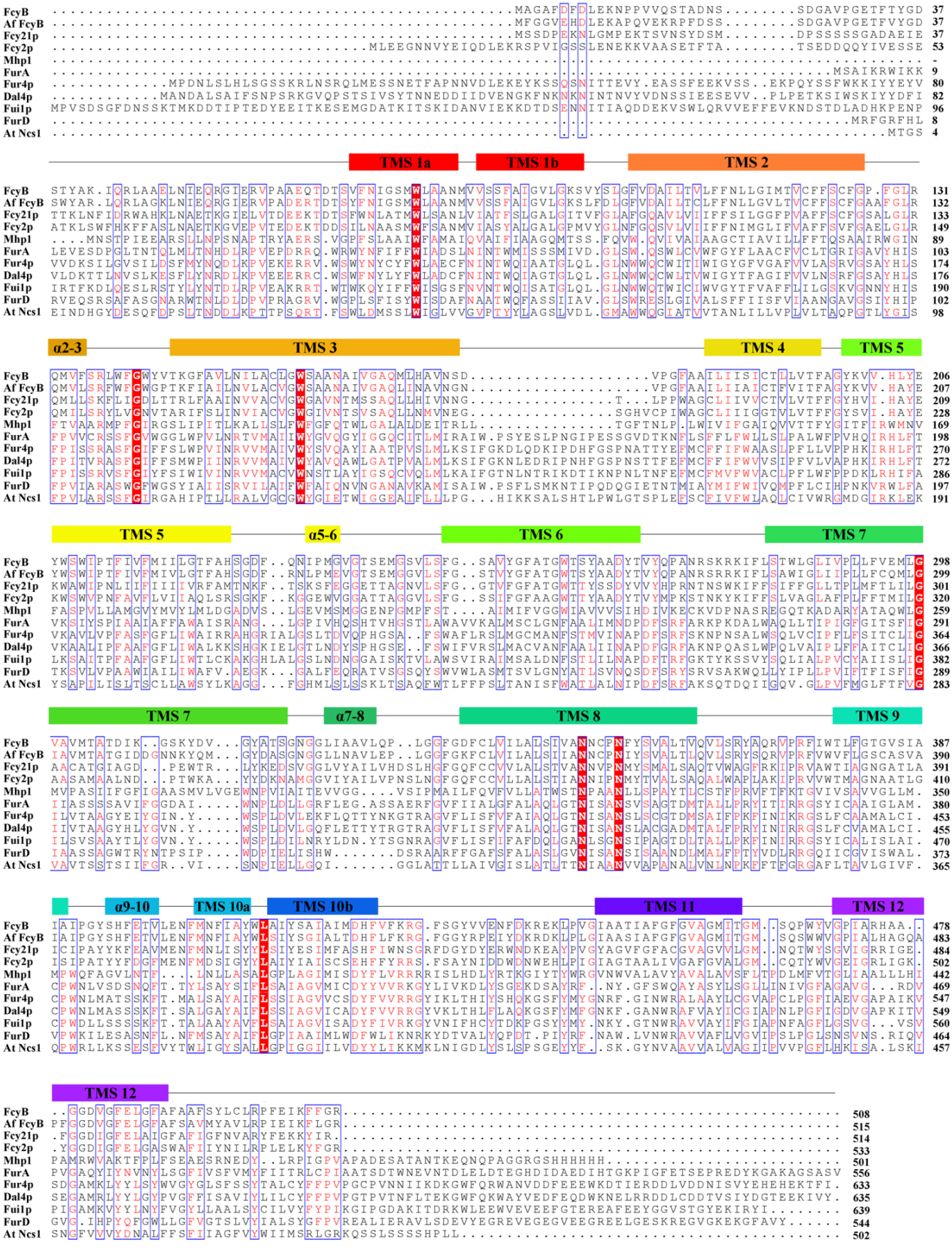
3.3.2 Rational design of FcyB mutants

The residues predicted by the computational analysis to participate in substrate binding were Ser85, Asn163, Trp159, Trp259 and Asn354. As shown in the multiple protein sequence alignment (**Figure 46**) that includes FcyB and characterized homologous NCS1 proteins, Trp159 is absolutely conserved in all proteins, whereas Trp259 is a Trp in Fcy-like proteins and Mhp1 and a Phe/Tyr in Fur-like proteins. Similarly to Trp259, Asn163 is an Asn in all Fcy-like proteins but varies to Gln/Glu/Leu in other homologues. Residues Asn350 and Asn354 are absolutely conserved within all NCS1 proteins, whereas Ser85 varies to Ala/Asn/Pro. Residues conserved through the whole NCS1 family could be critical for the transport function of the proteins and residues varying among proteins with different specificity profile could be involved in substrate specificity

Based on the computational analysis and the multiple sequence alignment we selected amino acid residues that could participate in the transport function and/or determination of substrate specificity. Those residues were replaced with either an Ala residue or an isofunctional or isostructural residue. For example, if a polar amino acid interacts with the substrate and is replaced by an Ala, it should result in reduction of substrate uptake capacity. The mutations devised were: N81A, N81S, S85A, V83N, W159A, N163Q, N163L, T191A, E206A, W259A, S261A, N350Q, N350D, N350A, N351A, P353A, N354Q, N354D and N354A. All mutations were constructed in a modified vector carrying the *fcyB* ORF and *5'* and *3' uapA* regulatory sequences. Studies on *fcyB* native promoter expression have shown low levels of transcription (Vlanti and Diallinas, 2008). Thus, the well-studied *uapA* promoter was used instead as it drives low but significant expression (Gorfinkiel et al., 1993). Transformants were selected by auxotrophy complementation of *pabaA1* or *riboB2* strains, with vectors carrying also *pabaA* of *A. nidulans* or *riboB* of *A. fumigatus*. For plasmid construction details see Materials and methods. The plasmids carrying the *fcyB*

mutant alleles were introduced in null mutant strains that did not express the major purine transporters, namely a Δ7 strain (*ΔfcyB ΔuapA ΔuapC ΔazgA ΔfurD ΔfurA ΔcntA*) for mutations V83N, N163Q, N163L and S261A or a ΔΔ strain (*ΔfcyB ΔuapA ΔuapC ΔazgA*) for the rest mutations; for details see Materials and methods. Transformants carrying single and multiple random plasmid integrations were selected based on Southern blot analysis and were functionally analysed.

Figure 46. Multiple sequence alignment of FcyB and NCS1 homologues of known function. The three-dimensional FcyB model was constructed on the basis of the alignment shown with Mhp1. Putative TMSs of FcyB are denoted in colored cylinders. Invariant and highly conserved amino acids are shaded in red and blue-lined boxes, respectively. The listed NCS1 homologues include: FcyB (*A. nidulans*, GI: 169798762); Af_FcyB (*A. fumigatus*, GI: 169798764); Fcy21p (*S. cerevisiae*, GI: 392860008); Fcy2p (*S. cerevisiae*, GI: 6320897); Mhp1 (*M. liquefaciens*, GI: 210060746); FurA (*A. nidulans*, GI: 2876438); Fur4p (*S. cerevisiae*, GI: 6319495); Dal4p (*S. cerevisiae*, GI: 392860728); Fui1p (*S. cerevisiae*, GI: 6319429); FurD (*A. nidulans*, GI: 149212441); AtNcs1 (*A. thaliana*, GI: 16648836).


FcyB
Af FcyB
Fcy21p
Fcy2p
Mhp1
FurA
Fur4p
Dal4p
Fui1p
FurD
At Ncs1
TMS 1a
TMS 1b
TMS 2
α2-3
TMS 3
TMS 4
TMS 5
α5-6
TMS 6
TMS 7
α7-8
TMS 8
TMS 9
α9-10
TMS 10a
TMS 10b
TMS 11
TMS 12

3.3.3 Functional analysis of the FcyB mutants

The transport function of FcyB becomes detectable by growth testing on adenine and hypoxanthine, only when the major transporter for salvageable purines, AzgA, is genetically inactivated. This is due to a significant overlap in the substrate specificity profile of these two transporters.

A strain expressing a functional *fcyB* via its native promoter in an *azgA* null background confers low growth on MM supplemented with adenine or hypoxanthine as sole nitrogen source (guanine is not usually used in growth tests as it has low solubility and because it is inefficiently catabolised), and is sensitive to 5FC. *fcyB* expressed via the *uapA* promoter leads to increased growth on MM supplemented with adenine or hypoxanthine. Loss of FcyB function leads to absence of growth on adenine and hypoxanthine and resistance on 5FC. Selected single-copy transformants expressing the mutant *fcyB* alleles were tested on adenine, hypoxanthine and 5FC at 37^{o}C and 25^{o}C. In **Figure 47A** growth tests at 25^{o}C are shown (the same results were obtained at both temperatures). From the growth test analysis, we could divide the mutants in the following groups: (*i*) wt-like mutants: not altering growth on either adenine and hypoxanthine, and conserving sensitivity to 5FC (E206A); (*ii*) total loss-of-function mutants: correspond to absence of growth on either adenine and hypoxanthine, and resistance to 5FC (N81A, S85A, W159A, N163L, W259A, N350Q, P353A and N354D); (*iii*) partial loss of function mutants: correspond to equally diminished growth on both purines and partial resistance on 5FC (N81S, V83N and S261A); (*iv*) altered specificity mutants: correspond to growth only on hypoxanthine or adenine but not on both, or partial resistance/sensitivity to 5FC (N163Q, T191A, N350D, N350A, N351A, N354Q and N354A).

GFP-tagged versions of the mutants were constructed to examine their cellular localization. Plasmid construction was performed by cloning of *fcyB* mutant alleles lacking a translation stop codon in a vector carrying GFP (see Materials and Methods). Single copy transformants were selected and analysed. In **Figure 47B** epifluorescence microscopy is shown for all mutants and an FcyB control, carried out at 25^{o}C. FcyB is normally localized in the plasma membrane and in a lower degree in internal structures. Mutants V83N, E206A, W259A, S261A and N350A had a wt-like topology, meaning mostly plasma membrane localization and less in internal structures. In contrast, we detected increased localization in internal structures and none or minimum plasma membrane localization in mutants N81S, N81A, S85A, N163Q, N163L, T191A, N350Q, N351A, N354Q, N354D and N354A. Absence or reduced plasma membrane localization of transporters is usually associated with loss of apparent transport function. The internal structures appeared mostly as rings were possibly due to localization in perinuclear ER. In order to confirm this hypothesis, we stained the nuclei with Hoechst 33342 in a strain expressing FcyB-GFP and N81A-GFP and, also, we constructed a strain expressing FcyB-GFP or N81A-GFP simultaneously with the histone protein HhoA tagged with mRFP (strain produced by genetic crossing with the H1-mRFP strain referred in Edgerton-Morgan and Oakley, 2012) (**Figure 48**). Epifluorescence microscopic analysis revealed that GFP ring-like

structures surround the nuclei in both occasions (Hoechst staining and HhoA tagging with mRFP) that is consistent with localization in the perinuclear endoplasmic reticulum (ER). ER localization is often associated with membrane protein misfolding. We have observed occasional ER localization for the wt FcyB-GFP protein and reduction of the apparent transport function that could be due to GFP tagging. Occasionally GFP can affect protein folding and function, e.g. PrnB proline transporter in *A. nidulans* (Tavoularis et al., 2001). So, concerning internally localized FcyB mutants it is possible that a mutation alone or synergistically to GFP led to partial structural collapse of the protein that resulted in increased ER-retention.

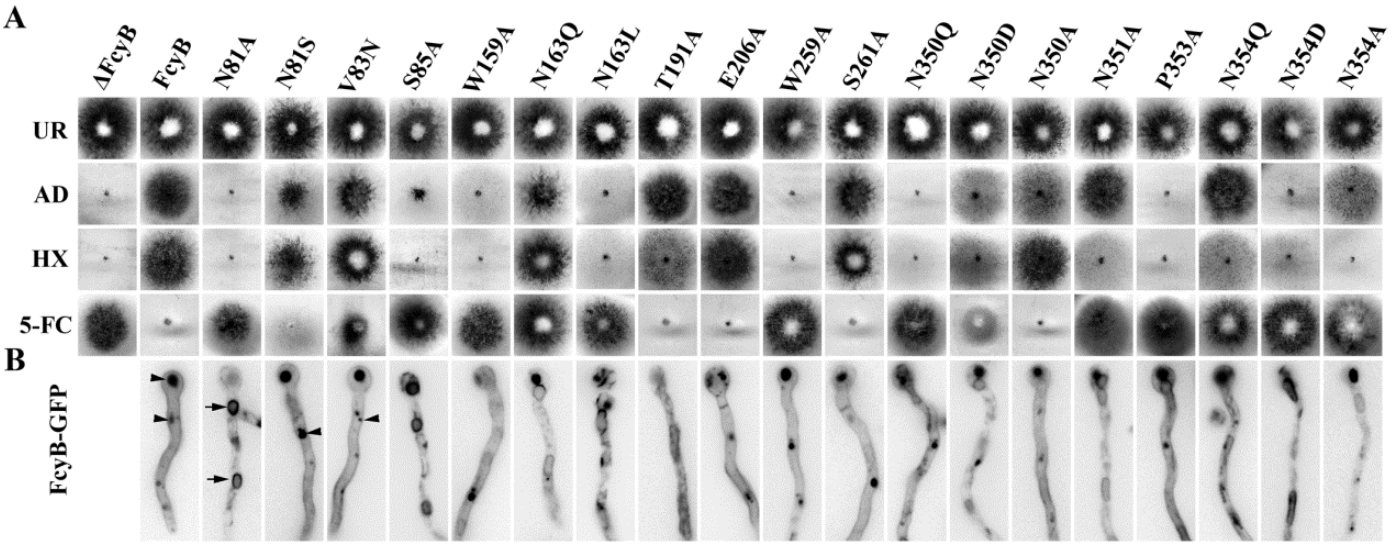


Figure 47. Functional analysis of FcyB mutants. (A) Growth tests of strains carrying single copy FcyB-plasmid integration on purines as sole nitrogen sources and resistance/sensitivity test on 5FC in the presence of $NaNO_3$ at 25°C. AD is adenine, and HX is hypoxanthine. Growth on urea is also shown (UR) as a control. Concentrations used are: UR: 5 mM, AD and HX: 0.1 mg/mL, 5-FC: 50 μM, $NaNO_3$: 10 mM. Positive (FcyB) and negative (ΔFcyB) isogenic control strains are shown. **(B)** Epifluorescence microscopy showing *in vivo* subcellular expression of FcyB-GFP mutant alleles and a wild-type control (FcyB) is presented as dark structures in a grayscale inverted mode. In selected samples, arrows and arrow heads depict perinuclear ER membrane rings and vacuoles, respectivel
y. Strains of single copy FcyB-GFP-plasmid integrations are shown. All samples were prepared from fresh conidiospores inoculated in liquid MM supplemented with $NaNO_3$ (10 mM) and were incubated for 14-16h at 25^{o}C.

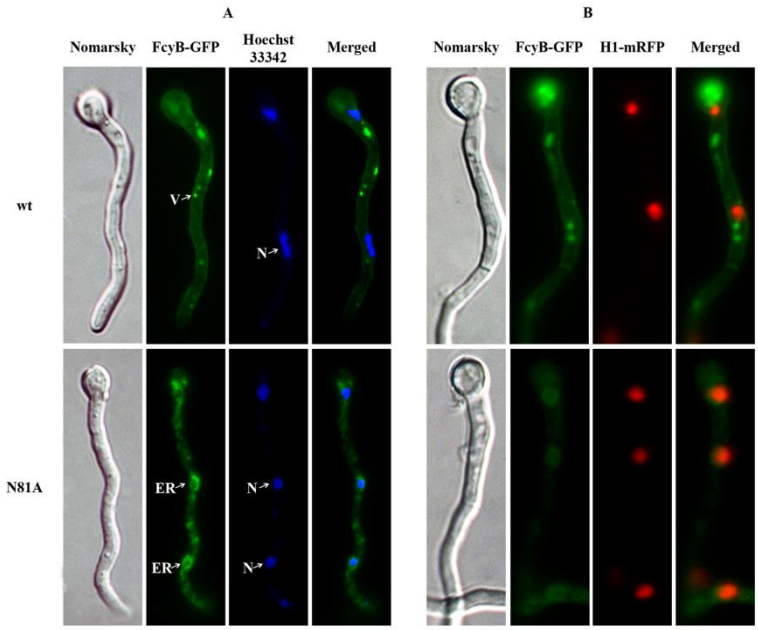


Figure 48. ***In vivo*** **subcellular localization of wild-type FcyB-GFP (upper panel) and mutant FcyB-N81A (lower panel) in relation to positioning of nuclei. (A)** Nuclei stained with Hoechst 33342. **(B)** Nuclei labeled with HhoA-mRFP (histone H1). Samples for microscopic analysis were prepared as before. Nuclei staining was performed according to manufacturer's instructions. V: indicates vacuoles, ER: indicates endoplasmic reticulum, N: indicates nucleus.

Western blot of the FcyB-GFP mutants was performed to monitor protein stability. **Figure 49** shows immunodetection with an antibody against GFP in total protein extracts; actin was used as an internal control for equal loading. Most mutants showed similar or increased protein levels (up to ~6-fold) compared with the wt. Mutants N350A and N354A showed rather lower protein levels compared with the wt, even though N350A localization is similar to wt. For some mutants with ER-retention (N163Q, N163L, N350Q, N351A, N354Q) we observed also increased protein levels, which may be accounted by protection from proteolysis by retention in the ER. Mutants V83N, E206A and W259A had increased protein levels and plasma membrane localization that could be be due to increased protein stability.

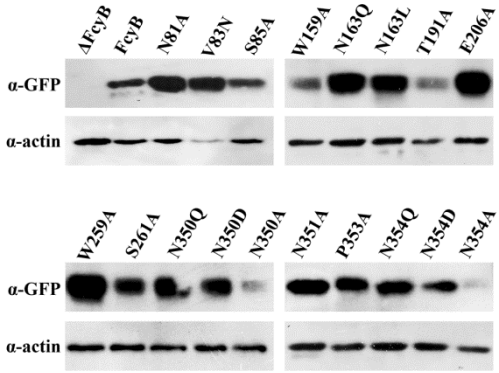


Figure 49. Western blot analysis of total proteins from FcyB-GFP mutants detected with anti-GFP antibody. Antibody against actin was used as an internal marker for equal loading. Cultures for total protein extraction were grown in MM supplemented with urea at 37°C for 8 h.

3.3.4 Kinetic analysis of FcyB mutants

Transport assays were performed using [^{3}H]-hypoxanthine to determine the initial uptake rates (*V*) of the mutants. All uptake experiments were performed during linear uptake phase of FcyB, that is for 1 min at 37^{o}C (Vlanti and Diallinas, 2008). As shown in **Figure 50**, all mutants, except E206A, showed decreased uptake rate compared with the wt. Mutants V83N, N163Q, T191A, S261A and N350A preserved ~30-50% of their *V*; mutant E206A showed 1.5-fold increased transport rate; while the remaining mutations led to <10% transport rate compared with the wt. These results are in good agreement with the growth test analysis, as only mutants with an uptake rate >30% grew on hypoxanthine (see **Figure 47**), similarly to other transporters studied previously in our lab. In some mutants reduced transport capacity could be associated with low plasma membrane localization and increased ER-retention (see **Figure 47**).

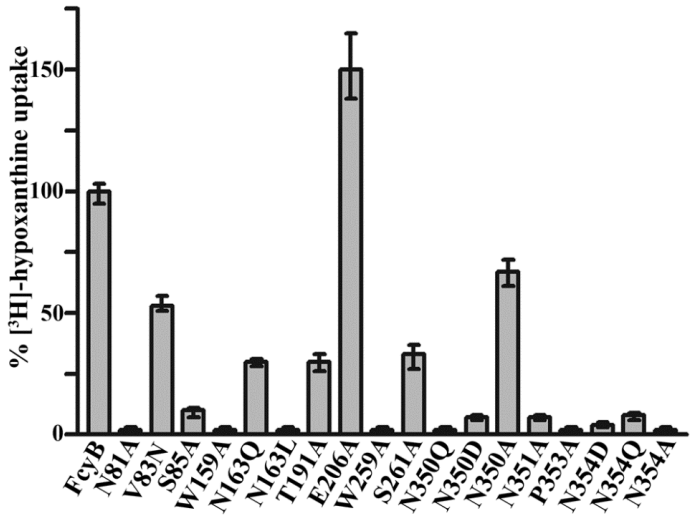


Figure 50. Apparent transport rates for FcyB mutants. Comparative initial uptake rates of [^{3}H]-hypoxanthine for 1 min at 37^{o}C in FcyB mutant alleles and a wild type (wt) (FcyB) control. 100% is the transport rate of the wt. Results are averages of at least three independent experiments in triplicate for each mutant. Strains used in the uptake assays correspond to the ones used for the growth test analysis, carrying single copy FcyB-plasmid integrations.

We then determined the binding and inhibition values ($K_{m/i}$) for the four physiological substrates (hypoxanthine, adenine, guanine, cytosine) of FcyB for mutants with measurable uptake rate (**Table 9**). For mutants with low hypoxanthine uptake rate (< 20%) we used multi-copy transformants (2-4 plasmid integrations). This was necessary because of the difficulty to measure precisely rates of uptake if they are too low, and stressing that increasing protein concetration in the membrane would change the V_{max}, but not the $K_{m/i}$. Mutations V83N, T191A, E206A, S261A and N351A led to minor changes (maximum up to ~2-fold increase or decrease) of the $K_{m/i}$ for one or

more substrates in each case; e.g. V83N led to ~2-fold higher K_i for adenine, but ~1.5-fold lower K_m for hypoxanthine, while retaining the same K_i for cytosine. On the contrary, replacement of the two Asn, Asn163 and Asn354, with a Gln and a Glu respectively, both led to ~30-fold decrease in K_i for guanine. Moreover, N163Q led to ~3-4-fold lower $K_{m/i}$ for hypoxanthine and cytosine and ~3-fold higher K_i for adenine. Mutation S85A led to significant decrease (~2-7-fold) in $K_{m/i}$ for all substrates except guanine that remained unchanged. Interestingly, mutation N350A led to ~2-fold higher K_m for hypoxanthine, 4-fold higher K_i for cytosine and ~1.5-fold lower K_i for guanine, but no measurable K_i for adenine (>1000 μM) that was also reflected in the growth test analysis (wt-like growth on hypoxanthine, but leaky growth on adenine; see also **Figure 47**).

Table 9. Kinetic and specificity profile of mutant versions of FcyB. $K_{m/i}$ values (μM) were determined as described in Materials and methods. >500 or >1000 stand for inhibition values close to 10-20% at 0.5 and 1 mM respectively. Results are averages of at least three independent experiments in triplicates for each concentration point. Standard deviation was <20%. HX: hypoxanthine, AD: adenine, GU: guanine, CY: cytosine. For mutants with low V (< 20%) multi-copy strains (2-4 plasmid integrations) were used.

	K_m (μM)	K_i (μM)		
	HX	**AD**	**GU**	**CY**
wt	11	7	17	20
V83N	16	2	7	18
S85A	78	26	16	45
N163Q	47	2	> 500	73
T191A	9	11	4	3
E206A	6	7	10	14
S261A	4	6	50	17
N351A	3	1	4	16
N350D	5	8	31	6
N350A	6	> 1000	25	5
N354D	3	1	> 500	4

In summary, the computational analysis combined with the mutational and functional analysis led to identification of several amino acids as critical for FcyB function. The amino acids thus identified are: Asn81, Ser85 (TMS1), Trp159 (TMS3), Trp259 (TMS6) and Pro353 (TMS8) as absolutely necessary for function; Asn163 (TMS3), Asn350 and Asn354 (TMS8) as critical for function and substrate specificity; Thr191 (TMS4), Ser261 (TMS6) and Asn351 (TMS8) as important for wt-like transport. Interestingly, Thr190 and Asn351 could contribute specifically to hypoxanthine uptake indirectly, as T190A and N351A mutations affect hypoxanthine uptake (no growth and low *V*) but do not affect significantly the binding affinity for any substrate. Some of the residues were proven to interact directly with the substrates, but others could participate in retaining the proper architecture of the binding site or set in

position neighboring critical residues in order to succeed substrate binding and finally transport.

3.3.5 Design of FcyB-specific novel inhibitors

The FcyB model constructed during this study was used by the teams of E. Mikros, P. Marakos and N. Pouli in the rational design and synthesis of novel purine analogues that may act as substrates/inhibitors for FcyB and potential antifungals. An efficient antifungal targets specifically the pathogenic fungus and not the host organism. Purine analogues are often considered as potential antifungals as they can lead to toxicity or loss of growth by inhibiting purine uptake or metabolism.

In particular, several new hypoxanthine- and adenine-analogues were synthesized and 27 were used in competition uptake assays with radiolabeled hypoxanthine in a strain expressing only *fcyB* under the regulation of *uapA* promoter, lacking all known transporters related to purines, pyrimidines or nucleosides (*uapA*$_p$-*FcyB-pabaA ΔuapA ΔuapC::AfpyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100,* see Materials and methods). As shown in **Figure 51** analogues 1549, 1552, 1682 and 1683 caused 80-95% inhibition of the [^{3}H]-hypoxanthine uptake indicating a binding affinity probably < 200 μM. Further kinetic experiments to determine the inhibition affinities (K_is) of these analogues, as well as physiological tests to check whether they possess antifungal properties, will define whether these analogues could be used as antifungal drugs.

Interestingly, the tick-borne bacterium *Borrelia burgdorferi* that causes the multistage inflammatory Lyme disease, lacks the enzymes responsible for purine synthesis and its pathogenicity was proved to depend on the uptake of purines from the host by two AzgA like proteins (Jain et al., 2012). Thus, the AzgA-like proteins of *B. burgdoferi* are suitable targets of purine analogues that can block the uptake of purines. In an effort to identify a potential inhibitor for AzgA among the purine analogues synthesized for FcyB, we tested them in competition uptake assays with radiolabeled hypoxanthine in a strain expressing *azgA* and lacking *fcyB*, but no inhibition was detected for any of the analogues (not shown). Apparently, these analogues can specifically inhibit FcyB-mediated hypoxanthine uptake and not AzgA, and thus they could be tested in pathogens expressing Fcy-like proteins.

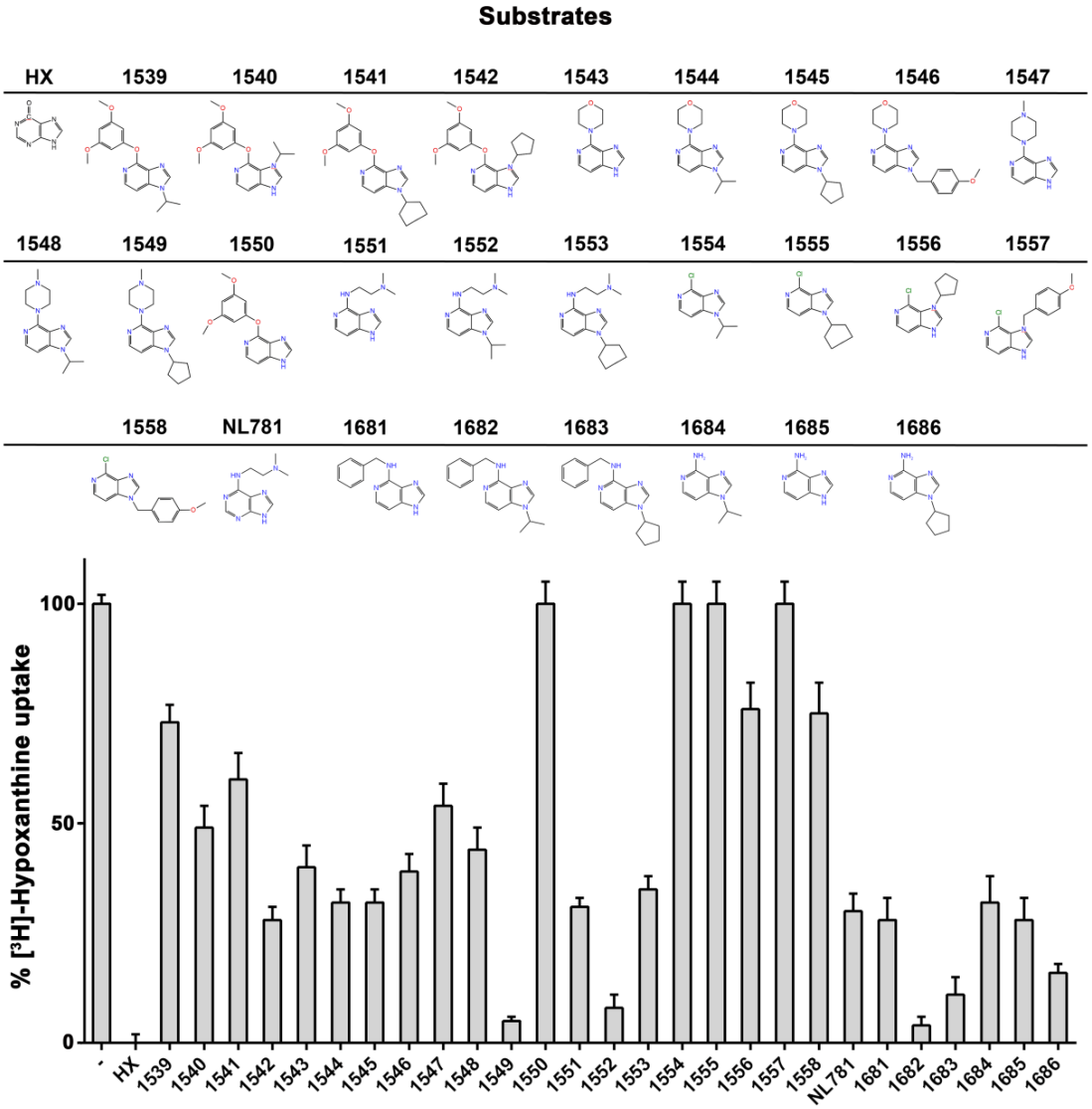


Figure 51. FcyB-mediated [^{3}H]-hypoxanthine uptake in the presence of various purine analogues. A multiply deleted strain, Δ7, expressing *fcyB* under *uapA* promoter was used. The chemical formula of each inhibitor is shown on the top part of the Figure. The reference numbers shown for each inhibitor in the Figure correspond to the ones deposited in the chemical library of School of Pharmacy, University of Athens. HX: stands for hypoxanthine and –: stands for no inhibitor.

3.4 Characterization of the orphan Fur proteins and functional analysis of FurD and FurE

The Fur family includes 7 members, as described in the Introduction, with FurD and FurA already characterized. In the current study, we attempted to characterize through over-expression the rest 5 Fur proteins of *A. nidulans*. Structural models for FurA, FurD and FurE, built by E. Mikros and T. Evangelidis, were used in order to predict residues that interact with substrates. Subsequently, a mutational analysis in FurD and FurE was performed in order to identify residues critical for the function or/and specificity.

3.4.1 Over-expression of the orphan Furs

Out of the seven Fur proteins only two (FurA and FurD) led to distinct phenotypes by null mutations (Hamari et al., 2009). Northern analysis performed in Hamari et al., (2009) showed under the conditions tested, that *furC*, *-E*, *-F* and *-G* are very poorly transcribed. Since low expression was detected along with the absence of distinct phenotypes, over-expression of the *fur* genes might lead to detectable phenotypes. Thus, all *fur* ORFs were amplified and cloned into a pGEM modified vector that carried the promoter of *gpdA* and the *3'* terminating region of *trpC*. GpdA is a glyceraldehyde-3-phosphate dehydrogenase with a well-studied promoter that is constitutive and results in high expression (Punt et al., 1988, 1990), and thus is often used to increase expression of poorly transcribed genes. TrpC is a trifunctional protein involved in tryptophan biosynthesis with well-studied gene structure and regulation. The *trpC* terminating region is often used in expression systems as it confers stability to the cognate mRNA produced due to a long poly-A tail (Mullaney et al., 1985). The vector used also carried the *pantoB* (*panB*) gene that complements the *pantoB100* auxotrophy for pantothenic acid. All plasmids were introduced in the Δ7 strain, a multiple deleted strain that lacks all major transporters for purines, pyrimidines and nucleosides (*ΔuapA ΔuapC::pyrG ΔazgA ΔfcyB::argB ΔfurD::riboB ΔfurA::riboB ΔcntA::riboB pabaA1 pantoB100*). The use of this strain allows to detect easily distinct phenotypes related with purine or pyrimidine uptake resulting from the overexpression of the relevant *fur* gene. Transformants selection was based on the complementation of *pantoB100* auxotrophy. For details on plasmid construction and strains see Materials and methods.

Selected single-copy transformants were analyzed by growth testing on various purines and toxic analogues at 37^{o}C and 25^{o}C. Growth tests at 25^{o}C and 37^{o}C are shown in **Figure 52**. In agreement with previous observations (Hamari et al., 2009), *furD* over-expression led to growth on uric acid and xanthine at 37^{o}C and increased sensitivity to 5-fluorouracil (FU), FC and 5-fluorouridine (FUd). *furA* over-expression led to toxicity on allantoin, probably due to over-accumulation of this substrate that is known to cause toxic effects (Darlington and Scazzocchio, 1967). Interestingly, over-expression of *furA* led to sensitivity in FU which was not expected.

Over-expression of *furE* led to growth on allantoin and uric acid and sensitivity to FU, FC, FUd and oxypurinol (OX). For *fur-C, -F* and *-G* we detected moderate sensitivity in FU in the order *furC* > *furG* > *furF*. Finally, for *furB* we could not detect any phenotype at conditions tested.

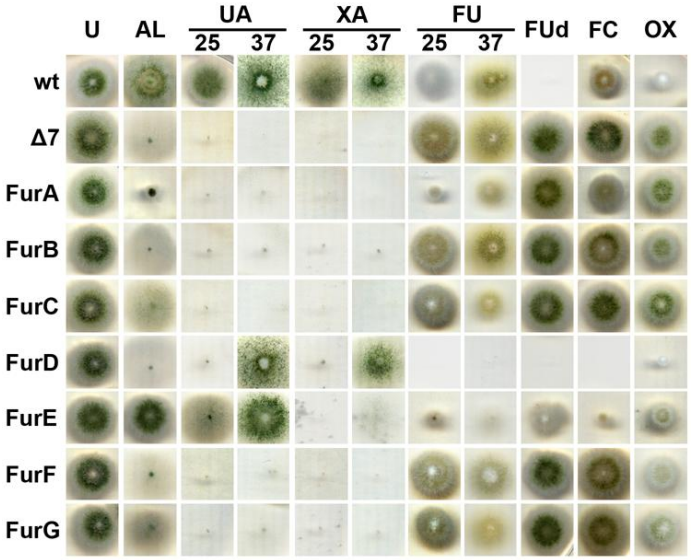


Figure 52. Growth tests of transformants overexpressing Fur transporters. Nitrogen sources or toxic analogues used are: urea (U; 5 mM), allantoin (AL; 0.1 mg/mL), uric acid (UA; 0.1 mg/mL), xanthine (XA; 0.2 mg/mL), 5-fluorouracil (FU; 50 μM), 5-fluorouridine (FUd; 10 μM), 5-fluorocytosine (FC; 50 μM) and oxypurinol (OX; 100 μM). Growth tests shown were at 25°C, or at 37°C when indicated. Toxic analogues were used in media containing urea as a nitrogen source. All strains used are single copy plasmid integration transformants. A wild-type strain was used as a positive control and Δ7 as a negative control.

3.4.2 Construction and analysis of multiple *fur* deleted strains

Over-expression of the *furs* revealed that Fur-C, -E, -F and -G are functional, each contributing to 5-flourouracil (5FU) sensitivity, and in the case of FurE, to growth on allantoin and xanthine. The minor sensitivity on 5FU probably reflects the low transport capacity or/and affinity of Fur-C, -F, and -G for this toxic pyrimidine analogue. Inability to detect any phenotype in strains carrying null mutations could be due to low expression of the corresponding genes, at least at conditions tested, in combination with low transport capacity. We constructed multiple null *fur* mutants by genetic crossing in order to test FU resistance when the major uracil transporter, FurD, is also absent. The strains were constructed in collaboration with J. Bobonis.

Double null mutants (*ΔfurDfurA*, *ΔfurDfurC*, *ΔfurDfurE*, *ΔfurDfurF*, *ΔfurDfurG*, *ΔfurEfurA*) and triple mutant *ΔfurDfurAfurE* were constructed and tested for resistance on 5FU by growth testing at 37^{o}C and 25^{o}C (same results obtained under both conditions). As shown in **Figure 53A**, *furD* knockout was highly resistant to FU, while the rest single null mutants were sensitive and more or less similar to the wt. The double mutants were resistant to 5FU in a degree similar to *ΔfurD*; minor differences in resistance were detected but were not consistent, so no secure

conclusions could be deduced. So, even though all Fur transporters (except FurB) contribute to 5FU uptake when over-expressed, only FurD seems to clearly contribute to 5FU uptake when Furs are expressed with their native promoters.

In order to further investigate whether Fur transporters contribute to uracil uptake and since uracil cannot be used as a nitrogen source, we constructed the same combinations of the null *fur* mutants in a *pyrG89* background (strain constructions were performed in collaboration with J. Bobonis). PyrG is an orotidine-5'-phosphate decarboxylase which catalyzes the conversion of orotidine monophosphate (OMP) to uridine monophosphate (UMP) (David et al., 2008). The *pyrG89* mutation blocks pyrimidine biosynthesis and leads to uracil requirement (Palmer and Cove, 1975) so that growth of *pyrG89* mutants depends on the efficiency of uracil accumulation from the media.

The *pyrG89* null mutants were tested on MM supplemented with uracil, growth on uridine was used as a control, at 37^{o}C and 25^{o}C (both conditions resulted in the same phenotypes). As shown in **Figure 53B**, Δ*furD* was the one with the least growth among the single null mutants, the rest had moderately less growth compared with the wt. In contrast, double null mutants Δ*furDfurE* and Δ*furDfurF* grew less compared to the Δ*furD* and the triple Δ*furDfurAfurE* mutant grew by far less than all the strains tested.

We finally tested the double *ΔfurAfurE* strain on allantoin, as we already were aware of the leaky growth of Δ*furA* on allantoin and that FurE is capable of recognizing and transporting allantoin when over-expressed. In **Figure 53C** is apparent that the leaky growth on allantoin is absent in a *ΔfurAfurE* genetic background.

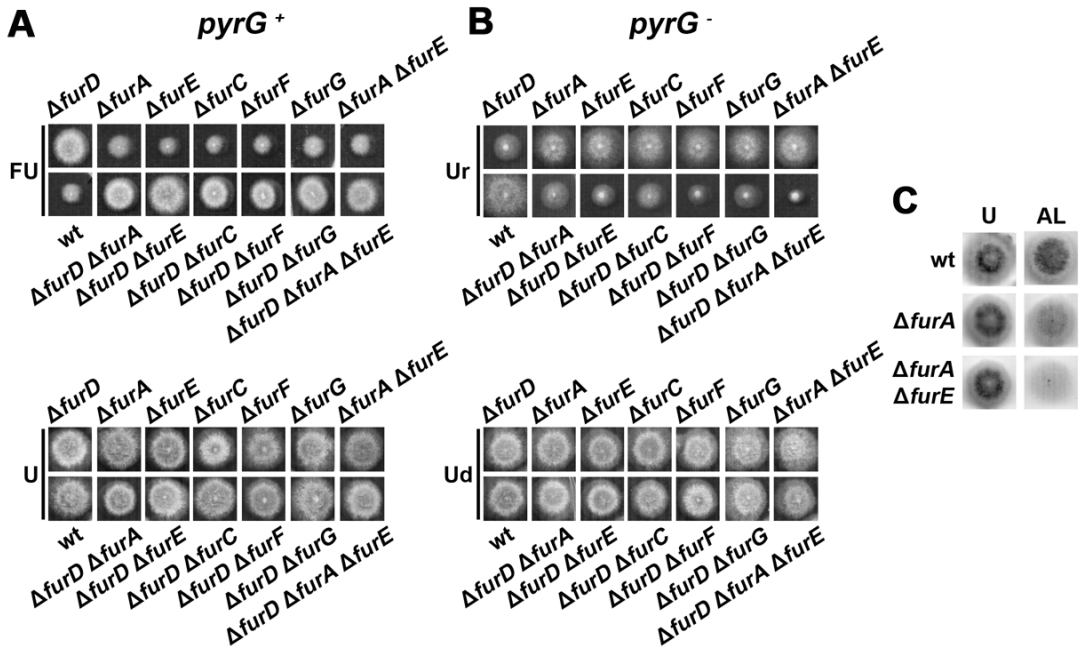


Figure 53. Growth tests of simple, double or triple null *Δfur* mutants. (A) *pyrG*$^{+}$ strains tested in 5-fluorouracil (FU) in presence of urea (upper panel). Growth on urea is shown as a control (lower panel).

(B) *pyrG89* (*pyrG*$^{-}$) strains grown in media supplemented with 2.5 mM uracil (Ur), or 10 mM uridine (Ud) as a control. In both cases urea was used as the sole nitrogen source. Growth tests shown are at 25°C. **(D)** Growth tests of single *ΔfurA* and double *ΔfurA ΔfurE* null mutants in the presence of 2 mM allantoin (AL) or urea (U) as sole nitrogen sources. Growth tests shown are at 25°C. Wt strains are shown in A, B and C as controls.

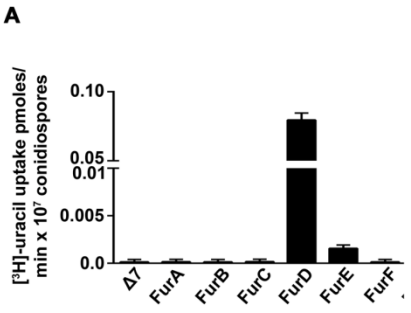
In summary, *fur* over-expression revealed that FurA, FurC, FurD, FurE, FurF and FurG contribute to 5FU uptake in the following order FurD ⩾ FurE, FurA > FurF, FurC, FurG. When Furs are expressed by their native promoters FurD is the major transporter for uracil, while the rest of Furs (except FurB) have minor contribution in uracil, mostly apparent in a *pyrG89* background. Finally, FurE was shown to be specific for the uptake of allantoin, uric acid and the toxic analogues 5FC, 5FUd and oxypurinol.

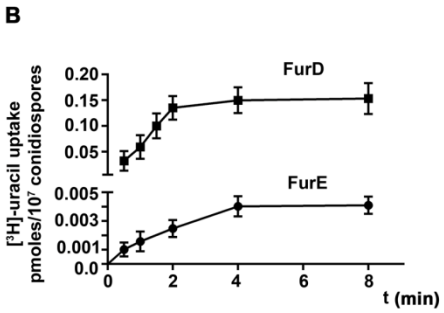
3.4.3 FurE kinetic profile

Transport assays using [^{3}H]-uracil were performed in all strains over-expressing Fur transporters. As shown in **Figure 54A**, besides from FurD-mediated uracil uptake that was, as expected, significant low but distinct FurE-mediated uptake was also detected. Even though we proved that the remaining Furs, except for FurB, can contribute to 5FU or/and uracil uptake, their direct involvement in [^{3}H]-uracil transport could not be detected. Given that the radiolabeled substrate used was at a concentration of 0.1 μM, this might be a concentration not sufficient to detect uptake of low K_m or/and capacity Fur transporters.

We performed time-dependent transport assays for FurE to establish its kinetic profile. In **Figure 54B**, is shown that FurE-mediated [^{3}H]-uracil transport is time dependent, as it is linear at first and reaches a plateau after 4-5 min; a profile similar to most transporters studied so far.

The binding and inhibition kinetic values ($K_{m/i}$) were determined for FurE recognized substrates (**Figure 54C**). FurE binds with low affinity uracil and allantoin (K_m ~1 mM) and with high affinity uric acid (20 μM), while no significant inhibition was detected for xanthine, hydantoin or thymine. The low measured uptake could be due to the low affinity for the radiolabeled substrate; but this does not controvert the results obtained as proper controls were used at all times. We finally, tested the H^+-uncoupler carbonylcyanide chlorophenylhydrazone (CCCP) and found that [^{3}H]-uracil uptake is H^+ dependent.





C

FurE specificity profile

Inhibitor (1mM)	[^{3}H]-uracil uptake (%)	$K_{m/i}$ (μM)
-	100	-
UR	45±5	~1000
AL	40±8	~1000
UA	27±3	20
XA	100±5	>1000
HY	74±10	>1000
TH	84±10	>1000
CCCP	0	-

Figure 54. Kinetic analysis of FurE transporter. (A) Uracil transport assays. Initial uptake rates of [^{3}H]-radiolabeled uracil (0.1 μM) for 1 min were measured in transformants over-expressing Fur transporters. Standard deviation is depicted with error bars. **(B)** Time course of FurE-mediated [^{3}H]-uracil uptake (0.1 μM). Time course of FurD is shown as a control. Standard deviation is depicted with error bars. **(C)** Specificity profile of FurE. Middle column: FurE-mediated [^{3}H]-uracil uptake in the presence of excess (1 mM) unlabelled uracil (UR), allantoin (AL), uric acid (UA), xanthine (XA), hydantoin (HY) or thymine (TH), or in the presence of the H^+-uncoupler carbonylcyanide chlorophenylhydrazone (30 μM, CCCP). Standard deviation values are shown. Right column: Approximate $K_{m/i}$ (μM) values of putative inhibitors. ~1000 or > 1000 means that at K_i values (inhibition close to 50%) are close to 1 mM or > 1 mM.

3.4.4 Transcriptional profile of FurE

Previous studies have shown that expression of *furE* mRNA from its native promoter is hardly detectable in conidiospores grown for 3 h at 37^{o}C in MM supplemented with urea (Hamari et al., 2009). In addition, transcriptomic analysis RNA-seq was performed by Sibthorp et al. (2013) from *A. nidulans* cultures grown in the presence of various nitrogen sources or under nitrogen starvation and showed that *furE* mRNA in the presence of neutral nitrogen sources or ammonium is marginally detectable, but its expression is upregulated upon nitrogen starvation.

We performed Northern analysis for *furE* under various conditions (nitrogen sources, temperature, pH) (**Figure 55**). In more detail, conidiospores of a wt strain were grown for 16 h at 37^{o}C and pH 6.8 in the presence of different nitrogen sources (NO_3, allantoin, uric acid) and uracil (in the presence of NO_3), or after reaching full growth followed by 4 h nitrogen starvation. In addition, conidiospores were grown for 16 h at 25^{o}C or 42^{o}C (pH 6.8) or at pH 5 or 8 (temperature 37^{o}C). As shown in **Figure 55**, *furE* is marginally detectable at 37^{o}C in the presence of nitrate, allantoin, uracil or at pH 8 and at 42^{o}C, but is not detected in the presence of uric acid, at pH 5 and at 25^{o}C. *furE* expression is highly upregulated under nitrogen starvation, a common phenomenon in nitrogen related transporters studied in *A. nidulans* (Amillis et al., 2004; Vlanti and Diallinas, 2008).

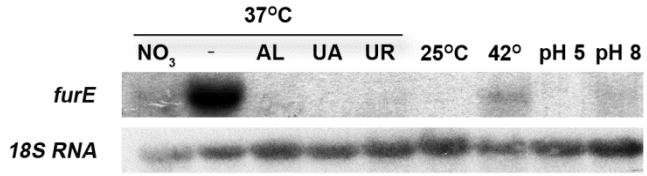


Figure 55. *furE* northern blot analysis. Total RNA extracts were analyzed and *furE* was detected with a specific probe. Cultures were incubated for 16 h at 37°C and pH 6.8 in the presence of nitrate (NO_3), or allantoin (AL), or uric acid (UA), or uracil (UR). Nitrogen starvation (-) was performed after collecting the mycelium from a culture supplemented with NO_3 grown for 16 h at 37°C and pH 6.8 in a new MM without any nitrogen source. Cultures at 25°C and 42°C (pH 6.8) or pH 5 and pH 8 (37°C) grown for 16 h were also performed. As an internal control for equal loading *18S RNA* was used.

3.4.5 Fur protein localization and stability

We constructed GFP tagged versions of the Fur proteins and used epifluorescence microscopy for studying their subcellular localization. As shown in **Figure 56A**, Fur-A, -C, -D, -E, -F and -G were localized in the plasma membrane. However, in the case of FurE and FurF some vacuolar localization was also present. In addition, in the case of FurF low fluorescence was detected. We never detected plasma membrane localization for FurB, which appeared to be mainly localized in cytoplasmic structures and foci beneath the plasma membrane, including probably perinuclear ER; thus, lack of FurB localization in the plasma or other membrane was compatible with absence of any characteristic phenotype related to transport, whether assessed by deletion or by overexpression. We also tested conditions triggering endocytosis in purine/pyrimidine transporters, namely presence of ammonium or substrate (Gournas et al., 2010). For Fur transporters, previous data (Hamari et al., 2009) and our results proposed as putative triggers for substrate endocytosis uracil, allantoin and uric acid. **Figure 56A** shows that FurA, FurD, FurE and FurF are sensitive to ammonium endocytosis in the following order FurD < FurA < FurE, FurF. FurE appears to be highly but differentially sensitive to substrate endocytosis, in an order that was uracil < uric acid < allantoin. FurD and FurF were moderately sensitive to uracil and FurA to allantoin. Interestingly, FurC and FurG were absolutely resistant to ammonium and substrate endocytosis.

We performed western blot analysis for the Fur proteins to estimate their turnover. Free GFP was used as an indicator for vacuolar degradation. The analysis, shown in **Figure 56B**, revealed that all Fur proteins (except FurB and FurF) are rather stable proteins. FurC and FurG appear as the most stable proteins, as their ratio Fur-GFP:free GFP is 1.7 and 5.2 respectively. FurA and FurD are rather stable with protein amounts of Fur-GFP similar to free GFP. FurE-GFP and FurF-GFP, as shown also by the microscopic analysis, are the ones with most increased vacuolar degradation. We did not detect any free GFP for FurB-GFP, while the protein amounts of intact FurB-GFP and FurF-GFP, were significantly lower compared to the rest Furs.

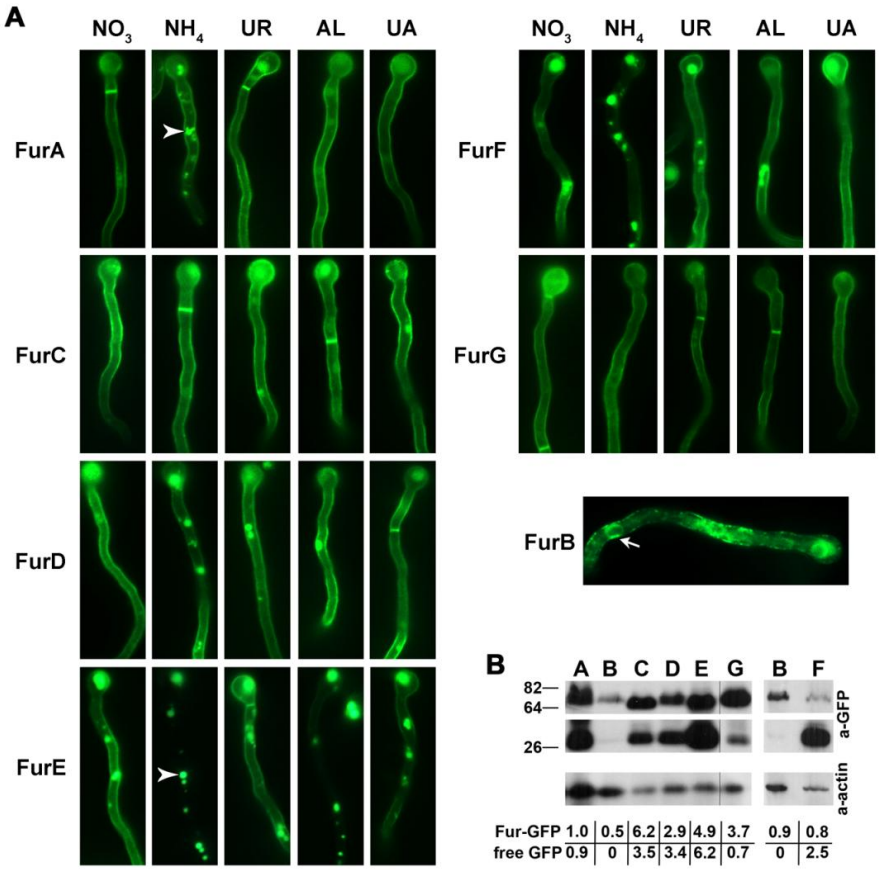


Figure 56. *In vivo* subcellular localization and protein expression levels of Fur transporters. (A) Epifluorescence microscopy of GFP-tagged Fur proteins expressed from the *gpdA* promoter. *Arrows* and *arrowheads* depict vacuoles and perinuclear endoplasmic reticulum (ER) membrane rings respectively. (UR: uracil, AL: allantoin, UA: uric acid). Cultures were grown in MM for 14–16 h at 25°C. **(B)** Western blot analysis of total protein extracts of strains expressing GFP-tagged Fur proteins using an anti-GFP antibody. At the left panel, all lanes correspond to the same blot. FurF is shown in a separate blot with FurB as an internal control. Antibody against actin was used as an internal marker for equal loading. Molecular weight markers in kDa are shown. Relative amounts of Fur-GFP and free-GFP, normalized to actin protein levels, are shown below the blots. Cultures for total protein extraction were grown in MM supplemented with urea at 25°C for 16 h.

3.4.6 FurB and AbFurB2 subcellular localization

As shown from the analysis presented above, FurB was detected in cytoplasmic foci and not in the plasma or any other membrane, and in addition, no transport function could be assigned to it. *In silico* analysis, performed by G. Diallinas and C. Scazzocchio, showed that FurB is present only in *A. nidulans* and *A. clavatus* (74% identity among them), whereas a rather close homologue (~54-56% identity), named FurB2, is present in few more Aspergilli (*A. carbonarius*, *A. niger*, *A. tubingensis*, *A. kawachii*, *A. acidus*, *A. aculeatus*, *A. brasiliensis*, *A. glaucus*, *A. terreus*). In order to check the functionality of the FurB2 subgroup, we selected FurB2 from *A. brasiliensis*, named herein AbFurB2, and expressed it in *A. nidulans*. *AbfurB2* was amplified from *A. brasiliensis* genomic DNA and cloned, in-frame with *gfp,* in the same vector used

to study all other Furs. The cognate plasmid was introduced, as previously described, in the Δ7 strain (for details see Materials and methods).

Figure 57 summarizes the functional analysis performed for a selected single-copy transformant. Growth test analysis was performed on various purines, used as nitrogen sources, or purine and pyrimidine analogues. As shown in **Figure 57A**, the strain expressing *AbfurB2* was moderately sensitive to 5FU and showed no distinct phenotype under any other condition tested (not shown). The AbFurB2-GFP localization was detected, using epifluorescence microscopy, mainly in the plasma membrane and at lower amounts in vacuoles (**Figure 57B**). This is a typical localization of a transporter, and obviously contrasts with FurB. We also used conditions triggering endocytosis (ammonium and uracil) and observed increased vacuolar localization only when ammonium was used. Finally, a western blot was performed for total protein extracts from strains expressing *AbfurB2-gfp* and *furB-gfp* (**Figure 57C**). AbFurB2-GFP protein levels were clearly higher than those of FurB-GFP; free-GFP, an indicator for vacuolar degradation, was also present for AbFurB2-GFP, whereas no free-GFP was detected in the strain expressing FurB-GFP.

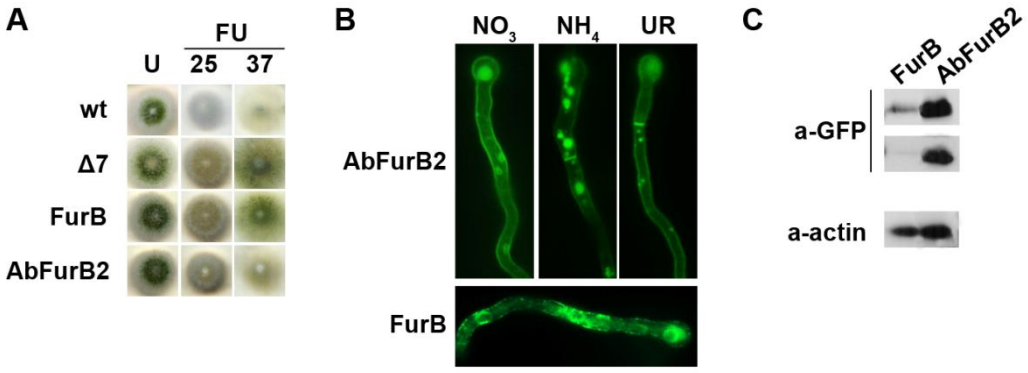


Figure 57. Functional characterization of AbFurB2 in comparison with FurB. (A) Growth test at 25°C and 37°C on MM containg FU in the presence of urea (U). **(B)** Epifluorescence microscopy of strains expressing AbFurB2-GFP and FurB-GFP in the presence of nitrate (NO_3), ammonium (NH_4) or uracil (UR). Cultures were grown in MM for 14–16 h at 25°C. **(C)** Western blot analysis of total protein extracts of strains expressing GFP-tagged AbFurB2 and FurB proteins using an anti-GFP antibody. Antibody against actin was used as an internal marker for equal loading. Molecular weight markers in kDa are shown. Cultures for total protein extraction were grown in MM supplemented with urea at 25°C for 16 h.

In summary, even though AbFurB2 and FurB share 56% identity on protein sequence level, FurB appears as a non-functional protein with localization in cytoplasmic structures, while AbFurB2 corresponds to a 5FU specific transporter with the expected plasma membrane localization and turnover. We performed protein sequence alignment of the two proteins and we did not identify any apparent differences that could explain the "instability" of FurB. Thus, it was tempting to conclude that FurB might contain intrinsic sequences that probably cause protein

misfolding, reflected as lack of proper localization and rapid turnover (See Discussion).

3.4.7 Homology modeling, molecular dynamics and substrate docking for FurA, FurD and FurE[4]

Structural models for FurA, FurD and FurE were built by E. Mikros and T. Evangelidis by homology modeling, using Mhp1 structure (PDB ID: 2JLO) as a template. The three Fur proteins share 22.1-24.2% identity in 421 amino acids overlap, out of the 501 amino acids of the shorter Mhp1, which is adequate for sustaining solid homology models. It should be noted that Fur-like proteins similarity with Mhp1 is significantly higher than those of Fcy-like proteins and Mhp1. The Fur structural models built are shown in **Figure 58**.

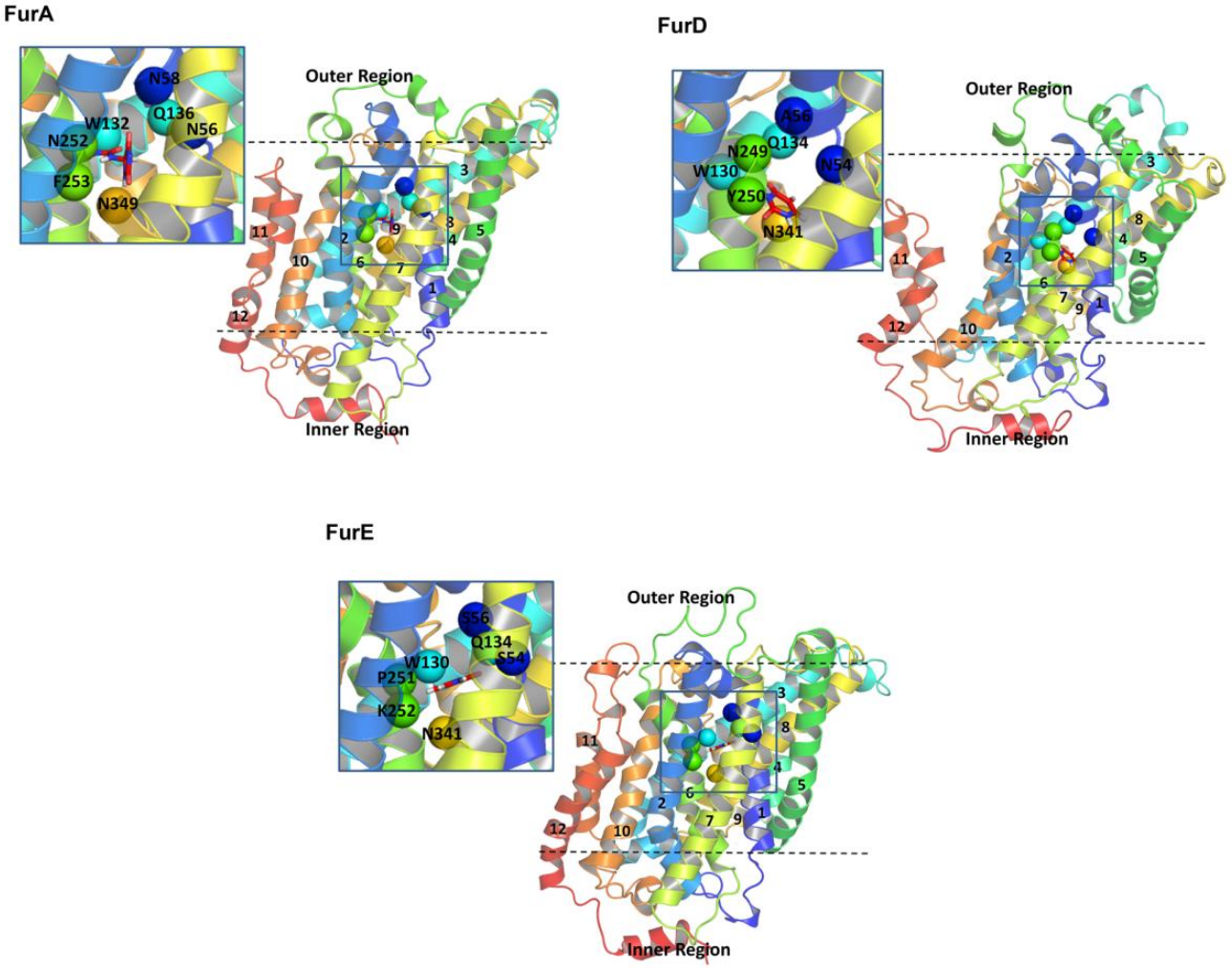


Figure 58. 3D structural models of Fur proteins. Upper left panel: topology of FurA occluded with allantoin and orientation with respect to the membrane. Upper right panel: topology of FurD occluded

[4] Details concerning the methodology used in the computational analysis can be found in Krypotou et al. (2015a), see Appendix.

with uracil and orientation with respect to the membrane. Lower panel: topology of FurE occluded with uracil and orientation with respect to the membrane. Models were built on the basis of alignments with Mhp1. The planes of the membrane are depicted with dashed lines. The substrate binding site including residues of TMS1, TMS3 TMS6 and TMS8 is located in the space between the hash and the bundle domains. Major residues of the substrate binding site are designated with spheres. In all cases, the putative substrate binding cavity is highlighted in zoom-in panels. Homology modelling was performed by T. Evangelidis.

The structure predicted is similar to the one predicted for FcyB. In particular, the topology of the Fur transporters consists of two distinct domains, a compact core (TMS1-10) and a C-terminal domain (TMS11-12). The core is subdivided in two inverted repeats made of TMS1-5 and TMS6-10, arranged in a two-fold pseudosymmetry axis, running in parallel with the membrane plane. The two repeats are completely intertwined and are formed by TMSs 1, 2, 6 and 7 (also referred as bundle motif) and TMSs 3, 4, 8 and 9 (also referred as hash motif) linked with helices TMS5 and 10. The substrate binding site is predicted to be located in the space between the two subdomains.

The Fur structural models were used for induced fit substrate docking (IFD) assays and molecular dynamics (performed by T. Evangelidis) in order to identify the putative binding site and validate the models constructed. The substrates used for the IFD were allantoin for FurA, uracil and uric acid for FurD, uracil, allantoin and uric acid for FurE. The most dominant positions occurred during the calculations are shown in **Figure 59A**. FurD interacted with uracil (K_m 0.5 μM) through two hydrogen bonds with Gln134 and Asn341 side chains. The high affinity binding of uracil could be due to additional interactions that could not be identified through IFD calculations. Uric acid binds with less affinity (K_i 100 μM) and was predicted to interact with FurD through H-bonding with Asn249 (bidentate), Asn341 side chains and Ala56 backbone. Uric acid was also stabilized in the binding pocket through π-π stacking interactions with Trp130. FurA is specific for allantoin, but its affinity cannot be measured because of unavailability of commercially sold radiolabeled substrate. Through IFD, allantoin was predicted to interact with FurA through H-bonding with Asn58, Gln136, Asn252 and Asn349 side chains. FurE was identified in this study as a high affinity transporter for uric acid (K_i 20 μM) and low affinity transporter for allantoin and uracil (K_m ~1000 μM). The IFD calculations predicted that FurE interacted with uric acid through H-bonding with Gln134, Lys252 and Asn341 side chains and could also interact occasionally with Trp130 through π-π stacking. Uracil was predicted to interact through H-bonding with Gln134 and Asn341 side chains; while allantoin interacted through H-bonding with Ser54, Ser56 and Asn341 side chains.

In the case of FurD, an extended molecular dynamics (MD) calculation was also performed (by T. Evangelidis) in the constant presence of uracil or uric acid to identify more interactions and if possible predict its trajectory in the binding pocket; the results obtained are shown in **Figure 59B**. Uracil was again predicted to interact through H-bonding with Gln134 and Asn341 side chains, but a H-bond interaction with Leu253 backbone was also added. The MD calculation also predicted

interactions through water bridges, specifically, mediated by Asp51 and Asn256 side chains (infrequent). The side chain of Thr252 was also detected to be stabilized in a position to interact with uracil through a constant water bridge connecting its side chain with Gly248 backbone. Additional stability to the protein-substrate binding might also be provided by a less frequent water bridge, between Asp51 and Asn256. Uric acid was predicted to interact through H-bonding with Asn249 and Asn341 side chains and Asn249 and Ala56 backbone (interactions also identified by IFD). Frequent water bridges also contributed to substrate binding connecting it with Thr252 side chain and backbone, Gln134 side chain and less frequently with Asn137 side chain. In summary, the MD calculations confirmed and extended the interactions predicted by IFD calculations for FurD-substrate interactions.

The residues identified herein for the three Fur transporters are located in TMSs 1, 3, 6 and 8 similarly to the ones identified previously for FcyB. Two residues, Gln in TMS3 and Asn in TMS8 (conserved in all three transporters), contribute through H-bonding for all substrates binding. The Trp identified in TMS3 (important for FcyB) also interacts through π-π stacking in most cases.

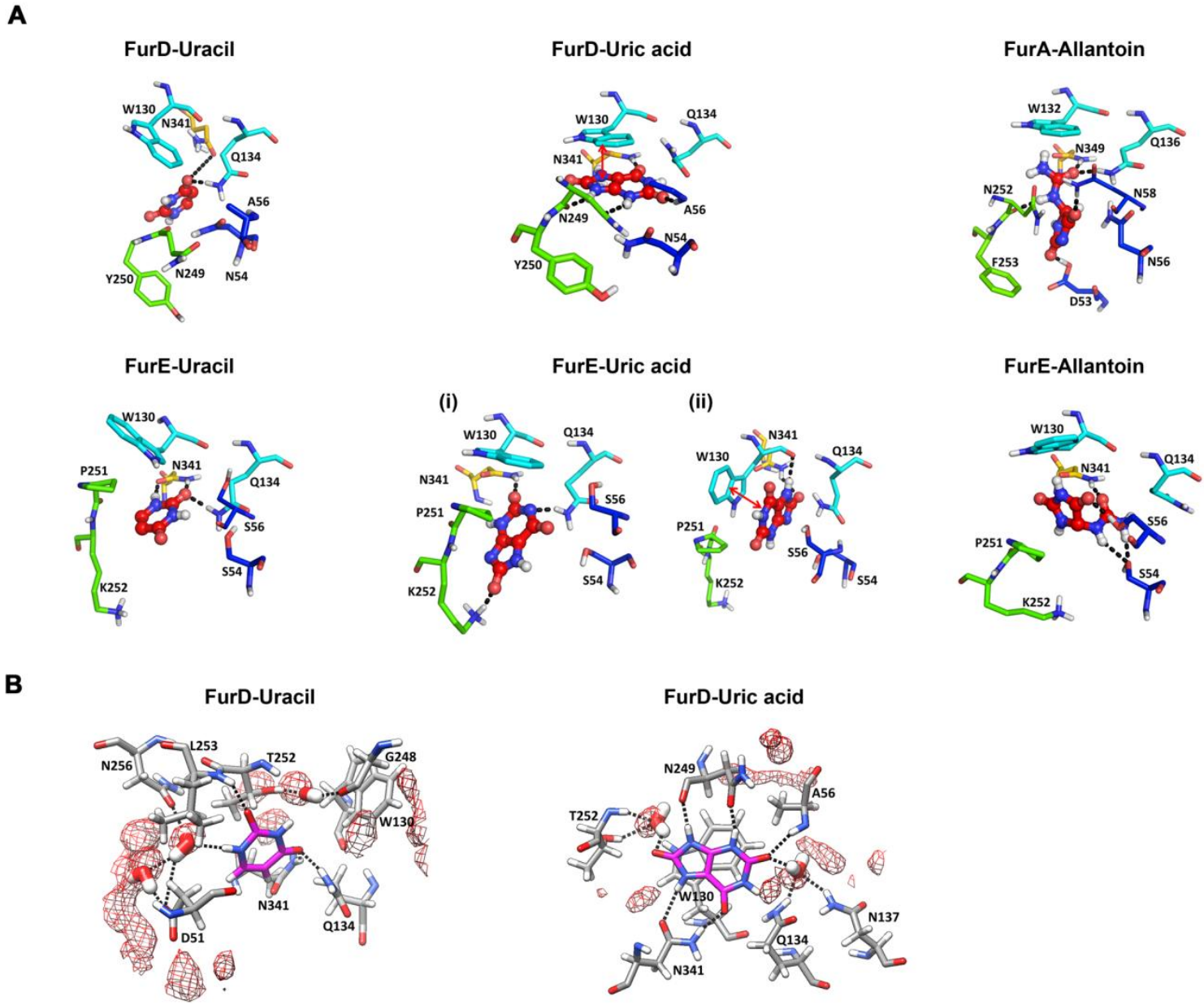


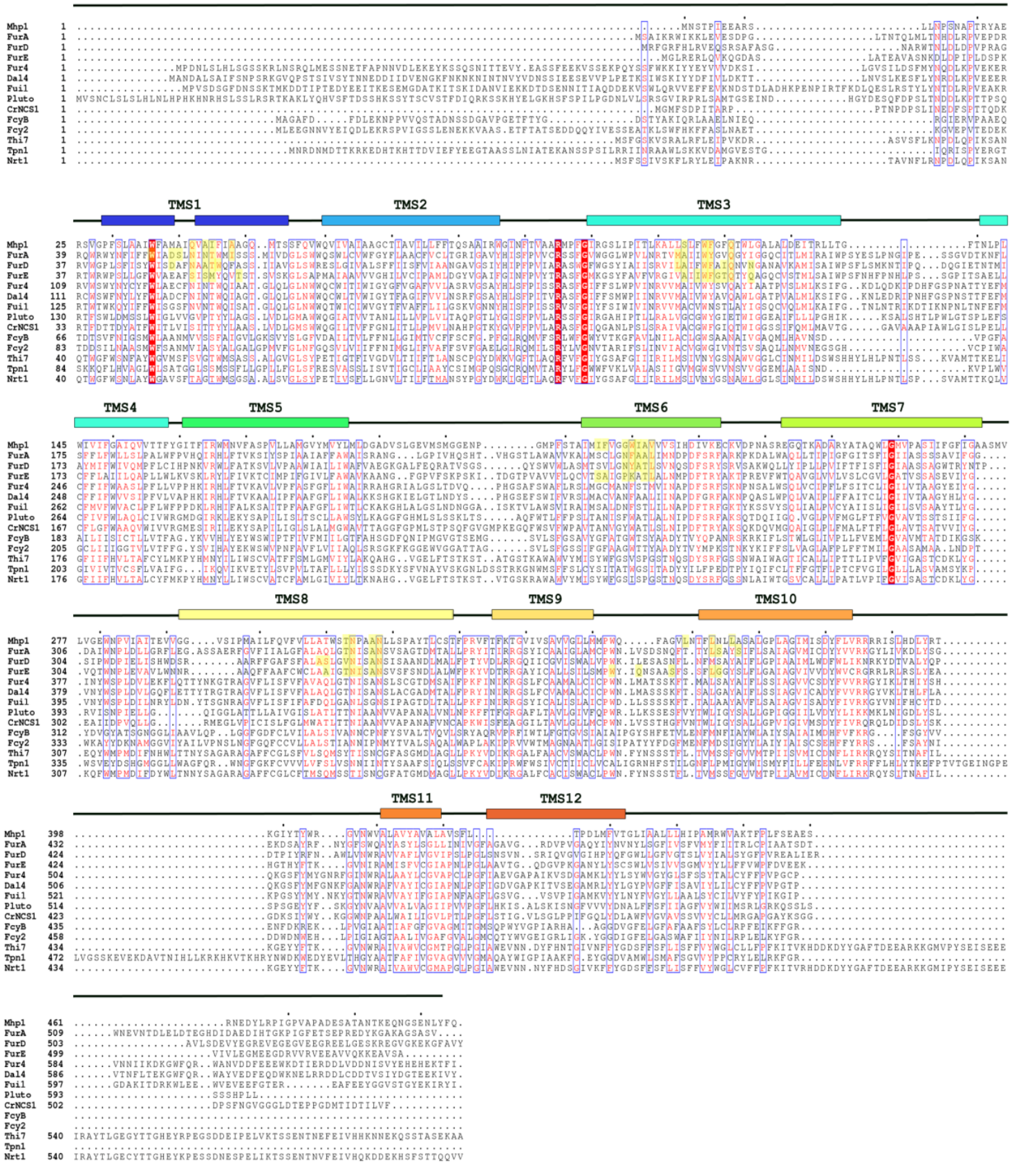
Figure 59. Modelled substrate interactions with the binding sites of Fur proteins. (A) Predominant Docking Poses from IFD calculations. Upper panels: FurD with uracil (left) and uric acid (middle), FurA with R-allantoin (right). Lower panels: FurE with uracil (left), uric acid (i) and (ii) (middle) and R-allantoin (right). Similar interactions were observed for S-allantoin (results not shown). Hydrogen bonds are depicted in black dashes, and π-π-stacking is highlighted with two-headed red arrows. **(B)** Representative snapshots from MD trajectory of FurD with uracil (left) and uric acid (right). Water oxygen density is depicted in mesh form and the water molecules that formed stable water bridges are shown in black dashes. IFD and molecular dynamics were performed by T. Evangelidis.

3.4.8 Mutational analysis of FurD and FurE

The computational analysis identified a few residues of FurD and FurE that are involved in substrate binding. Those residues are for FurD: Ala56 (TMS1), Trp130 and Gln134 (TMS3), Asn249 and Tyr250 (TMS6), and Asn341 (TMS8); for FurE were: Ser54 and Ser56 (TMS1), Trp130 and Gln134 (TMS3), Pro251 and Lys252 (TMS6) and Asn341 (TMS8). In a multiple protein sequence alignment with representatives of NCS1 family, **Figure 60**, we located the residues predicted by the computational analysis (highlighted with yellow). Trp130 and Asn341 (FurD) were absolutely conserved in all NCS1 proteins, while residues Asn249 and Tyr250 (FurD) varied. Gln134 is conserved in the Fur family, but is replaced by Asn in Fcy family and by Glu in plant NCS1 protein.

In order to determine the significance of residues in transport function or/and substrate specificity, identified by the IFD and MD analysis, and those in crucial positions that vary among NCS1 members, we selected FurD to perform an extended mutational analysis. The choice of FurD was due to the fact that FurE is a low capacity transporter and in addition, there is not available radiolabeled substrate for FurA. The amino acid substitutions carried out in FurD were: N54A, T57A, W130A, Q134E, Q134A, N249A, Y250A, Y250F, Y250W, Y250K and N341A. The rationale of the mutations were to replace each amino acid with one isofunctional or isostructural residue or one present in homologue proteins.

We also performed a few mutations in FurE and in particular changed residues that were specifically predicted to interact with the substrates, but were significantly different from the amino acids usually found at the same positions in other Fur transporters. The FurE mutations made were: S54N/S56N/M57T, S54N/S56A/M57T, P251N, K252F and P251N/K252F. The rationale behind the mutational changes were to introduce in FurE the cognate residues found in FurA or FurD under the *gpdA* promoter. The plasmids thus constructed were introduced in the Δ7 strain. Selected single copy transformants were functionally analyzed. Construction and characterization of the FurD mutant strains was performed in collaboration with J. Bobonis.


TMS1
TMS2
TMS3
TMS4
TMS5
TMS6
TMS7
TMS8
TMS9
TMS10
TMS11
TMS12

3.4.9 Functional analysis of the FurD and FurE mutants

Figure 61 summarizes the functional analysis performed for the FurD mutants by growth testing. Functionality of FurD is reflected as growth on uric acid and xanthine at 37^{o}C and sensitivity to FU, FC and FUd at 37^{o}C and 25^{o}C. From their phenotypes on the substrates tested the mutants could be categorized in the following groups: (*i*) wt-like: corresponds to similar to wt growth on uric acid and xanthine and sensitivity to FU, FUd and FC (N54A and T57A); (*ii*) total loss-of-function: corresponds to total absence of growth on uric acid and xanthine and resistance to FU, FUd and FC (W130A and Y250K); (*iii*) partial loss-of-function: corresponds to lack of growth on uric acid or/and xanthine, and/or resistance to one or more toxic analogues (Q134A, Q134E, Y250F, Y250W and N341A); (*iv*) broadened specificity: corresponds to growth on allantoin and retention of the ability to recognize some or all the physiological substrates (N249A and Y250A). Interestingly, two mutants proved thermosensitive, namely Y250F grew on xanthine and N249A grew on all substrates (uric acid, xanthine and allantoin) only at 25^{o}C.

The growth test analysis performed for the FurE mutants is shown in **Figure 62**. Functional (over-expressed) FurE confers growth on uric acid, allantoin, sensitivity to FU and moderate sensitivity to FUd and FC. Mutations S54N/S56N/M57T and S54N/S56A/M57T led to a loss-of-function phenotype, namely absence of growth on uric acid and allantoin, and resistance to all toxic analogues. Mutation K252F led to a higher capacity or/and affinity transporter with broader specificity, namely better growth on uric acid at 25^{o}C, growth on xanthine and allantoin and increased sensitivity to FU, FUd and FC. Mutation P251N led to partial lack of growth on uric acid and allantoin, while exhibited increased capacity for FUd and FC. Finally, the double mutation P251N/K252F resulted in absence of growth on uric acid, growth on allantoin and increased sensitivity to FUd and FC.

Figure 60. Sequence alignment of selected NCS1 members with distinct specificities. Putative TMSs of FurD are denoted in colored rectangles. The sequences shown are: Mhp1 (D6R8X8; *M. liquefaciens*), FurA (Q5BFM0; *A. nidulans*), FurD (A6N844; *A. nidulans*), FurE (Q5ATG4; *A. nidulans*), Fur4 (P05316; *S. cerevisiae*), Dal4 (Q04895; *S. cerevisiae*), Fui1 (Q04895; *S. cerevisiae*), PLUTO (Q9LZD0; *A. thaliana*), CrNCS1 (A8J166; *C. reinhardtii*), FcyB (C8V329; *A. nidulans*), Fcy2 (P17064; *S. cerevisiae*), Thi7 (Q05998; *S. cerevisiae*), Tpn1 (P53099; *S. cerevisiae*) and Nrt1 (Q08485; *S. cerevisiae*). Amino acids of the binding site of Mhp1, FurA, FurD and FurE are highlighted with yellow color.

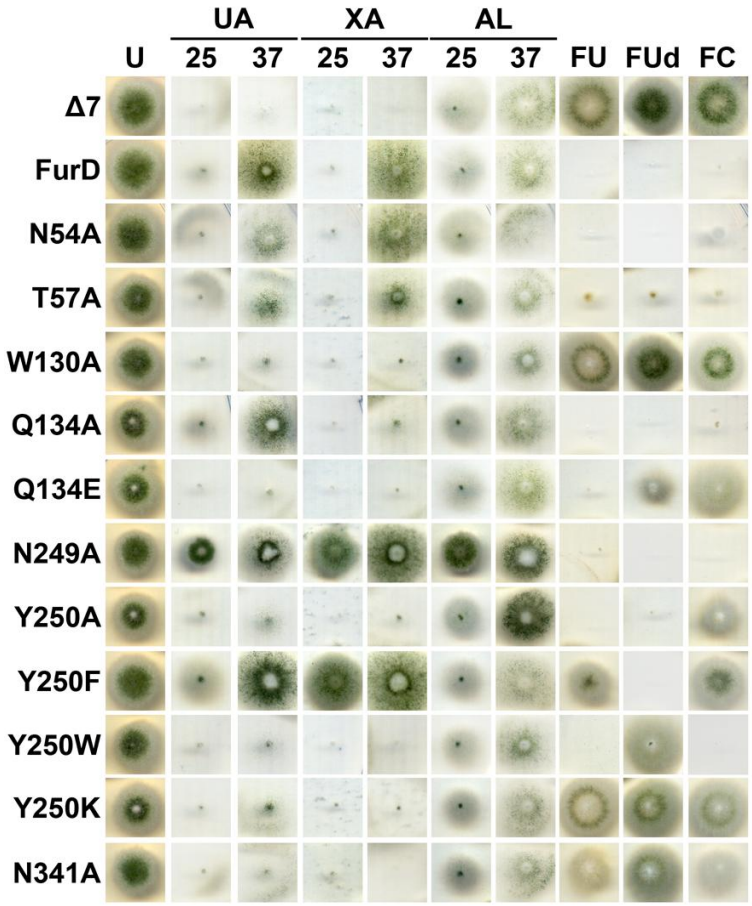


Figure 61. Functional analysis of FurD mutants. Growth tests of FurD mutations single copy transformants. Nitrogen sources or toxic analogues used are as in **Figure 52**. Growth tests shown were at 25°C, or at 37°C when indicated. Toxic analogues were used in media containing urea as a nitrogen source. Isogenic positive (FurD) and negative (Δ7) controls were used.

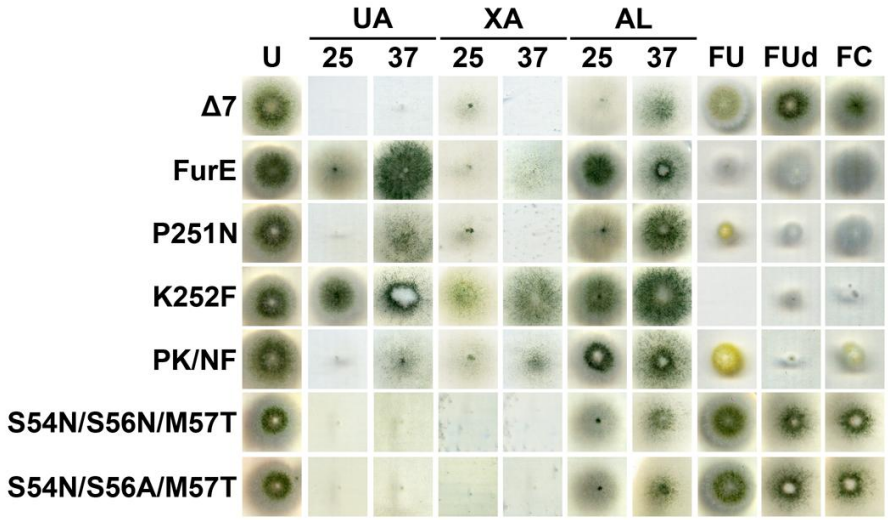


Figure 62. Functional analysis of FurE mutants. Growth tests of FurE mutations single copy transformants. Nitrogen sources or toxic analogues used are as in **Figure 52**. Growth tests shown were at 25°C, or at 37°C when indicated. Toxic analogues were used in media containing urea as a nitrogen source. Isogenic positive (FurE) and negative (Δ7) controls were used.

3.4.10 Kinetic analysis of FurD mutants

Radiolabeled uracil uptake was performed for all FurD mutants to determine their initial uptake rates (*V*) and their binding and inhibition values ($K_{m/i}$) for their substrates. As shown in **Table 10**, mutations W130A, Y250K and N341A led to absence of uracil transport, in agreement with resistance to FU. Mutants Q134A, N249A, Y250F and Y250W retained more or less (75-90%) their uracil uptake rate compared with the wt. Mutants N54A, T57A and Q134E had significantly reduced uptake rates that were 5-15% compared with the wt, while Y250A-mediated uptake rate was 25%. Briefly, all Ala substitutions, except Q134A and N249A, led to significant reduction of the initial uptake rate of uracil (0-25% *V* compared to the wt).

Table 10. Initial uptake rates (*V*) and $K_{m/i}$ values (μM) were determined for FurD mutants using [^{3}H]-uracil uptake competition. Results are averages of at least three independent experiments in triplicate for each concentration point. Standard deviation was < 20%. n.m., non measurable. UR: uracil, UA: uric acid, XA: xanthine, AL: allantoin, HY: hydantoin, TH: thymine. Single copy plasmid integration transformants were used.

		K_m (μM)	K_i (μM)				
	V (%)	UR	UA	XA	AL	HY	TH
FurD	100	0.5	99	100	>1000	40	3
N54A	5 ± 2	40	500	500	>1000	800	-
T57A	15 ± 5	0.5	18	110	>1000	-	-
W130A	0	n.m	n.m	n.m	n.m	n.m	n.m
Q134A	90 ± 5	2	66	140	>1000	100	2
Q134E	8 ± 3	8	216	300	>1000	500	40
N249A	75 ± 5	2.4	510	190	>1000	250	5
Y250A	25 ± 2	20	>1000	>1000	>1000	>1000	130
Y250F	90 ± 3	4.5	194	300	>1000	67	6
Y250W	80 ± 6	3	10	200	>1000	-	1
Y250K	0	n.m	n.m	n.m	n.m	n.m	n.m
N341A	0	n.m	n.m	n.m	n.m	n.m	n.m

$K_{m/i}$ values were determined for the FurD mutants with measurable uptake capacity. FurD binds uracil with high affinity (0.5 μM), uric acid, xanthine and hydantoin with medium affinity (40-100 μM), and thymine also with high affinity (3 μM). Mutations N54A, Q134E, N249A and Y250F led to lower affinity for all substrates with changes of variable-fold. In more detail, Ala substitution of Asn54 led to 80-fold lower affinity for uracil, 5-fold for uric acid and xanthine, and 20-fold for hydantoin. Mutation Q134E led to 16-fold lower affinity for uracil, ~2-3-fold for uric acid and xanthine, ~13-fold for hydantoin and thymine. Mutation N249A led to ~5-fold lower affinity for uracil and uric acid, ~2-fold for xanthine, ~6-fold for hydantoin, while it had almost no effect on thymine binding. Substitution of Tyr250 with a Phe

led to lower affinity for uracil (9-fold) and for uric acid or xanthine (~2-3-fold), but had almost no effect on hydantoin or thymine binding. The rest of the substitutions of Tyr250 had variable effects on the binding affinities. In particular, Y250A resulted in ~40-fold lower affinity for uracil and thymine, but no binding (>1 mM) for the rest of substrates. On the contrary, mutation Y250W led to 10-fold increase in uric acid affinity, but 6-fold and 2-fold decrease in uracil and xanthine binding, respectively, and no effect on thymine binding.

Ala substitutions of Thr57 and Gln134 differentially affected each substrate. For T57A, xanthine binding was wt-like, but uric acid binding affinity was ~5-fold increased. For Q134A, the affinity for uracil was 4-fold reduced, for uric acid 1.5-fold increased, and for xanthine and hydantoin 1.5-fold and ~2.5-fold increased, respectively. Finally, even though N249A and Y250A mutants grew on allantoin, we did not detect any significant allantoin binding (>1 mM) similar to wt.

In summary, the most severe effect on substrate transport and binding was due to mutations in Asn54 (TMS1), Trp130 (TMS3), Tyr250 (TMS6) and Asn341 (TMS8), indicating the involvement of these residues in direct interactions with substrates, as also predicted by the substrate docking analysis. A variable effect on binding of different substrates, probably related with substrate specificity, was obtained by mutations in Asn249 and Tyr250.

3.4.11 Kinetic analysis of FurE mutants

Similarly to FurD mutants, we performed kinetic analysis for all FurE mutants to determine their initial uptake rates (V) and their $K_{m/i}$ values (**Table 11**). As shown herein, FurE is a high affinity transporter for uric acid (20 μM) and shows low affinity for uracil and allantoin (~1 mM). [^{3}H]-uracil uptake was used to kinetically characterize the FurE mutants, since radiolabeled uric acid was not available (due to high affinity for uric acid, more accurate results could have been obtained with that substrate). Mutants in TMS1 (S54N/S56N/M57T and S54N/S56A/M57T) lost transport function since no measurable uptake was detected. Differently from the above, mutants P251N, K252F and P251N/K252F led to a surprising increase in uracil initial uptake rate (~13-20-fold) compared with the wt. This increase could be associated with an increase in uracil K_m, ~37-fold for P251N, 125-fold for K252F and ~333-fold for P251N/K252F. These mutations also increased the affinity for other substrates.

In addition, mutation P251N resulted in 4-fold lower K_i for uric acid, but 55-fold, higher K_i for allantoin, ~10-fold higher K_i for hydantoin and thymine. Mutation K252F, similarly to P251N, led to 4-fold lower K_i for uric acid, ~18-fold higher K_i for allantoin, and ~10-fold higher K_i for hydantoin and thymine, while it could also bind xanthine with a K_i of ~1000 μM. Finally, the double mutation P251N/K252F led to 12.5-fold lower K_i for uric acid, and higher K_i for the rest of substrates: 3-fold for allantoin, 25-fold for hydantoin, 2000-fold for thymine and could bind xanthine with a K_i of 800 μM.

Apparently, residues Pro251 and Lys252 affect substrate binding, and their substitution with residues present in FurA (Asn and Phe respectively) resulted in high affinity binding of uracil and other substrates, as well as increased initial uptake rate.

Table 11. Initial uptake rates (*V*) and $K_{m/i}$ values (μM) were determined for FurE mutants using [^{3}H]-uracil uptake competition. Results are averages of at least three independent experiments in triplicate for each concentration point. Standard deviation was < 20%. All symbols as indicated in **Table 10**. Single copy plasmid integration transformants were used.

		K_m (μM)	K_i (μM)				
	V (%)	UR	UA	XA	AL	HY	TH
FurE	100 ± 2	~1000	20	>1000	~1000	>1000	>1000
P251N	2000 ± 15	27	80	>1000	18	97	140
K252F	2000 ± 25	8	80	~1000	56	130	117
PK/NF	1280 ± 5	3	250	800	330	40	5
SN/SN/MT	0	n.m	n.m	n.m	n.m	n.m	n.m
SN/SA/MT	0	n.m	n.m	n.m	n.m	n.m	n.m

3.4.12 A putative selective gate in FurD

TMS10 in Mhp1 has been described as a moving flexible element (Weyand et al., 2008) that could be related to the entry of the substrate from the outer environment to the binding pocket. In addition, as described in the Introduction, a single mutation in Leu353 of Mhp1 resulted in transport of a bulky analogue, that could bind but could not be transported by the wt protein (Simmons et al., 2014). We located the respective area in FurD by protein sequence multiple alignment (**Figure 63**) and identified residues highly or partially conserved with Mhp1.

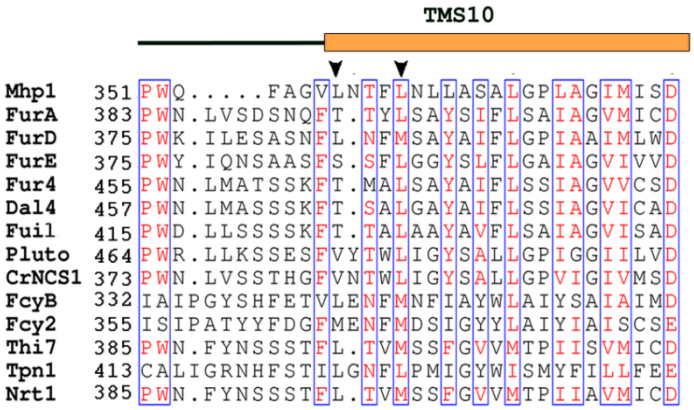


Figure 63. Sequence alignment of TMS10 of selected NCS1 transporters. The sequences shown are: Mhp1 (D6R8X8; *M. liquefaciens*), FurA (Q5BFM0; *A. nidulans*), FurD (A6N844; *A. nidulans*), FurE (Q5ATG4; *A. nidulans*), Fur4 (P05316; *S. cerevisiae*), Dal4 (Q04895; *S. cerevisiae*), Fui1 (Q04895; *S. cerevisiae*), PLUTO (Q9LZD0; *A. thaliana*), CrNCS1 (A8J166; *Chlamydomonas reinhardtii*), FcyB (C8V329; *A. nidulans*), Fcy2 (P17064; *S. cerevisiae*), Thi7 (Q05998; *S. cerevisiae*), Tpn1 (P53099; *S.*

cerevisiae) and Nrt1 (Q08485; *S. cerevisiae*). Arrow heads indicate the positions Leu353 in Mhp1 could align with the rest sequences.

In particular, Leu353 of Mhp1 could be aligned with two residues in FurD, Leu386 or Met389. The first residue, Leu386, varied to Thr/Leu/Ser among Fur-like proteins and Val/Leu/Met in Fcy-like and plant homologues. On the other hand, the residue at the position corresponding to 389 of FurD was a Met only in FurD and Fcy-like proteins, and a Leu in the other homologues. In order to identify a potential role of these residues, we constructed the following mutations in FurD: L386A, L386S, M389A, M389L, L386-N387/L386-E387*-N388 (Glu387 is an inserted residue, corresponding to the extra amino acid found in Fcy-like transporters). The mutant alleles were introduced to the proper recipient strain as previously described.

Selected transformants were analysed by growth tests (**Figure 64**) similarly to the FurD mutants. Mutations L386A, L386S, M389L and L386-N387/L386-E387*-N388 led to loss-of-function phenotypes, but mutation M389A led to a functional transporter with broadened specificity. More specifically, a single copy transformant expressing FurD-M389A could grow on uric acid and xanthine at 25^{o}C and 37^{o}C, showed moderate growth on allantoin and hypoxanthine (2 mM) and was sensitive to FU, FUd, FC and 8-azaguanine (AZ). Multi-copy expression of FurD-M389A led to increased growth on allantoin and hypoxanthine.

These results indicate a probable role of Met389 in substrate specificity, despite the modelling evidence showing that it is not part of the *bona fidae* substrate binding site. A possible explanation could be that M389 is a dynamic element of an outward-facing, channel-like, filter or "gate", which controls substrate access to the major binding site, and thus affects substrate selectivity, similar to what was observed for UapA (Kosti et al., 2010).

In order to expand the observations obtained by using M389A, we combined this mutation with those that were shown to affect substrate specificity of the substrate-binding site (N249A, Y250A and Y250F). The mutants were constructed as previously described and transformants were analyzed by growth tests (**Figure 64**). Plasmid and strain construction of the double mutants was performed in collaboration with G. Sioupouli.

The double mutant N249A/M389A combined properties from both mutations, as it grew on uric acid, xanthine, allantoin, hypoxanthine (1-2 mM) and was sensitive to all toxic analogues. It thus seems that at least the double mutant N249A/M389A has a broader specificity profile, characteristic of a promiscuous Fur transporter. On the other hand, mutant Y250A/M389A grew on allantoin and hypoxanthine (2 mM) and was sensitive to all toxic analogues, whereas mutant Y250F/M389A grew on xanthine and uric acid, similarly to the single mutant Y250F, grew on allantoin and hypoxanthine (2 mM) only at 37^{o}C, and was sensitive to all toxic analogues.

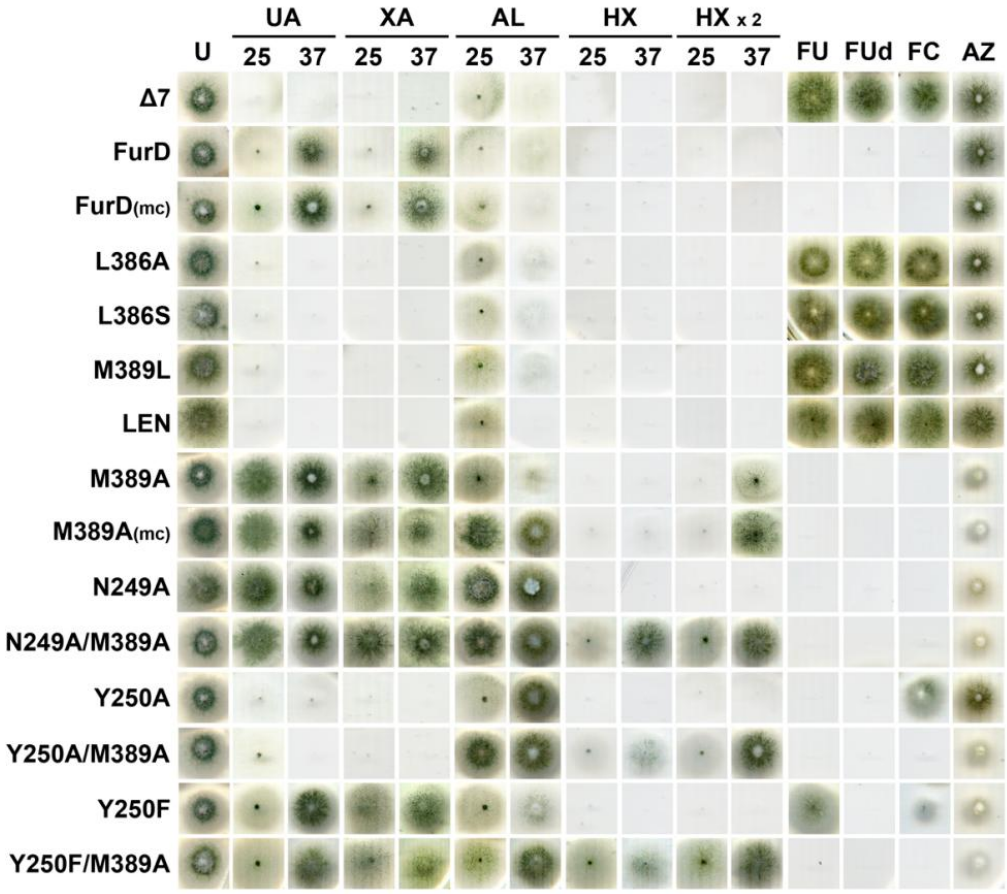


Figure 64. Functional analysis of FurD TMS10 mutants. Growth tests of FurD TMS10 mutations transformants. Nitrogen sources or toxic analogues used are as in **Figure 52**, hypoxanthine (HX) and 8-azaguanine (AZ: 0.1 mg/mL). Growth tests shown were at 25°C, or at 37°C when indicated. Toxic analogues were used in media containing urea as a nitrogen source. HX corresponds to a concentration of 1 mM and HXx2 to 2 mM. Isogenic positive (FurD) and negative (Δ7) controls were used.

The single and double mutants were also kinetically analyzed by [^{3}H]-uracil uptake; initial uptake rates (*V*) and $K_{m/i}$ values were determined (**Table 12**) for mutants with measurable uptake. The single mutant M389A and the double mutants N249A/M389A and Y250F/M389A retained their initial uptake rate at 70-90% compared with the wt, while N249A/M389A expressed reduced *V* at 20% compared with the wt. Mutation M389A led to 16-fold lower affinity for uracil, but 2-fold higher affinity for uric acid and xanthine, and to a measurable affinity for allantoin (700 μM). Addition of the second mutation N249A led to wt-like uracil affinity, 5-fold lower affinity for uric acid, 3-fold higher affinity for xanthine and medium affinity for allantoin (100 μM). On the other hand, double mutation Y250A/M389A led to 13-fold lower affinity for uracil, >10-fold lower affinity for xanthine and uric acid and medium affinity for allantoin (330 μM). Finally, Y250F/M389A behaved similarly to M389A, namely led to wt-like uracil uptake, and ~1.2 fold lower affinity for uric acid and xanthine, and to low affinity for allantoin (700 μM). Interestingly, only for the double mutant Y250A/M389A we could detect measurable uracil uptake inhibition in the presence of excess hypoxanthine (35%).

Table 12. Initial uptake rates (*V*) and $K_{m/i}$ values (µM) were determined for FurD TMS10 mutants using [^{3}H]-uracil uptake competition. Competition of [^{3}H]-uracil uptake by excess hypoxanthine (HX) is also shown (right column). Results are averages of at least three independent experiments in triplicate for each concentration point. Standard deviation was < 20%. UR: uracil, UA: uric acid, XA: xanthine, AL: allantoin.

		K_m (µM)	K_i (µM)			% Uracil uptake
	V (%)	UR	UA	XA	AL	HX (1mM)
FurD	100	0.5	99	100	>1000	100
M389A	90 ± 5	8	40	50	700	88 ± 3
N249A	75 ± 5	2.4	510	190	>1000	80 ± 5
N249A/M389A	70 ± 6	0.5	500	34	100	88 ± 6
Y250A	25 ± 2	20	>1000	>1000	>1000	77 ± 6
Y250A/M389A	20 ± 5	6.5	~1000	>1000	330	35 ± 8
Y250F	90 ± 3	4.5	194	300	>1000	77 ± 5
Y250F/M389A	85 ± 4	0.5	85	77	700	60 ± 6

In summary, Met389 and probably other elements in TMS10 could be implicated in substrate selectivity directly or indirectly. Combination of M389A in TMS10, which is a mutation that apparently allows new substrates to be transported, with specificity mutations in the major binding site, led to a transporter with broader specificity and increased affinity for most substrates, both physiological and non-physiological.

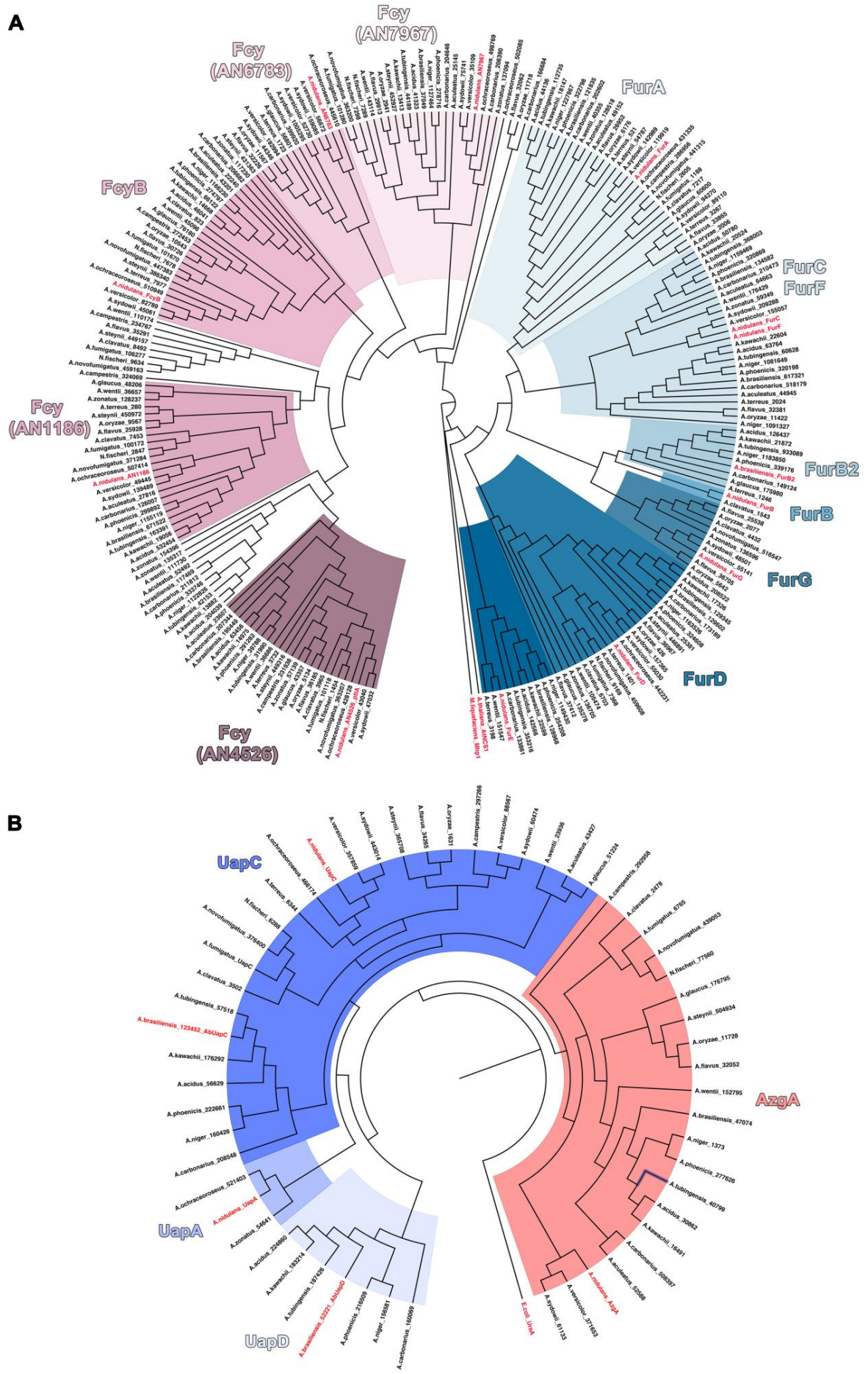
4 Discussion

4.1 An evolutionary perspective of fungal nucleobase transporters

The purine/pyrimidine transporters identified by phylogenetic and functional analyses in *A. nidulans* belong to three distinct families: (*i*) NCS1 family with 12 members (5 Fcy- and 7 Fur-like proteins), (*ii*) NCS2/NAT family with 2 members (UapA/C-like proteins), (*iii*) AzgA-like family with one member (AzgA). To highlight the distinct evolutionary profile of each transporter family, a phylogenetic analysis of the homologous proteins found in Aspergilli is shown in **Figure 65**.

The NCS1 family is predicted to have expanded by multiple gene duplications in fungi and resulted in multimembered Fcy and Fur subfamilies (Krypotou et al., 2015a). FurA and FurD fungal orthologues of the Fur subfamily are ubiquitously found in almost all Aspergilli, whereas the other Furs are absent from many fungi (**Figure 65A**). Interestingly, FurA and FurD, correspond to the major allantoin and uracil transporters respectively in *A. nidulans* (Amillis et al., 2007; Hamari et al., 2009), whereas the function of FurC, FurF and FurG as uracil transporters and FurE as an allantoin-uracil-uric acid transporter was revealed, only upon over-expression, suggesting a low capacity and/or low affinity for the substrates. Similarly to the Fur subfamily, Fcy-like proteins are variably conserved among the Aspergilli (**Figure 65A**) and we can also detect homologous proteins that are not present in *A. nidulans* (**Figure 65A**). Of the Fcy subfamily, FcyB was characterized as a high-affinity, medium capacity, hypoxanthine-adenine-guanine-cytosine transporter (Vlanti and Diallinas, 2008). The function of a second member of the family was recently revealed upon over-expression (G. Sioupouli and G. Diallinas, unpublished results).

In contrast to the NCS1 family, AzgA- and NAT/NCS2-like proteins constitute single-membered (AzgA) or much less expanded (NAT/NCS2) families (**Figure 65B**). A single AzgA-like protein is highly conserved in all Aspergilli except *A. zonatus*. All Aspergilli include a UapC-like protein and three species also include UapA-like members (*A. nidulans*, *A. zonatus*, *A. ochraceoroseus*). The more distinct UapD subgroup has members in fewer species, particularly the black Aspergilli.


A
Fcy
(AN7967)
Fcy
(AN6783)
FurA
FcyB
FurC
FurF
FurB2
Fcy
(AN1186)
FurB
FurG
FurD
Fcy
(AN4526)
B
UapC
AzgA
UapA
UapD

Thus, the three transporter families seem to illustrate distinct examples of how genetic and evolutionary mechanisms have variably affected the emerging and establishment of novel functions and specificities. An organism can acquire new genes through at least three distinct, potentially overlapping mechanisms: lateral gene transfer from another organism, evolution by modification of a preexisting gene after duplication, or *de novo* generation from noncoding DNA. Both two first scenarios could concern the NCS1 family. Existence of multiple gene copies relaxes the selective pressure and allows random evolutionary processes to take place. Thus, gene duplications are quite abundant in nature and several scenarios are possible for the fate of a duplicated gene, few of which are: inactivation of one copy by accumulation of deleterious mutations (pseudogene), improvement of a preexisting function by accumulation of neutral mutations or emerging of a novel function (for a review see Andersson et al. 2015)

The phylogeny of the transporters studied herein suggests that gene duplications do not occur ubiquitously or randomly (e.g. a single AzgA gene in all fungi *vs.* multiple NCS1 genes). An increasing number of studies currently supports the existence of "*hotspots*" in the chromosomes of organisms, where genetic crossovers are more likely to occur and can potentially lead to gene duplications. These "hotspots" seem to partly depend on local DNA sequence and their frequency varies from one organism to another (Hey, 2004). Another hypothesis that could explain the diverse phylogenetic profile of the transporter families studied herein is that some protein structures might have more potentiality to acquire new function or to conserve old functions in spite of random mutations, while other are very "intollerant" and are much more easily inactivated and then subsequently lost.

The creation of new genes after duplication is proposed to be based on three aspects "innovation-amplification-divergence", where the parental gene has a trace of a function additional to its main function; this trace can be inherited in the duplicated gene and evolved in an environment where this is beneficial (Andersson et al., 2015).

Figure 65. **Maximum likelihood phylogenetic trees of the NCS1 (A), NCS2 and AzgA-like proteins (B) of the Aspergilli. (A)** Fur-like proteins are highlighted in shades of blue and Fcy-like proteins in shades of pink according to identified Fcy proteins in *A. nidulans*, distinct proteins with homologues absent from *A.nidulans* are in white background. **(B)** UapA-like proteins are highlighted in shades of blue and AzgA-like proteins in red. Characterized proteins are highlighted in red. Mhp1 from *M. liquefaciens*, AtNCS1 from *A. thaliana* in (A) and UraA from *E. coli* in (B) are used as outgroups. Sequences were retrieved with BlastP using FcyB, FurA and UapA as probes in JGI database. The trees were build with PhymL after alignment of protein sequences with Clustal Omega and curation with TrimAI (Capella-Gutiérrez et al., 2009; Guindon et al., 2010; Sievers et al., 2011).

In particular, the phylogenetic profile of Fur family (**Figure 66**) suggests that Fur proteins originate from an ancestral protein closer to FurE (Krypotou et al., 2015a). Characteristics of the functional promiscuity of FurE (allantoin, uracil, uric acid) were thus, probably, inherited and evolved in Fur proteins with increased specificity (e.g. allantoin or uracil) and higher transport capacity (e.g. FurA and FurD, respectively) or lower transport capacity (FurC, FurF, FurG). Interestingly, a single Fur protein without an apparent function was also found (FurB), suggesting that some genes/proteins might have acquired deleterious mutations and are prompt to become extinct. A similar case to FurB was also found within the UapD-subgroup of the NCS2 family (discussed separately later).

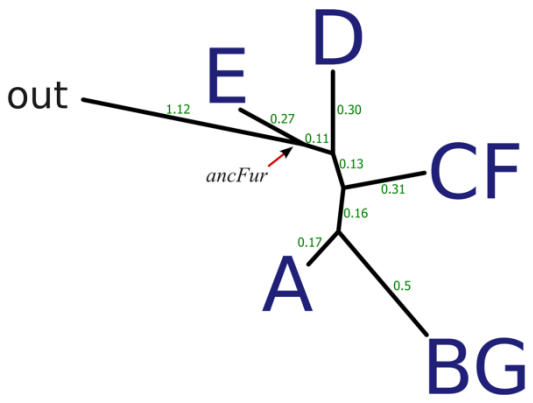


Figure 66. Simplified scheme of a maximum likelihood phylogenetic tree of fungal Fur proteins. The tree shows the phylogeny of the Fur proteins of *A. nidulans* (represented as capital letters). FurE (E) is predicted to be closer to the ancestral Fur protein. The Figure was adapted from (Krypotou et al., 2015a).

It is also interesting that substrate specificities of fungal nucleobase transporters seem to have evolved independently more than once, often in transporters with distinct structural topologies. More specifically, the specificity profile of AzgA (hypoxanthine-adenine-guanine) overlaps that of FcyB (hypoxanthine-adenine-guanine-cytosine), that of UapA/C (uric acid-xanthine) overlaps that of FurD (uracil-uric acid - xanthine), while the *A. nidulans* and *S. cerevisiae* uracil, allantoin or nucleoside transporters belong to distinct Fur sub-groups or even to different major transporter families (see Hamari et al., 2007).

furB and uapD: proto-neogenes or proto-pseudogenes?

Northern analysis (Hamari et al., 2009) and RNA-seq transcriptomic analysis (Sibthorp et al., 2013) have shown that *furB*-mRNA is not detected under any of the conditions tested, with the only exception a low but detectable expression under 72 hours nitrogen starvation (Sibthorp et al., 2013). Expression of FurB from a

constitutive non-native promoter led to low protein levels, associated with cytoplasmic foci and internal structures and never with plasma or other organellar membrane, as expected for a transporter. Thus, is not surprising that no function was ever detected. A close homologue of FurB from the FurB2 subgroup is a functional transporter specific for uracil with plasma membrane localization, which suggests that FurB has accumulated a low number of mutations leading to its low expression and misfolding. The absence of FurB from most Aspergilli might even suggest that this gene is in the process of becoming extinct.

The case of FurB resembles that of AbUapD from *A. brasiliensis*. Both proteins were found localized in internal structures, non-functional and with a patchy phylogenetic profile. Transcription data show that the *uapD* orthologue of *A. niger* is transcriptionally inactive (Delmas et al., 2012; van Munster et al., 2014), similarly to FurB. Moreover, the fact that chimeric molecules, made mostly by UapA and short segments of AbUapD, were also unstable and functionless, suggests that AbUapD might contain intrinsic characteristics that lead to misfolding and high protein turnover. The lack of any function or expression and the incongruous phylogenetic profile suggest that UapDs might constitute a group of genes that accumulated deleterious mutations and might eventually evolve to pseudogenes or new genes. Interestingly, we detected a *bona fidae* UapD pseudogene in *Pyrenophora teres* that includes several stop codons and truncations and lacks a start codon, while its orthologue in *Pyrenophora tritici* shows an open reading frame.

A third, similar case of unstable and functionless transporter has also been reported recently as a homologue of the major acetate transporter, AcpA (Sá-Pessoa et al., 2015).

Thus, it seems that occasionally some genes can be evolutionary conserved despite the fact that they might encode for proteins with apparently no function. This idea contrasts the dogma that genes without a function get extinct or rapidly become pseudogenes. As mentioned above, proteins of novel function can arise after a number of random mutations in a preexisting gene, a process that can be succeeded only after gene duplication in order to assure that at least one functional gene copy will remain and one will be available for evolutionary "experiments". FurB and UapD, and similar genes, could be the intermediates of an evolutionary process that could eventually lead to a pseudogene or a gene with new properties. Thus, these genes could be cases of *proto-pseudogenes* or *proto-neogenes* (Krypotou et al., 2015b).

4.2 The cryptic physiological function of Furs

Overexpression of the Fur proteins revealed the functionality and in particular the specificity profile of the orphan Furs, except FurB; however, their physiological role remains unclear. The only data available related to their physiological function are from expression analysis experiments. In particular, RNA-seq transcriptomic analysis (Sibthorp et al., 2013) revealed that all Fur proteins are significantly expressed under

nitrogen starvation conditions, but only FurA, FurD and FurE are expressed in the presence of NO_3 (Hamari et al., 2009; Sibthorp et al., 2013). Thus, the minor phenotypes revealed herein by growth tests on uracil and 5-fluorouracil of multiply *fur* deleted strains were justified (see Results).

Interestingly, next to FurE lie two genes (AN8417 and AN8418) predicted to encode for proteins with hydrolytic activity. In particular, AN8417 is predicted as a hydantoin racemase. The three genes FurE, AN8417 and AN8418 appear to form a cluster conserved in other Aspergilli, while they are lowly expressed under the same conditions (Sibthorp et al., 2013). Similarly, next to FurF lies a putative alternative allantoinase (AN9327), an orthologue of the PuuE alternative allantoinase from *Pseudomonas fluorescens*; and next to FurC lies a putative second paralogue of the enzyme (AN3354). *furC* and *furF* along with their neighbouring putative enzymes appear to be expressed under the same nitrogen conditions (Sibthorp et al., 2013). It seems that the Fur proteins are conserved as an enzyme-transporter cluster and their physiological role could be related with the role of the neighboring enzymes. Thus, a further analysis that involves the functional characterization and the expression profile of the proteins of each cluster could reveal the physiological role of Furs.

A more detailed analysis concerning the enzymes lying next to Furs can be found in Galanopoulou et al. (2014).

4.3 Structural determinants of the function and specificity of NAT/NCS2 and NCS1 proteins

Based on the crystal structures of bacterial homologues, *A. nidulans* NCS1 proteins are predicted to comprise an Mhp1-like (or LeuT-like) fold with 12 TMSs, whereas NCS2 and AzgA proteins are compatible with a UraA-like fold with 14 TMSs (see also later). Even though, at first sight, the structural folds of Mhp1 and UraA seem quite different (**Figure 67**, see also Introduction), they do share several structural and mechanistic features and are predicted to be of common origin (Västermark and Saier, 2014). In particular, the 5+5 symmetrical structure of the LeuT-fold is predicted to have arisen after a duplication of the 5-TMS subunit; while, the 7+7 symmetrical structure of UraA-fold seems to have arisen after two consecutive duplications: (*i*) duplication of an N-terminal hairpin of a similar to LeuT-fold 5-TMS subunit creating a 7-TMS subunit and (*ii*) duplication of the 7-TMS subunit to create the 7+7 final structure (Västermark and Saier, 2014). Thus, TMSs 1, 3, 6 and 8 that are predicted to comprise the substrate binding site in LeuT-fold proteins, seem to structurally correspond to TMSs 3, 5, 10 and 12 in UraA-fold proteins (Västermark and Saier, 2014). However, the substrate binding site in UraA is located between TMSs 1, 3, 8 and 10. Thus, the two structures have functionally diverged, even though they are considered to have a common overall structural origin.

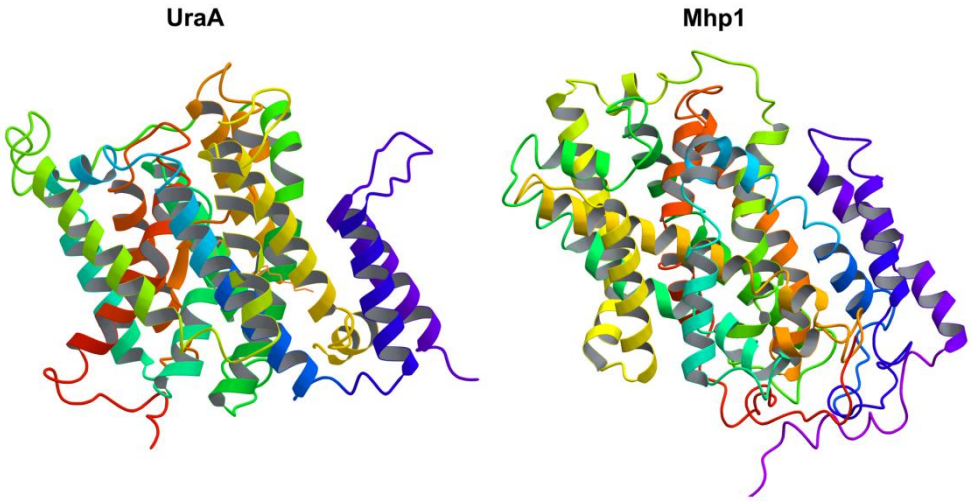


Figure 67. Overview of the crystal structures of UraA and Mhp1. TMSs are highlighted in different colors. (Lu et al., 2011; Shimamura et al., 2008)

The current and previous studies from our laboratory have identified few residues critical for the function and specificity of the NCS1 (FurD, FcyB), NCS2 (UapA, UapC) and AzgA proteins of *A. nidulans*. One class concerns residues located in the TMSs hosting the substrate-binding site, a finding that might be expected. More interestingly, a distinct class of residues affecting mostly and more dramatically the specificity of all *A. nidulans* purine transporters concerns residues located outside the binding site. This finding led to the novel concept that transporters contain channel-like gates controlling substrate selectivity (see also later).

4.3.1 AzgA-like proteins acquire an NCS2-like structure

A solid theoretical model was built for AzgA by homology modelling on the crystal structure of an NCS2/NAT member, UraA, even though AzgA-like and NCS2 proteins constitute distinct transporter families (Krypotou et al., 2014). Computational analysis of the theoretical structural model of AzgA identified few residues to interact through hydrogen bonding with the substrates, namely Asn131 (TMS3), Asp339 (TMS8), Glu394 (TMS10) and the amide group of the backbone of Val393 (TMS10) (Krypotou et al., 2014). In addition to these residues, a hydrogen bond network among Asp339-Glu394-Asp342 is predicted to be involved in H^+ binding and substrate translocation.

Our mutational and functional analysis, presented herein, proposes that Asn131, Asp339, Asp342 and Glu394 are involved directly or indirectly in substrate and/or H^+ binding, since they were found irreplaceable for function. Interestingly, their substitutions differentially affected the binding affinities of AzgA physiological substrates, especially adenine, proposing a different contribution in the binding of each substrate. Among the residues mutated, additional three were found to be necessary for AzgA function, Gly129 (TMS3), Thr390 and Gly402 (TMS10). These three amino acids are predicted to be in close proximity to the substrate-binding site

and could contribute to substrate translocation or formation of a proper binding cavity architecture. Mutations of the rest of amino acids led to more or less significant phenotypes, but because of the non-physiological protein localization we could not reach definite conclusions. Additional AzgA loss-of-function mutants could be easily acquired exploiting the high rate occurring resistance sectors of AzgA-expressing strains grown in the presence of the toxic analogue 8-azaguanine. An AzgA mutant library based on thorough screening could reveal residues critical for AzgA transport function not predicted by the computational analysis.

The functionally critical residues identified herein are in good agreement with those identified by the work of Papakostas et al. (2013) for the bacterial AzgA-homologues (PurP and YjcD). In particular, the *E. coli* PurP and its isofunctional YicO were identified as adenine transporters and YjcD and its isofunctional YgfQ as hypoxanthine and guanine transporters (Papakostas et al., 2013). Mutational and functional analysis revealed an Asp (Asp267 for PurP and Asp271 for YjcD) in TMS8 and a Glu (Glu318 for PurP and Glu322 for YjcD) in TMS10 that were essential for function. The Asp and Glu residues from the bacterial proteins correspond to Asp342 and Glu394 in AzgA (also found critical for function) (**Figure 68**). Other residues found to impair more or less the function of YjcD/PurP and AzgA were: Thr35/Thr38, Thr275/Thr271 in YjcD/PurP and Ser323 in YjcD that correspond to Thr49, Thr390 and Ser395 in AzgA.

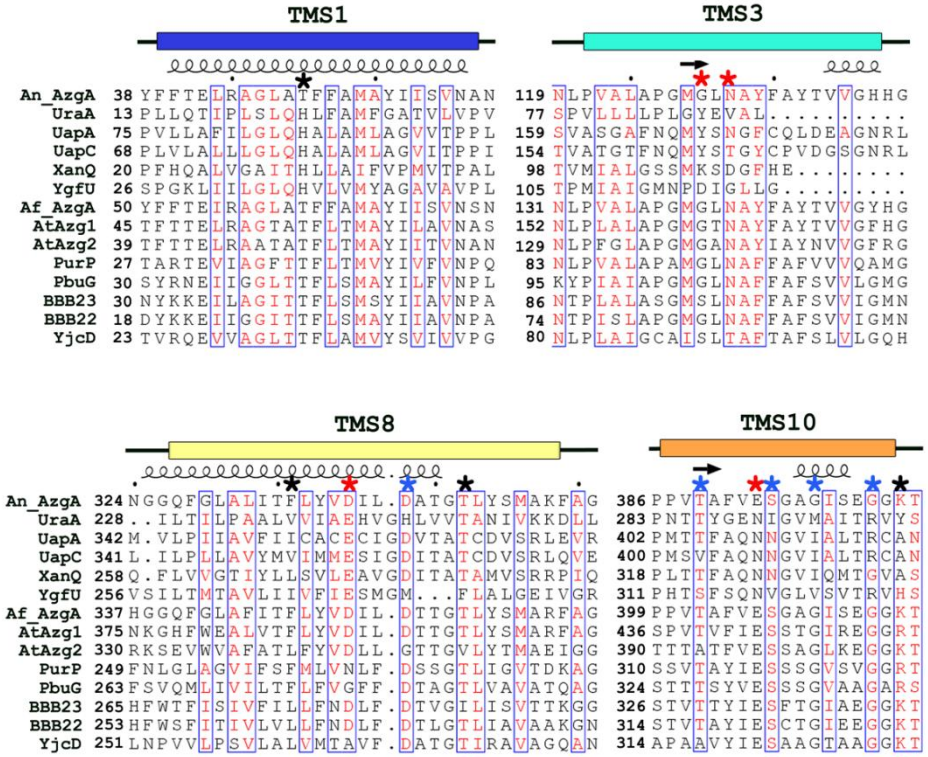


Figure 68. Multiple protein sequence alignment of AzgA predicted TMSs that constitute the binding site, and AzgA-like sequences and characterized NAT proteins. Invariant and highly conserved amino acids are shaded in red and blue-lined boxes, respectively. Amino acids essential for substrate binding and transport are highlighted with red asterisks. Residues critical for function and less

or not important residues are marked with blue and black asterisks, respectively. The alignment is part of the full length sequence alignment shown in **Figure 34**.

Interestingly, the two bacterial homologues have distinct specificity profiles, PurP is specific only for adenine and YjcD only for hypoxanthine and guanine, in contrast with fungal and plant AzgA-like proteins that can take up all three substrates. Based on their protein sequence alignment, several residues vary among PurP or YjcD and the rest AzgAs that could reflect their different specificities. For example, YjcD has Ser, Thr, Ala and Ala instead of Gly (Gly129), Asn (Asn131), Asp (Asp339) and Thr (Thr390) respectively found in the fungal and plant AzgAs (AzgA residues in parenthesis); while, PurP shares more common residues with the other AzgAs with only an Asn residue varying from the conserved Asp (Asp339) (**Figure 68**). The varying residues could be related with the different specificity preferences of PurP and YjcD affecting their substrate specificity.

The well-characterized NAT homologues, UapA from *A. nidulans* and XanQ from *E. coli* have revealed a set of amino acid residues that are absolutely essential for function. The residues identified herein for AzgA and in the work of Papakostas et al., (2013) for PurP and YjcD seem to topologically correspond to several of the residues proven important for UapA and XanQ. For example, the functionally irreplaceable Glu356/Glu272 (TMS8) and Asn409/Asn325 (TMS10) of UapA/XanQ correspond to Asp339 and Glu394 of AzgA (**Figure 68**).

Thus, it appears that the NAT proteins are structurally related to the AzgA-family, as they share similar structural fold and amino acid residues critical for function with respective topology. Based on phylogenetic analyses (Cecchetto et al., 2004; Krypotou et al., 2014), the AzgAs are also predicted phylogenetically more related to NATs than other families. AzgA and NATs common structural fold could be due to events of convergent evolution or the AzgA-family could have diverged early from the NAT family, since they exhibit a completely different specificity profile.

4.3.2 Fcy and Fur transporters differ in substrate specificity elements

Theoretical structural models for FcyB, FurA, FurD and FurE were built based on homology modelling on the crystal structure of the bacterial benzyl-hydantoin transporter Mhp1 (Krypotou et al., 2012, 2015a). The models were used in substrate docking calculations to identify residues participating in substrate binding (Krypotou et al., 2012, 2015a). Our mutational and functional analysis for FcyB and FurD, presented herein, identified residues critical for function or/and substrate specificity located in TMSs 1, 3, 6 and 8 that compose the binding site. A multiple sequence alignment of characterized NCS1 proteins and predicted binding site TMSs for FurD are shown in **Figure 69**.

The residues identified in TMS1 as critical for function were Ser85 in FcyB and Asn54 in FurD, the latter corresponds to Gln42 in Mhp1, which was also found to be important for function. Based on the multiple sequence alignment shown in **Figure 69**, Asn54 in FurD corresponds to Val83 in FcyB, which was not shown as critical for function. Topologically respective residues in TMS3, Asn163/Gln134 in FcyB/FurD, were identified to participate in substrate binding and specificity. This position in TMS3 is differentially conserved in NCS1 proteins and varies from an Asn in fungal Fcy-like proteins to a Gln in fungal Fur-like proteins and Mhp1, an Asp in plant NCS1 homologues and a Leu in the uridine transporter Fui1 from *S. cerevisiae* (**Figure 69**). The replacement of Asn163 in FcyB with a Leu or a Gln, residues found in FurD and other NCS1 homologues, did not lead to the binding or transport of "new" substrates suggesting that substrate specificity of FcyB is not determined solely by residues of the binding site.

Residues Trp159/Trp130 (TMS3) and Asn354/Asn341 (TMS8) in FcyB/FurD, corresponding to Trp117 and Asn318 in Mhp1, were proved irreplaceable for function. These residues are absolutely conserved in the NCS1 family and are predicted to contribute in substrate binding by π-π stacking interactions (Trp) and H-bonding (Asn). In FcyB, a second Asn residue (Asn350) conserved in all NCS1 proteins, additional to Asn354 (TMS8), was proved critical for function. Both Asn residues in FcyB affected the transport capacity as well as substrate specificity, whereas Asn341 in FurD was proved to be involved in transport function. In FcyB, Asn354 seems to interact through hydrogen bonding with all substrates except adenine, while Asn350 could indirectly affect substrate binding by affecting the architecture of the substrate binding site. Pro353 in FcyB, an also irreplaceable residue for transport, is probably affecting similarly to Asn350 substrate binding or by setting the neighboring Asn354 in the proper position for binding.

Interestingly, the conserved aromatic residues in TMS6 were shown to have different roles in FcyB and FurD. The position corresponding to Trp259/Tyr250 in FcyB/FurD is always occupied by an aromatic amino acid, highly conserved among NCS1 subgroups, a Trp in Fcys and plant homologues, a Phe in Furs and a Tyr in FurD. The only exception is FurE, which instead of an aromatic residue has a Lys. Trp259 in FcyB was shown to be irreplaceable for function and probably participates in π-π interactions with substrates. Tyr250 in FurD was shown to participate in substrate binding and to determine substrate specificity. Contrasting the irreplaceable Trp259 of FcyB, replacement of Tyr250 in FurD with an Ala led to a transporter with the capacity to take up a non-physiological substrate, allantoin, but not xanthine or uric acid. Other substitutions made herein (with a Trp or a Phe) also led to functional versions of FurD, which variably affected substrate affinities for the physiological substrates. Similarly, substitution of the neighboring Asn (Asn249) with an Ala led, again, to a fully functional transporter with altered specificity and increased capacity to uptake all its physiological substrates and allantoin.

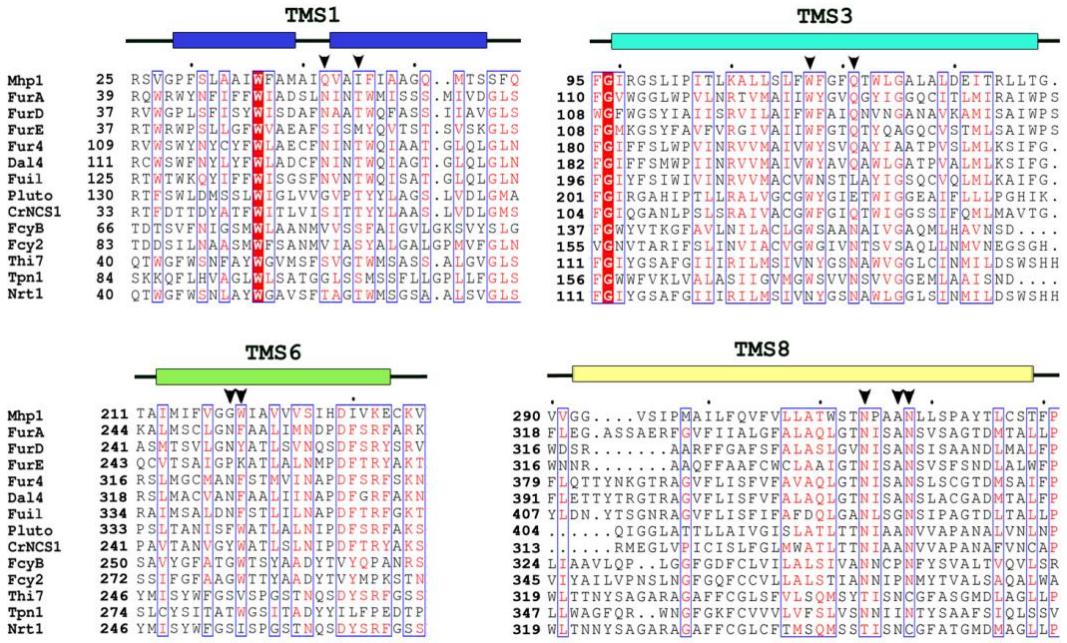


Figure 69. Multiple sequence alignment of FurD predicted TMSs that constitute the binding site, and characterized NCS1 proteins. Invariant and highly conserved amino acids are shaded in red and blue-lined boxes, respectively. Amino acids essential for function or specificity for FurD and FcyB are highlighted with black arrows. The alignment is part of the full length sequence alignment shown in **Figure 60**.

In order to verify whether the residues of TMS6 corresponding to Asn249 and Tyr250 of FurD determine the specificity, also, in other Fur proteins, we replaced the respective residues (Pro251 and Lys252), found in FurE, with those found in most Furs. Thus, mutations P251N and K252F, led to high affinity binding of uracil and allantoin (38-100-fold increase) and also binding of new substrates (xanthine, hydantoin, thymine). In conclusion, this region in TMS6 was identified as critical in substrate specificity for both FurD and FurE, but not FcyB.

Residues of the substrate binding site of PLUTO/AtNCS1, a uracil-adenine-guanine transporter from *A. thaliana*, mutated in the work of Witz et al. (2014) indicated residues similar with those of FcyB and FurD to participate in substrate binding (**Table 13**). Briefly, the substitutions of the relevant residues led to identification of a Glu in TMS3 as critical for function. The additional residues mutated, including the aromatic residues of TMS3 and TMS6, were shown to be of less importance for PLUTO function as they had minor effect in transport capacity. Interestingly, Trp223 (TMS3) and Phe341 (TMS6) are predicted to be involved in substrate specificity.

A Met residue in TMS10 of FurD (Met389) was shown to be critical for the specificity of FurD. Actually, when mutated to an Ala led to an extention of substrate specificity (M389A led to low affinity uptake of allantoin and hypoxanthine). This residue corresponds to a highly conserved Leu residue in Mhp1, the rest Furs and plant NCS1 proteins, but a Met in all Fcys and FurD (see alignment in **Figure 63**). Additionally, it was shown that in Mhp1 this Leu was responsible for the incapacity

of Mhp1 to take up a bulky inhibitor (Simmons et al., 2014). When we combined the FurD mutation M389A with specificity mutants of the binding site, we created FurD alleles with increased capacity and affinity for non-physiological substrates (allantoin and hypoxanthine). Thus, similarly to Mhp1 and UapA, we identified a residue in FurD not located in the major binding site that affects its substrate specificity profile indicating the existence of elements outside the binding site participating in substrate selectivity (see also later).

In summary, for the NCS1 family, residues of TMS1 and TMS3 are predicted to be involved in transport function and capacity, while residues of TMS6 and TMS8 are predicted to affect substrate affinity and specificity (**Table 13**). The variable role of residues identified in FurD and FcyB in topologically respective positions, mainly in TMS6 and TMS8, could be due to the functional diversity of the two subfamilies, Fur and Fcy. The flexible TMS10 was shown, at least for Mhp1 and FurD, to be involved in substrate selectivity without grossly affecting substrate affinities for physiological substrates or transport capacity.

Table 13. Residues important for transport function in Mhp1, FcyB, FurD and PLUTO. Residues irreplaceable for function are in red, more or less critical for function are in black and those affecting substrate specificity are in bold.

	TMS1	TMS3		TMS6		TMS8		TMS10
Mhp1	Gln42	Trp117	Gln121	-	Trp220	Asn314	Asn318	**Leu363**
FcyB	Ser85	Trp159	Asn163	-	Trp259	Asn350	**Asn354**	-
FurD	Asn54	Trp130	Gln134	**Asn249**	**Tyr250**	-	Asn341	**Met389**
PLUTO	Gly147	**Trp223**	Glu227	**Phe341**	Trp342	Asn426	Asn430	-

4.4 Transporter gating elements as channel-like selectivity filters

In an increasing number of transporters, recent findings point out residues located elsewhere than in the major binding site, which control occlusion or opening of the binding site, and thus affect the transport of substrates (e.g. LeuT, Mhp1; see also Introduction). Herein, we identified a residue in TMS10 of FurD that when mutated affects the uptake of a non-physiological substrates (see above and Results). Based on the crystal structures of Mhp1 and biochemical data (Shimamura et al., 2010; Simmons et al., 2014; Weyand et al., 2008), TMS10 is predicted to act as a flexible gate that bends over the occluded substrate binding site and triggers conformational changes from an outward to an inward state. The apparent ability of TMS10 in both Mhp1 and FurD to distinguish physiological from non-physiological substrates resembles to selectivity filters usually present in channels. Additionally, a plethora of genetic data have been published for UapA that report the ability of specific mutants mainly outside the binding site to bind or/and uptake non physiological substrates (for a review see Diallinas 2014).

The idea supporting the existence of channel-like gates in transporters was first reported by Papageorgiou et al. (2008) based on observations concerning UapA specificity mutants, and recent data on different transporters seem to support it. Based on our results for FurD, we propose that the specificity of a transporter is not determined solely by the residues comprising its binding pocket, but also by more distant elements. Based on our results, these elements are predicted to interact through hydrogen bonding or other active interactions with the putative substrate and thus co-operate in transport catalysis. A potential steric clash or "misfit" of a substrate/ligand with the gating elements, even though it can fit or even bind in the binding pocket might block the necessary conformational cascade needed for transport. In conclusion, the current study supports and contributes to the hypothesis that active selectivity *gates* are an intrinsic element of the transport function present in transporters and dynamically determine, along with the major binding site, the overall kinetics and specificity profile of a transporter.

4.5 Could transporter studies lead to rational design of more effective drugs?

The majority of drugs used currently are small hydrophobic compounds that enter the cell through the plasma membrane by slow non-facilitated diffusion and were selected based on random screening methods. Only few cases have been reported, where the drugs target specific transporters or channels in order to enter the cell or to block the transport proteins themselves (for a review see Diallinas 2014).

Molecules targeting specifically membrane proteins present only at the surface of pathogenic or cancer cells could be more efficient in disease treatment without any undesirable side effects. In order to design these molecules, what is necessary, is the elucidation of the function, properties and molecular mechanism of transport proteins found in pathogenic microbial organisms and cancer cells. Computational, biochemical and genetic approaches are necessary in order to develop efficient drug molecules.

An example of combined knowledge in the process of drug discovery was presented in the current study for the FcyB transporter, where rationally designed purine analogues inhibited competitively and specifically FcyB-mediated hypoxanthine uptake. Recently, unpublished results of G. Sioupouli and G. Diallinas confirmed the specific high affinity binding of few purine analogues. FcyB shows increased homology with Fcy-like proteins of *A. fumigatus* and *C. albicans*, two widespread pathogens, and thus, any inhibitor designed for FcyB could be finally tested for the two pathogens.

We believe that the current study has contributed in the knowledge of the transport mechanism met in fungal NCS1, NCS2 and AzgA-like transporters and has revealed elements critical for transport function that could be exploited in rational design of novel antifungals.

5 References

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., and Iwata, S. (2003). Structure and mechanism of the lactose permease of Escherichia coli. Science *301*, 610–615.

Amillis, S., Cecchetto, G., Sophianopoulou, V., Koukaki, M., Scazzocchio, C., and Diallinas, G. (2004). Transcription of purine transporter genes is activated during the isotropic growth phase of Aspergillus nidulans conidia. Mol. Microbiol. *52*, 205–216.

Amillis, S., Hamari, Z., Roumelioti, K., Scazzocchio, C., and Diallinas, G. (2007). Regulation of expression and kinetic modeling of substrate interactions of a uracil transporter in Aspergillus nidulans. Mol. Membr. Biol. *24*, 206–214.

Amillis, S., Kosti, V., Pantazopoulou, A., Mikros, E., and Diallinas, G. (2011). Mutational analysis and modeling reveal functionally critical residues in transmembrane segments 1 and 3 of the UapA transporter. J. Mol. Biol. *411*, 567–580.

Andersson, D.I., Jerlström-Hultqvist, J., and Näsvall, J. (2015). Evolution of New Functions De Novo and from Preexisting Genes. Cold Spring Harb. Perspect. Biol. *7*, a017996.

Apostolaki, A., Erpapazoglou, Z., Harispe, L., Billini, M., Kafasla, P., Kizis, D., Peñalva, M.A., Scazzocchio, C., and Sophianopoulou, V. (2009). AgtA, the dicarboxylic amino acid transporter of Aspergillus nidulans, is concertedly down-regulated by exquisite sensitivity to nitrogen metabolite repression and ammonium-elicited endocytosis. Eukaryot. Cell *8*, 339–352.

Aryal, P., Sansom, M.S.P., and Tucker, S.J. (2015). Hydrophobic gating in ion channels. J. Mol. Biol. *427*, 121–130.

Bayram, O., Braus, G.H., Fischer, R., and Rodriguez-Romero, J. (2010). Spotlight on Aspergillus nidulans photosensory systems. Fungal Genet. Biol. *47*, 900–908.

Blondel, M.-O., Morvan, J., Dupré, S., Urban-Grimal, D., Haguenauer-Tsapis, R., and Volland, C. (2004). Direct sorting of the yeast uracil permease to the endosomal system is controlled by uracil binding and Rsp5p-dependent ubiquitylation. Mol. Biol. Cell *15*, 883–895.

Boudker, O., Ryan, R.M., Yernool, D., Shimamoto, K., and Gouaux, E. (2007). Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. Nature *445*, 387–393.

Brèthes, D., Chirio, M.C., Napias, C., Chevallier, M.R., Lavie, J.L., and Chevallier, J. (1992). In vivo and in vitro studies of the purine-cytosine permease of Saccharomyces cerevisiae. Functional analysis of a mutant with an altered apparent transport constant of uptake. Eur. J. Biochem. *204*, 699–704.

Cain, N.E., and Kaiser, C.A. (2011). Transport activity-dependent intracellular sorting of the yeast general amino acid permease. Mol. Biol. Cell *22*, 1919–1929.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973.

Cecchetto, G., Amillis, S., Diallinas, G., Scazzocchio, C., and Drevet, C. (2004). The AzgA purine transporter of Aspergillus nidulans. Characterization of a protein belonging to a new phylogenetic cluster. J. Biol. Chem. *279*, 3132–3141.

Celik, L., Schiøtt, B., and Tajkhorshid, E. (2008). Substrate binding and formation of an occluded state in the leucine transporter. Biophys. J. *94*, 1600–1612.

Chaptal, V., Kwon, S., Sawaya, M.R., Guan, L., Kaback, H.R., and Abramson, J. (2011). Crystal structure of lactose permease in complex with an affinity inactivator yields unique insight into sugar recognition. Proc. Natl. Acad. Sci. U. S. A. *108*, 9361–9366.

Le Coutre, J., and Kaback, H.K. (2000). Structure-function relationships of integral membrane proteins: membrane transporters vs channels. Biopolymers *55*, 297–307.

Cunningham, P., and Naftalin, R.J. (2013). Implications of aberrant temperature-sensitive glucose transport via the glucose transporter deficiency mutant (GLUT1DS) T295M for the alternate-access and fixed-site transport models. J. Membr. Biol. *246*, 495–511.

Cunningham, P., Afzal-Ahmed, I., and Naftalin, R.J. (2006). Docking studies show that D-glucose and quercetin slide through the transporter GLUT1. J. Biol. Chem. *281*, 5797–5803.

Dang, S., Sun, L., Huang, Y., Lu, F., Liu, Y., Gong, H., Wang, J., and Yan, N. (2010). Structure of a fucose transporter in an outward-open conformation. Nature *467*, 734–738.

Darlington, A.J., and Scazzocchio, C. (1967). Use of analogues and the substrate-sensitivity of mutants in analysis of purine uptake and breakdown in Aspergillus nidulans. J. Bacteriol. *93*, 937–940.

Darlington, A.J., Scazzocchio, C., and Pateman, J.A. (1965). Biochemical and genetical studies of purine breakdown in Aspergillus. Nature *206*, 599–600.

David, H., Ozçelik, I.S., Hofmann, G., and Nielsen, J. (2008). Analysis of Aspergillus nidulans metabolism at the genome-scale. BMC Genomics *9*, 163.

Delmas, S., Pullan, S.T., Gaddipati, S., Kokolski, M., Malla, S., Blythe, M.J., Ibbett, R., Campbell, M., Liddell, S., Aboobaker, A., et al. (2012). Uncovering the genome-wide transcriptional responses of the filamentous fungus Aspergillus niger to lignocellulose using RNA sequencing. PLoS Genet. *8*, e1002875.

Deng, D., Xu, C., Sun, P., Wu, J., Yan, C., Hu, M., and Yan, N. (2014). Crystal structure of the human glucose transporter GLUT1. Nature *510*, 121–125.

Diallinas, G. (2013). Allopurinol and xanthine use different translocation mechanisms and trajectories in the fungal UapA transporter. Biochimie *95*, 1755–1764.

Diallinas, G. (2014). Understanding transporter specificity and the discrete appearance of channel-like gating domains in transporters. Front. Pharmacol. *5*, 207.

Diallinas, G., and Gournas, C. (2008). Structure-function relationships in the nucleobase-ascorbate transporter (NAT) family: Lessons from model microbial genetic systems. Channels *2*, 363–372.

Diallinas, G., and Scazzocchio, C. (1989). A gene coding for the uric acid-xanthine permease of Aspergillus nidulans: inactivational cloning, characterization, and sequence of a cis-acting mutation. Genetics *122*, 341–350.

Diallinas, G., Gorfinkiel, L., Arst, H.N., Cecchetto, G., and Scazzocchio, C. (1995). Genetic and molecular characterization of a gene encoding a wide specificity purine permease of Aspergillus nidulans reveals a novel family of transporters conserved in prokaryotes and eukaryotes. J. Biol. Chem. *270*, 8610–8622.

Diallinas, G., Valdez, J., Sophianopoulou, V., Rosa, A., and Scazzocchio, C. (1998). Chimeric purine transporters of Aspergillus nidulans define a domain critical for function and specificity conserved in bacterial, plant and metazoan homologues. EMBO J. *17*, 3827–3837.

Dubyak, G.R. (2004). Ion homeostasis, channels, and transporters: an update on cellular mechanisms. Adv. Physiol. Educ. *28*, 143–154.

Edgerton-Morgan, H., and Oakley, B.R. (2012). γ-Tubulin plays a key role in inactivating APC/C(Cdh1) at the G(1)-S boundary. J. Cell Biol. *198*, 785–791.

Ethayathulla, A.S., Yousef, M.S., Amin, A., Leblanc, G., Kaback, H.R., and Guan, L. (2014). Structure-based mechanism for Na(+)/melibiose symport by MelB. Nat. Commun. *5*, 3009.

Faaland, C.A., Race, J.E., Ricken, G., Warner, F.J., Williams, W.J., and Holtzman, E.J. (1998). Molecular characterization of two novel transporters from human and mouse kidney and from LLC-PK1 cells reveals a novel conserved family that is homologous to bacterial and Aspergillus nucleobase transporters. Biochim. Biophys. Acta *1442*, 353–360.

Forrest, L.R., Krämer, R., and Ziegler, C. (2011). The structural basis of secondary active transport mechanisms. Biochim. Biophys. Acta *1807*, 167–188.

Frisvad, J.C. (2014). Taxonomy, chemodiversity, and chemoconsistency of Aspergillus, Penicillium, and Talaromyces species. Front. Microbiol. *5*, 773.

Fukuda, M., Takeda, H., Kato, H.E., Doki, S., Ito, K., Maturana, A.D., Ishitani, R., and Nureki, O. (2015). Structural basis for dynamic mechanism of nitrate/nitrite antiport by NarK. Nat. Commun. *6*, 7097.

Gabriel, F., Sabra, A., El-Kirat-Chatel, S., Pujol, S., Fitton-Ouhabi, V., Brèthes, D., Dementhon, K., Accoceberry, I., and Noël, T. (2014). Deletion of the uracil permease gene confers cross-resistance to 5-fluorouracil and azoles in Candida lusitaniae and highlights antagonistic interaction between fluorinated nucleotides and fluconazole. Antimicrob. Agents Chemother. *58*, 4476–4485.

Gabrielsen, M., Kurczab, R., Ravna, A.W., Kufareva, I., Abagyan, R., Chilmonczyk, Z., Bojarski, A.J., and Sylte, I. (2012). Molecular mechanism of serotonin transporter inhibition elucidated by a new flexible docking protocol. Eur. J. Med. Chem. *47*, 24–37.

Gadsby, D.C. (2009). Ion channels versus ion pumps: the principal difference, in principle. Nat. Rev. Mol. Cell Biol. *10*, 344–352.

Galanopoulou, K., Scazzocchio, C., Eleftheria, M., Liu, W., Borbolis, F., Karachaliou, M., Oestreicher, N., Hatzinikolaou, D.G., Diallinas, G., and Amillis, S. (2014). Purine utilization proteins in the Eurotiales : Cellular compartmentalization , phylogenetic conservation and divergence. Fungal Genet. Biol. *69*, 96–108.

Garrett, R., and Grisham, C. (2012). Biochemistry (Cengage Learning).

Ghaddar, K., Merhi, A., Saliba, E., Krammer, E.-M., Prévost, M., and André, B. (2014). Substrate-induced ubiquitylation and endocytosis of yeast amino acid permeases. Mol. Cell. Biol. *34*, 4447–4463.

Goldberg, A.L. (2003). Protein degradation and protection against misfolded or damaged proteins. Nature *426*, 895–899.

Goldman, G.H., and Osmani, S.A. (2008). The Aspergilli: Genomics, medical aspects, biotechnology, and research methods. (CRC Press).

Gorfinkiel, L., Diallinas, G., and Scazzocchio, C. (1993). Sequence and regulation of the uapA gene encoding a uric acid-xanthine permease in the fungus Aspergillus nidulans. J. Biol. Chem. *268*, 23376–23381.

Goudela, S., Tsilivi, H., and Diallinas, G. (2006). Comparative kinetic analysis of AzgA and Fcy21p, prototypes of the two major fungal hypoxanthine-adenine-guanine transporter families. Mol. Membr. Biol. *23*, 291–303.

Goudela, S., Reichard, U., Amillis, S., and Diallinas, G. (2008). Characterization and kinetics of the major purine transporters in Aspergillus fumigatus. Fungal Genet. Biol. *45*, 459–472.

Gournas, C., Papageorgiou, I., and Diallinas, G. (2008). The nucleobase-ascorbate transporter (NAT) family: genomics, evolution, structure-function relationships and physiological role. Mol. Biosyst. *4*, 404–416.

Gournas, C., Amillis, S., Vlanti, A., and Diallinas, G. (2010). Transport-dependent endocytosis and turnover of a uric acid-xanthine permease. Mol. Microbiol. *75*, 246–260.

Guan, L., Mirza, O., Verner, G., Iwata, S., and Kaback, H.R. (2007). Structural determination of wild-type lactose permease. Proc. Natl. Acad. Sci. U. S. A. *104*, 15294–15298.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

Hamari, Z., Amillis, S., Drevet, C., Apostolaki, A., Vágvölgyi, C., Diallinas, G., and Scazzocchio, C. (2009). Convergent evolution and orphan genes in the Fur4p-like family and characterization of a general nucleoside transporter in Aspergillus nidulans. Mol. Microbiol. *73*, 43–57.

Hey, J. (2004). What's So Hot about Recombination Hotspots? PLoS Biol. *2*, e190.

Huang, Y., Lemieux, M.J., Song, J., Auer, M., and Wang, D.-N. (2003). Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli. Science *301*, 616–620.

Iancu, C. V, Zamoon, J., Woo, S.B., Aleshin, A., and Choe, J. (2013). Crystal structure of a glucose/H+ symporter and its mechanism of action. Proc. Natl. Acad. Sci. U. S. A. *110*, 17862–17867.

Jain, S., Sutchu, S., Rosa, P. a, Byram, R., and Jewett, M.W. (2012). Borrelia burgdorferi harbors a transport system essential for purine salvage and mammalian infection. Infect. Immun. *80*, 3086–3093.

Jardetzky, O. (1966). Simple allosteric model for membrane pumps. Nature *211*, 969–970.

Jund, R., Weber, E., and Chevallier, M.R. (1988). Primary structure of the uracil transport protein of Saccharomyces cerevisiae. Eur. J. Biochem. *171*, 417–424.

Kaback, H.R. (2015). A chemiosmotic mechanism of symport. Proc. Natl. Acad. Sci. U. S. A. *112*, 1259–1264.

Kaback, H.R., Smirnova, I., Kasho, V., Nie, Y., and Zhou, Y. (2011). The alternating access transport mechanism in LacY. J. Membr. Biol. *239*, 85–93.

Käfer, E. (1965). Origins of translocations in Aspergillus nidulans. Genetics *52*, 217–232.

Karachaliou, M., Amillis, S., Evangelinos, M., Kokotos, A.C., Yalelis, V., and Diallinas, G. (2013). The arrestin-like protein ArtA is essential for ubiquitination and endocytosis of the UapA transporter in response to both broad-range and specific signals. Mol. Microbiol. *88*, 301–317.

Karena, E., and Frillingos, S. (2011). The role of transmembrane segment TM3 in the xanthine permease XanQ of Escherichia coli. J. Biol. Chem. *286*, 39595–39605.

Keener, J.M., and Babst, M. (2013). Quality control and substrate-dependent downregulation of the nutrient transporter Fur4. Traffic *14*, 412–427.

Keller, R., Ziegler, C., and Schneider, D. (2014). When two turn into one: evolution of membrane transporters from half modules. Biol. Chem. *395*, 1379–1388.

Kim, H., Han, K., Kim, K., Han, D., Jahng, K., and Chae, K. (2002). The veA gene activates sexual development in Aspergillus nidulans. Fungal Genet. Biol. *37*, 72–80.

De Koning, H., and Diallinas, G. (2000). Nucleobase transporters (review). Mol. Membr. Biol. *17*, 75–94.

Kosti, V., Papageorgiou, I., and Diallinas, G. (2010). Dynamic elements at both cytoplasmically and extracellularly facing sides of the UapA transporter selectively control the accessibility of substrates to their translocation pathway. J. Mol. Biol. *397*, 1132–1143.

Kosti, V., Lambrinidis, G., Myrianthopoulos, V., Diallinas, G., and Mikros, E. (2012). Identification of the substrate recognition and transport pathway in a eukaryotic member of the nucleobase-ascorbate transporter (NAT) family. PLoS One *7*, e41939.

Koukaki, M., Giannoutsou, E., Karagouni, A., and Diallinas, G. (2003). A novel improved method for Aspergillus nidulans transformation. J. Microbiol. Methods *55*, 687–695.

Koukaki, M., Vlanti, A., Goudela, S., Pantazopoulou, A., Gioule, H., Tournaviti, S., and Diallinas, G. (2005). The nucleobase-ascorbate transporter (NAT) signature motif in UapA defines the function of the purine translocation pathway. J. Mol. Biol. *350*, 499–513.

Krishnamurthy, H., and Gouaux, E. (2012). X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. Nature *481*, 469–474.

Krishnamurthy, H., Piscitelli, C.L., and Gouaux, E. (2009). Unlocking the molecular secrets of sodium-coupled transporters. Nature *459*, 347–355.

Krypotou, E., and Diallinas, G. (2014). Transport assays in filamentous fungi: Kinetic characterization of the UapC purine transporter of Aspergillus nidulans. Fungal Genet. Biol. *63*, 1–8.

Krypotou, E., Kosti, V., Amillis, S., Myrianthopoulos, V., Mikros, E., and Diallinas, G. (2012). Modeling, substrate docking, and mutational analysis identify residues essential for the function and specificity of a eukaryotic purine-cytosine NCS1 transporter. J. Biol. Chem. *287*, 36792–36803.

Krypotou, E., Lambrinidis, G., Evangelidis, T., Mikros, E., and Diallinas, G. (2014). Modelling, substrate docking and mutational analysis identify residues essential for function and specificity of the major fungal purine transporter AzgA. Mol. Microbiol. *93*, 129–145.

Krypotou, E., Evangelidis, T., Bobonis, J., Pittis, A.A., Gabaldón, T., Scazzocchio, C., Mikros, E., and Diallinas, G. (2015a). Origin, diversification and substrate specificity in the family of NCS1/FUR transporters. Mol. Microbiol. *96*, 927–950.

Krypotou, E., Scazzocchio, C., and Diallinas, G. (2015b). Functional characterization of NAT/NCS2 proteins of Aspergillus brasiliensis reveals a genuine xanthine–uric acid transporter and an intrinsically misfolded polypeptide. Fungal Genet. Biol. *75*, 56–63.

Kumar, H., Kasho, V., Smirnova, I., Finer-Moore, J.S., Kaback, H.R., and Stroud, R.M. (2014). Structure of sugar-bound LacY. Proc. Natl. Acad. Sci. U. S. A. *111*, 1784–1788.

Lauwers, E., Erpapazoglou, Z., Haguenauer-Tsapis, R., and André, B. (2010). The ubiquitin code of yeast permease trafficking. Trends Cell Biol. *20*, 196–204.

Lu, F., Li, S., Jiang, Y., Jiang, J., Fan, H., Lu, G., Deng, D., Dang, S., Zhang, X., Wang, J., et al. (2011). Structure and mechanism of the uracil transporter UraA. Nature *472*, 243–246.

MacGurn, J. a, Hsu, P.-C., and Emr, S.D. (2012). Ubiquitin and Membrane Protein Turnover: From Cradle to Grave. Annu. Rev. Biochem. *81*, 231–259.

Madej, M.G., Soro, S.N., and Kaback, H.R. (2012). Apo-intermediate in the transport cycle of lactose permease (LacY). Proc. Natl. Acad. Sci. U. S. A. *109*, E2970–E2978.

Madej, M.G., Dang, S., Yan, N., and Kaback, H.R. (2013). Evolutionary mix-and-match with MFS transporters. Proc. Natl. Acad. Sci. U. S. A. *110*, 5870–5874.

Martzoukou, O., Karachaliou, M., Yalelis, V., Leung, J., Byrne, B., Amillis, S., and Diallinas, G. (2015). Oligomerization of the UapA Purine Transporter Is Critical for ER-Exit, Plasma Membrane Localization and Turnover. J. Mol. Biol. *427*, 2679–2696.

Meintanis, C., Karagouni, A.D., and Diallinas, G. Amino acid residues N450 and Q449 are critical for the uptake capacity and specificity of UapA, a prototype of a nucleobase-ascorbate transporter family. Mol. Membr. Biol. *17*, 47–57.

Mirza, O., Guan, L., Verner, G., Iwata, S., and Kaback, H.R. (2006). Structural evidence for induced fit and a mechanism for sugar/H+ symport in LacY. EMBO J. *25*, 1177–1183.

De Montigny, J., Straub, M.L., Wagner, R., Bach, M.L., and Chevallier, M.R. (1998). The uracil permease of Schizosaccharomyces pombe: a representative of a family of 10 transmembrane helix transporter proteins of yeasts. Yeast *14*, 1051–1059.

Moran, Y., Barzilai, M.G., Liebeskind, B.J., and Zakon, H.H. (2015). Evolution of voltage-gated ion channels at the emergence of Metazoa. J. Exp. Biol. *218*, 515–525.

Mourad, G.S., Tippmann-Crosby, J., Hunt, K. a, Gicheru, Y., Bade, K., Mansfield, T. a, and Schultes, N.P. (2012). Genetic and molecular characterization reveals a unique nucleobase cation symporter 1 in Arabidopsis. FEBS Lett. *586*, 1370–1378.

Mullaney, E.J., Hamer, J.E., Roberti, K.A., Yelton, M.M., and Timberlake, W.E. (1985). Primary structure of the trpC gene from Aspergillus nidulans. Mol. Gen. Genet. *199*, 37–45.

Van Munster, J.M., Daly, P., Delmas, S., Pullan, S.T., Blythe, M.J., Malla, S., Kokolski, M., Noltorp, E.C.M., Wennberg, K., Fetherston, R., et al. (2014). The role of carbon starvation in the induction of enzymes that degrade plant-derived carbohydrates in Aspergillus niger. Fungal Genet. Biol. *72*, 34–47.

Naftalin, R.J. (2010). Reassessment of models of facilitated transport and cotransport. J. Membr. Biol. *234*, 75–112.

Nayak, T., Szewczyk, E., Oakley, C.E., Osmani, A., Ukil, L., Murray, S.L., Hynes, M.J., Osmani, S.A., and Oakley, B.R. (2006). A versatile and efficient gene-targeting system for Aspergillus nidulans. Genetics *172*, 1557–1566.

Newstead, S., Drew, D., Cameron, A.D., Postis, V.L.G., Xia, X., Fowler, P.W., Ingram, J.C., Carpenter, E.P., Sansom, M.S.P., McPherson, M.J., et al. (2011). Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2. EMBO J. *30*, 417–426.

Palmer, L.M., and Cove, D.J. (1975). Pyrimidine biosynthesis in Aspergillus nidulans: isolation and preliminary characterisation of auxotrophic mutants. Mol. Gen. Genet. *138*, 243–255.

Pantazopoulou, A., and Diallinas, G. (2009). The first transmembrane segment (TMS1) of UapA contains determinants necessary for expression in the plasma membrane and purine transport. Mol. Membr. Biol. *23*, 337–348.

Pantazopoulou, A., Lemuh, N.D., Hatzinikolaou, D.G., Drevet, C., Cecchetto, G., Scazzocchio, C., and Diallinas, G. (2007). Differential physiological and developmental expression of the UapA and AzgA purine transporters in Aspergillus nidulans. Fungal Genet. Biol. *44*, 627–640.

Papageorgiou, I., Gournas, C., Vlanti, A., Amillis, S., Pantazopoulou, A., and Diallinas, G. (2008). Specific Interdomain Synergy in the UapA Transporter Determines Its Unique Specificity for Uric Acid among NAT Carriers. J. Mol. Biol. *382*, 1121–1135.

Papakostas, K., Botou, M., and Frillingos, S. (2013). Functional identification of the hypoxanthine/guanine transporters YjcD and YgfQ and the adenine transporters PurP and YicO of Escherichia coli K-12. J. Biol. Chem. *288*, 36827–36840.

Pedersen, B.P., Kumar, H., Waight, A.B., Risenmay, A.J., Roe-Zurz, Z., Chau, B.H., Schlessinger, A., Bonomi, M., Harries, W., Sali, A., et al. (2013). Crystal structure of a eukaryotic phosphate transporter. Nature *496*, 533–536.

Piscitelli, C.L., and Gouaux, E. (2012). Insights into transport mechanism from LeuT engineered to transport tryptophan. EMBO J. *31*, 228–235.

Punt, P.J., Dingemanse, M.A., Jacobs-Meijsing, B.J.M., Pouwels, P.H., and van den Hondel, C.A.M.J.J. (1988). Isolation and characterization of the glyceraldehyde-3-phosphate dehydrogenase gene of Aspergillus nidulans. Gene *69*, 49–57.

Punt, P.J., Dingemanse, M.A., Kuyvenhoven, A., Soede, R.D.M., Pouwels, P.H., and van den Hondel, C.A.M.J.J. (1990). Functional elements in the promoter region of the Aspergillus nidulans gpdA gene encoding glyceraldehyde-3-phosphate dehydrogenase. Gene *93*, 101–109.

Quistgaard, E.M., Löw, C., Moberg, P., and Nordlund, P. (2013). Metal-mediated crystallization of the xylose transporter XylE from Escherichia coli in three different crystal forms. J. Struct. Biol. *184*, 375–378.

Reddy, V.S., Shlykov, M.A., Castillo, R., Sun, E.I., and Saier, M.H. (2012). The major facilitator superfamily (MFS) revisited. FEBS J. *279*, 2022–2035.

Saier, M.H., Reddy, V.S., Tamang, D.G., and Västermark, A. (2014). The transporter classification database. Nucleic Acids Res. *42*, D251–D258.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning : a laboratory manual (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory).

Sá-Pessoa, J., Amillis, S., Casal, M., and Diallinas, G. (2015). Expression and specificity profile of the major acetate transporter AcpA in Aspergillus nidulans. Fungal Genet. Biol. *76*, 93–103.

Schein, J.R., Hunt, K. a, Minton, J. a, Schultes, N.P., and Mourad, G.S. (2013). The nucleobase cation symporter 1 of Chlamydomonas reinhardtii and that of the evolutionarily distant Arabidopsis thaliana display parallel function and establish a plant-specific solute transport profile. Plant Physiol. Biochem. *70*, 52–60.

Shi, Y. (2013). Common folds and transport mechanisms of secondary active transporters. Annu. Rev. Biophys. *42*, 51–72.

Shi, L., Quick, M., Zhao, Y., Weinstein, H., and Javitch, J.A. (2008). The mechanism of a neurotransmitter:sodium symporter--inward release of Na+ and substrate is triggered by substrate in a second binding site. Mol. Cell *30*, 667–677.

Shimamura, T., Yajima, S., Suzuki, S., Rutherford, N.G., O'Reilly, J., Henderson, P.J.F., and Iwata, S. (2008). Crystallization of the hydantoin transporter Mhp1 from Microbacterium liquefaciens. Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun. *64*, 1172–1174.

Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N.G., Hadden, J.M., Sharples, D., Sansom, M.S.P., Iwata, S., Henderson, P.J.F., and Cameron, A.D. (2010). Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. Science *328*, 470–473.

Sibthorp, C., Wu, H., Cowley, G., Wong, P.W.H., Palaima, P., Morozov, I.Y., Weedall, G.D., and Caddick, M.X. (2013). Transcriptome analysis of the filamentous fungus Aspergillus nidulans directed to the global identification of promoters. BMC Genomics *14*, 847.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Simmons, K.J., Jackson, S.M., Brueckner, F., Patching, S.G., Beckstein, O., Ivanova, E., Geng, T., Weyand, S., Drew, D., Lanigan, J., et al. (2014). Molecular mechanism of ligand recognition by membrane transport protein, Mhp1. EMBO J. *21*, 11–27.

Singh, S.K. (2014). LeuT: A prokaryotic stepping stone on the way to a eukaryotic neurotransmitter transporter structure. Channels *2*, 380–389.

Singh, S.K., and Pal, A. (2015). Biophysical Approaches to the Study of LeuT, a Prokaryotic Homolog of Neurotransmitter Sodium Symporters. Methods Enzymol. *557*, 167–198.

Singh, S.K., Piscitelli, C.L., Yamashita, A., and Gouaux, E. (2008). A competitive inhibitor traps LeuT in an open-to-out conformation. Science *322*, 1655–1661.

Slotboom, D.J. (2014). Structural and mechanistic insights into prokaryotic energy-coupling factor transporters. Nat. Rev. Microbiol. *12*, 79–87.

Smirnova, I., Kasho, V., and Kaback, H.R. (2011). Lactose permease and the alternating access mechanism. Biochemistry *50*, 9684–9693.

Smirnova, I., Kasho, V., Sugihara, J., and Kaback, H.R. (2013). Trp replacements for tightly interacting Gly-Gly pairs in LacY stabilize an outward-facing conformation. Proc. Natl. Acad. Sci. U. S. A. *110*, 8876–8881.

Smirnova, I., Kasho, V., and Kaback, H.R. (2014). Real-time conformational changes in LacY. Proc. Natl. Acad. Sci. U. S. A. *111*, 8440–8445.

Solcan, N., Kwok, J., Fowler, P.W., Cameron, A.D., Drew, D., Iwata, S., and Newstead, S. (2012). Alternating access mechanism in the POT family of oligopeptide transporters. EMBO J. *31*, 3411–3421.

Southern, E. (2006). Southern blotting. Nat. Protoc. *1*, 518–525.

Springael, J.Y., and André, B. (1998). Nitrogen-regulated ubiquitination of the Gap1 permease of Saccharomyces cerevisiae. Mol. Biol. Cell *9*, 1253–1263.

Stelzl, L.S., Fowler, P.W., Sansom, M.S.P., and Beckstein, O. (2014). Flexible gates generate occluded intermediates in the transport cycle of LacY. J. Mol. Biol. *426*, 735–751.

Stinnett, S.M., Espeso, E.A., Cobeño, L., Araújo-Bazán, L., and Calvo, A.M. (2007). Aspergillus nidulans VeA subcellular localization is dependent on the importin alpha carrier and on light. Mol. Microbiol. *63*, 242–255.

Sun, J., Bankston, J.R., Payandeh, J., Hinds, T.R., Zagotta, W.N., and Zheng, N. (2014). Crystal structure of the plant dual-affinity nitrate transporter NRT1.1. Nature *507*, 73–77.

Sun, L., Zeng, X., Yan, C., Sun, X., Gong, X., Rao, Y., and Yan, N. (2012). Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. Nature *490*, 361–366.

Susarla, B.T.S., and Robinson, M.B. (2008). Internalization and degradation of the glutamate transporter GLT-1 in response to phorbol ester. Neurochem. Int. *52*, 709–722.

Tavoularis, S., Scazzocchio, C., and Sophianopoulou, V. (2001). Functional expression and cellular localization of a green fluorescent protein-tagged proline transporter in Aspergillus nidulans. Fungal Genet. Biol. *33*, 115–125.

Todd, R.B., Davis, M.A., and Hynes, M.J. (2007). Genetic manipulation of Aspergillus nidulans: meiotic progeny for genetic analysis and strain construction. Nat. Protoc. *2*, 811–821.

Tsukaguchi, H., Tokui, T., Mackenzie, B., Berger, U. V, Chen, X.Z., Wang, Y., Brubaker, R.F., and Hediger, M.A. (1999). A family of mammalian Na+-dependent L-ascorbic acid transporters. Nature *399*, 70–75.

Valdez-Taubas, J., Diallinas, G., Scazzocchio, C., and Rosa, A.L. (2000). Protein expression and subcellular localization of the general purine transporter UapC from Aspergillus nidulans. Fungal Genet. Biol. *30*, 105–113.

Valdez-Taubas, J., Harispe, L., Scazzocchio, C., Gorfinkiel, L., and Rosa, A.L. (2004). Ammonium-induced internalisation of UapC, the general purine permease from Aspergillus nidulans. Fungal Genet. Biol. *41*, 42–51.

Vastermark, A., Wollwage, S., Houle, M.E., Rio, R., and Saier, M.H. (2014). Expansion of the APC superfamily of secondary carriers. Proteins Struct. Funct. Bioinforma. *82*, 2797–2811.

Västermark, Å., and Saier, M.H. (2014). Evolutionary relationship between 5+5 and 7+7 inverted repeat folds within the amino acid-polyamine-organocation superfamily. Proteins Struct. Funct. Bioinforma. *82*, 336–346.

Vlanti, A., and Diallinas, G. (2008). The Aspergillus nidulans FcyB cytosine-purine scavenger is highly expressed during germination and in reproductive compartments and is downregulated by endocytosis. Mol. Microbiol. *68*, 959–977.

Vlanti, A., Amillis, S., Koukaki, M., and Diallinas, G. (2006). A novel-type substrate-selectivity filter and ER-exit determinants in the UapA purine transporter. J. Mol. Biol. *357*, 808–819.

Wagner, R., Montigny, J., Wergifosse, P., Souciet, J.., and Potier, S. (1998). The ORF YBL042 of Saccharomyces cerevisiae encodes a uridine permease. FEMS Microbiol. Lett. *159*, 69–75.

Wang, K.H., Penmatsa, A., and Gouaux, E. (2015). Neurotransmitter and psychostimulant recognition by the dopamine transporter. Nature *521*, 322–327.

Weber, E., Rodriguez, C., Chevallier, M.R., and Jund, R. (1990). The purine-cytosine permease gene of Saccharomyces cerevisiae: primary structure and deduced protein sequence of the FCY2 gene product. Mol. Microbiol. *4*, 585–596.

Weiss, J.N. (1997). The Hill equation revisited: uses and misuses. FASEB J. *11*, 835–841.

Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E.P., Rutherford, N.G., Hadden, J.M., O'Reilly, J., et al. (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. Science *322*, 709–713.

Weyand, S., Shimamura, T., Beckstein, O., Sansom, M.S.P., Iwata, S., Henderson, P.J.F., and Cameron, A.D. (2011). The alternating access mechanism of transport as observed in the sodium-hydantoin transporter Mhp1. J. Synchrotron Radiat. *18*, 20–23.

Widdas, W.F. (1952). Inability of diffusion to account for placental glucose transfer in the sheep and consideration of the kinetics of a possible carrier transfer. J. Physiol. *118*, 23–39.

Witz, S., Jung, B., Fürst, S., and Möhlmann, T. (2012). De novo pyrimidine nucleotide synthesis mainly occurs outside of plastids, but a previously undiscovered nucleobase importer provides substrates for the essential salvage pathway in Arabidopsis. Plant Cell *24*, 1549–1559.

Witz, S., Panwar, P., Schober, M., Deppe, J., Pasha, F.A., Lemieux, M.J., and Möhlmann, T. (2014). Structure-function relationship of a plant NCS1 member--homology modeling and mutagenesis identified residues critical for substrate specificity of PLUTO, a nucleobase transporter from Arabidopsis. PLoS One *9*, e91343.

Wong, F.H., Chen, J.S., Reddy, V., Day, J.L., Shlykov, M.A., Wakabayashi, S.T., and Saier, M.H. (2012). The amino acid-polyamine-organocation superfamily. J. Mol. Microbiol. Biotechnol. *22*, 105–113.

Yamashita, A., Singh, S.K., Kawate, T., Jin, Y., and Gouaux, E. (2005). Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. Nature *437*, 215–223.

Yan, H., Huang, W., Yan, C., Gong, X., Jiang, S., Zhao, Y., Wang, J., and Shi, Y. (2013). Structure and mechanism of a nitrate transporter. Cell Rep. *3*, 716–723.

Yernool, D., Boudker, O., Jin, Y., and Gouaux, E. (2004). Structure of a glutamate transporter homologue from Pyrococcus horikoshii. Nature *431*, 811–818.

Yin, Y., He, X., Szewczyk, P., Nguyen, T., and Chang, G. (2006). Structure of the multidrug transporter EmrD from Escherichia coli. Science *312*, 741–744.

Yoo, H.S., Cunningham, T.S., and Cooper, T.G. (1992). The allantoin and uracil permease gene sequences of Saccharomyces cerevisiae are nearly identical. Yeast *8*, 997–1006.

Van Zeebroeck, G., Rubio-Texeira, M., Schothorst, J., and Thevelein, J.M. (2014). Specific analogues uncouple transport, signalling, oligo-ubiquitination and endocytosis in the yeast Gap1 amino acid transceptor. Mol. Microbiol. *93*, 213–233.

Zhao, Y., Mao, G., Liu, M., Zhang, L., Wang, X., and Zhang, X.C. (2014). Crystal structure of the E. coli peptide transporter YbgH. Structure *22*, 1152–1160.

Zheng, H., Wisedchaisri, G., and Gonen, T. (2013). Crystal structure of a nitrate/nitrite exchanger. Nature *497*, 647–651.

Zhou, Z., Zhen, J., Karpowich, N.K., Goetz, R.M., Law, C.J., Reith, M.E.A., and Wang, D.-N. (2007). LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. Science *317*, 1390–1393.

6

Appendix:
Curriculum Vitae and publications

Curriculum Vitae

Education

- PhD in Molecular Microbiology: Faculty of Biology, University of Athens (2012-2016)
- BSc: Faculty of Biology, University of Athens (2006-2011), degree (7.84/10)

Laboratory skills and techniques

- Research activity since 2010 focused on the study of transporters using the model fungus *Aspergillus nidulans* in the laboratory of Prof. Diallinas.
- Handling and culture of microorganisms.
- Molecular biology-related techniques: DNA, RNA and protein isolation, cloning, PCR and PCR mutagenesis, SDS-page, Southern, Northern and Southern blot analysis.
- Kinetic analysis of transport function by radiolabeled substrate uptake experiments
- Live cell imaging of GFP tagged proteins by epifluorescence microscopy.
- Basic knowledge of protein structure model prediction by homology modeling.
- Average knowledge of phylogenetic tree construction and analysis.
- Great organizing and management skills, I have been for at least three years lab manager in Prof. Diallinas' laboratory.
- Experience in guiding and supervising undergraduate students and less experienced members of the lab during their scientific projects.
- Demonstrating and teaching skills, as I have worked as an assistant in undergraduate practical courses for 3 years.

Personal skills

- Proficient user of English
- Independent user of French.
- Very good user of computer

Grants and scholarships

- PhD research was co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALIS.
- Scholarship during my BSc studies by *John S. Latsis Public Benefit Foundation* (2006-2010).

Published papers

1. Krypotou E, Kosti V, Amillis S, Myrianthopoulos V, Mikros E, Diallinas G. Modeling, substrate docking, and mutational analysis identify residues essential for the function and specificity of a eukaryotic purine-cytosine NCS1 transporter. J Biol Chem. 2012, **287**: 36792-803.

2. Krypotou E, Diallinas G. Transport assays in filamentous fungi: kinetic characterization of the UapC purine transporter of *Aspergillus nidulans*. *Fungal Genet Biol.* 2014, **63**: 1-8. doi: 10.1016/j.fgb.2013.12.004.

3. Krypotou E, Lambrinidis G, Evangelidis T, Mikros E, Diallinas G. Modelling, substrate docking and mutational analysis identify residues essential for function and specificity of the major fungal purine transporter AzgA. *Mol Microbiol*. 2014, **93**: 129-45. doi: 10.1111/mmi.12646.

4. Krypotou E, Scazzocchio C, Diallinas G. Functional characterization of NAT/NCS2 proteins of *Aspergillus brasiliensis* reveals a genuine xanthine-uric acid transporter and an intrinsically misfolded polypeptide. *Fungal Genet Biol*. 2015, **75**: 56-63. doi: 10.1016/j.fgb.2015.01.009.

5. Krypotou E., Evangelidis T., Bobonis J., Pittis A.A., Gabaldón T., Scazzocchio C., Mikros E., Diallinas G. Origin, diversification and substrate specificity in the family of NCS1/FUR

transporters. *Mol Microbiol*. 2015, **96**: 927-50. doi: 10.1111/mmi.12982.

Conferences

Oral presentation:

Krypotou E., Lambrinides G., Evangelidis T., Mikros E., Diallinas G. Modeling, substrate docking and mutational analysis identify residues essential for function and specificity of a major fungal purine transporter.

64th Congress of the Hellenic Society of Biochemistry and Molecular Biology, Evgenides Foundation, Athens, Greece, December 6-8, 2013

Poster presentations:

1. Krypotou E., Kosti V., Myrianthopoulos V., Mikros E., Amillis S., Diallinas G. A combined computational and genetic approach for studying structure-function relationships in the cytosine-purine transporter FcyB

   62nd Congress of the Hellenic Society of Biochemistry and Molecular Biology, Athens, Greece, 2011

2. Krypotou E., Kosti V., Amillis S., Myrianthopoulos V., Mikros E., Diallinas G. Modeling, substrate docking and mutational analysis identify residues essential for the function and specificity of the purine-cytosine transporter FcyB.

   63rd Congress of the Hellenic Society of Biochemistry and Molecular Biology, Heraklion, Greece, November 9-11, 2012

3. Krypotou E., Lamprinidis G., Evangelidis T., Mikros E., Scazzocchio C., Diallinas G. Cryptic functions and evolution of substrate specificities in the Fur-like transporter family in *Aspergillus nidulans.*

   12th European Conference on Fungal Genetics (ECGF12) Seville, Spain. March 23-27, 2014

4. Evangelidis T., Krypotou E., Kairis M., Lougiakis N., Gavriil E.S., Lambrinidis G.,

Asimakopoulos A., Benaki D., Mikros E., Diallinas G., Marakos P., Pouli N. Design, Synthesis and *In Vivo* Evaluation of the Inhibitory Potential of Antifungal Agents Targeting the *Aspergillus nidulans* Purine–Cytosine/H+ FcyB Symporter.
EFMC-ISMC 2014, XXIII International Symposium on Medicinal Chemistry, Lisbon, Portugal - September 7-11, 2014

5. Krypotou E., Evangelidis T., Bobonis J., Pittis A.A., Gabaldón T., Scazzocchio C., Mikros E., Diallinas G. Origin, diversification and substrate binding specificity in the family of NCS1/FUR transporters.
4th International Workshop on Expression, Structure and Function of Membrane Proteins Florence, Italy - June 28 - July 2, 2015

6. Gavriil E.S., Krypotou E., Lambrinidis G., Lougiakis N., Sioupouli G., Asimakopoulos A., Benaki D., Mikros E., Diallinas G., Marakos P., Pouli N. Design, synthesis and biological evaluation of purine analogues as potential inhibitors substrates of the *Aspergillus nidulans* purine-cytosineH+ FcyB symporter.
FMC 2015, Frontiers in Medicinal Chemistry, Antwerp, Belgium - September 14-16, 2015

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 287, NO. 44, pp. 36792–36803, October 26, 2012

Modeling, Substrate Docking, and Mutational Analysis Identify Residues Essential for the Function and Specificity of a Eukaryotic Purine-Cytosine NCS1 Transporter*[S]

Received for publication, July 12, 2012, and in revised form, September 7, 2012 Published, JBC Papers in Press, September 11, 2012, DOI 10.1074/jbc.M112.400382

Emilia Krypotou[‡1], Vasiliki Kosti[‡1], Sotiris Amillis[‡], Vassilios Myrianthopoulos[§], Emmanuel Mikros[§], and George Diallinas[‡2]

From the [‡]Faculty of Biology, University of Athens, Panepistimiopolis, Athens 15784, Greece and [§]School of Pharmacy, University of Athens, Panepistimiopolis, Athens 15771, Greece

Background: The purine-cytosine FcyB transporter is a prototype member of the NCS1 family.
Results: Using homology modeling, substrate docking, and rational mutational analysis, we identify residues critical for function and specificity.
Conclusion: Important aspects concerning the molecular mechanism and evolution of transporter specificity are revealed.
Significance: The first systematic approach on structure-function-specificity relationships in a eukaryotic NCS1 member is shown.

The recent elucidation of crystal structures of a bacterial member of the NCS1 family, the Mhp1 benzyl-hydantoin permease from *Microbacterium liquefaciens*, allowed us to construct and validate a three-dimensional model of the *Aspergillus nidulans* purine-cytosine/H^+ FcyB symporter. The model consists of 12 transmembrane α-helical, segments (TMSs) and cytoplasmic N- and C-tails. A distinct core of 10 TMSs is made of two intertwined inverted repeats (TMS1–5 and TMS6–10) that are followed by two additional TMSs. TMS1, TMS3, TMS6, and TMS8 form an open cavity that is predicted to host the substrate binding site. Based on primary sequence alignment, three-dimensional topology, and substrate docking, we identified five residues as potentially essential for substrate binding in FcyB; Ser-85 (TMS1), Trp-159, Asn-163 (TMS3), Trp-259 (TMS6), and Asn-354 (TMS8). To validate the role of these and other putatively critical residues, we performed a systematic functional analysis of relevant mutants. We show that the proposed substrate binding residues, plus Asn-350, Asn-351, and Pro-353 are irreplaceable for FcyB function. Among these residues, Ser-85, Asn-163, Asn-350, Asn-351, and Asn-354 are critical for determining the substrate binding affinity and/or the specificity of FcyB. Our results suggest that Ser-85, Asn-163, and Asn-354 directly interact with substrates, Trp-159 and Trp-259 stabilize binding through π-π stacking interactions, and Pro-353 affects the local architecture of substrate binding site, whereas Asn-350 and Asn-351 probably affect substrate binding indirectly. Our work is the first systematic approach to address structure-function-specificity relationships in a eukaryotic member of NCS1 family by combining genetic and computational approaches.

Purines and pyrimidines (nucleobases) are absolutely essential metabolites for all cells, being not only the precursors of nucleotide and nucleic acid biosynthesis but also being involved in cell signaling, nutrition, response to stress, and cell homeostasis. Most cells in all domains of life possess specific nucleobase transporters, distinct from those involved in the uptake of nucleosides (1, 2). Nucleobase transporters are important in protozoan and plant development and in the susceptibility of human, protozoan, or fungal cells to purine-related drugs (3–10). Fungi possess three evolutionary distinct families of transporters specific for nucleobases (2, 11, 12). Members of one of them, the so called nucleobase cation symport (NCS1,[3] also known as purine-related transporter) family, are ubiquitously distributed in all fungi and are present in some eubacteria, archaea (11–14), and plants (15) but are absent in protozoa and metazoa (see uniprot online). Besides purines, cytosine, and uracil, other known substrates for NCS1 transporters include the antifungals 5-fluorocytosine (5-FC), and 5-fluorouracil, benzyl-hydantoin, allantoin, thiamine, pyridoxal-based compounds, and nicotinamide riboside. All NCS1 members of known function probably act as H^+ symporters, although the *Microbacterium liquefaciens* benzyl-hydantoin transporter was reported to be a Na^+ symporter (11).

The NCS1 family includes two major subfamilies, the Fcy-like and the Fur-like transporters (11). Three Fcy-like proteins of Ascomycetes have been well characterized genetically and studied with respect to regulation of expression, transport kinetics, and substrate specificity. These are the Fcy2p (16), Fcy21p (17), and FcyB (13) permeases of *Saccharomyces cerevi-*

* This work was supported in part by the John S. Latsis Public Benefit Foundation.
[S] This article contains supplemental Tables S1 and S2 and Figs. 1–3.
[1] Co-financed by the European Union (European Social Fund) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework-Research Funding Program: THALIS and Heracleitus II, Investing in knowledge society through the European Social Fund.
[2] To whom correspondence should be addressed. Tel.: 30(210)7274649; Fax: 30(210)7274702; E-mail: diallina@biol.uoa.gr.

[3] The abbreviations used are: NCS1, nucleobase cation symport 1; TMS, transmembrane segment; 5-FC, 5-fluorocytosine; IFD, induced fit docking; MM, minimal media; ER, endoplasmic reticulum; mRFP, monomeric red fluorescent protein.

siae, *Candida albicans*, and *Aspergillus nidulans*, respectively. All three are involved in high affinity transport of adenine, hypoxanthine, guanine, cytosine, and 5-FC. The only physiological difference between Fcy2p/Fcy21p and FcyB is that the two yeast transporters constitute the sole high capacity gateways for purine uptake, whereas in *A. nidulans* high capacity purine uptake is catalyzed by another transporter, called AzgA (17), FcyB acting basically as a cytosine supplier and only secondarily as a purine carrier (13).

Despite some classical genetic approaches in Fcy2p that identified several residues critical for substrate or cation binding and transport (18–20), very little was known with respect to structure-function relationships in NCS1-like transporters until recently. In 2008, however, the crystal structure of a bacterial member of the NCS1 family, namely the Mhp1 benzyl-hydantoin permease from *M. liquefaciens* (21), was reported (22). Surprisingly, the Mhp1 topology has proved to be very similar to that of several recently revealed structures of other bacterial transporters that showed no sequence similarity and exhibited entirely different specificities. These include the amino acid transporter LeuT (23), the galactose transporter vSGLT (24), the betaine transporter BetP (25), and two amino acid transporters, AdiC (26) and ApcT (27). Dysfunction of members of this growing superfamily in humans is associated with neurological (28) and kidney disorders (29), cancer (30), and drug resistance (31). The core of the fold shared by these transporters is an "inverted repeat" motif with two sets of five transmembrane helices oppositely orientated with respect to the membrane (often called the 5HIR-fold) (32–34). The two extra TMSs (TMS11 and TMS12 in the NCS1 family) do not seem to participate in transport activity, and their role is unclear.

Most interestingly, several different conformations have been observed for these transporters in the recent years. These can be categorized into three classes: outward-facing, as observed in LeuT (35, 36), Mhp1 (22), and AdiC (26, 37); occluded, where a trapped substrate is blocked from exiting on either side of the protein as seen in LeuT (23), Mhp1 (22), BetP (25), and AdiC (37); inward-facing, as seen in Mhp1 (34), vSGLT (24), ApcT (27), and LeuT (36). Up to date, Mhp1 and LeuT are the only transporters trapped into three structural conformations associated with transport catalysis. From analyses of these three structures and molecular dynamics simulations, a mechanism has been proposed for the transport cycle in Mhp1 or LeuT (34, 36, 38). Switching from the outward- to the inward-facing state goes through occluded states and is primarily achieved by a rigid body movement of several TMSs (*e.g.* 3, 4, 8, and 9 in Mhp1) relative to a rather rigid bundle of helices (1, 2, 6, and 7 in Mhp1). In the occluded transient states, "thin gates" involving only a few residues in specific TMSs (*e.g.* 5 and 10 in Mhp1) and parts of loops control the opening and closing of the substrate binding site to the exterior or interior. This forms the basis of an alternating access mechanism applicable to probably all transporters of the 5HIR superfamily (32–34, 36, 39). Evidently, details of substrate transport, especially those concerning the occlusion mechanism, remain largely controversial because models have been derived by comparing transporters with divergent amino acid sequences that transport a wide variety of substrates and that possibly select and transport different ions (Na^+ *versus* H^+) for driving the symport of their major substrates (34, 36, 38). Furthermore, what is often ignored is that all available structures concern prokaryotic transporters, which despite being similar to important mammalian carriers, also present significant structural and biochemical differences. For example, eukaryotic transporters possess long N- and C-tails or longer hydrophilic loops, absent in prokaryotes, that are often associated with transporter subcellular trafficking, function, specificity, stability, or topology (40–42).

In this work we construct and validate a three-dimensional model of the *A. nidulans* purine-cytosine FcyB transporter based on the outward facing substrate-occluded crystal structure of the Mhp1 benzyl-hydantoin permease from *M. liquefaciens* (PDB entry 2JLO). Subsequently, we use this model to perform independent docking studies and thus identify residues that might be involved in substrate binding and transport. Most importantly, we experimentally validate our speculations on specific amino acid residues by a systematic functional analysis of relevant mutants. Our work is discussed with respect to the molecular determinants underlying substrate specificity in the NCS1 family.

EXPERIMENTAL PROCEDURES

Homology Modeling—Proteins with significantly similar structural-fold with FcyB were identified using HHpred. All 5HIR transporters showed significant similarity with FcyB, but only Mhp1 showed a 100% probability (*E* value $1.8E^{-55}$; *p* value $6.6\ E^{-60}$). Homology model building was performed using MODELLER v.9.8 software (43).

Protein Preparation—The protein was prepared for docking calculations using the Protein Preparation Workflow (Schrödinger Suite 2011 Protein Preparation Wizard) implemented in the Schrödinger suite and accessible from within the Maestro program (Maestro, Version 9.2, Schrödinger, LLC, NY, 2011). Briefly, hydrogen atoms were added, and the orientation of hydroxyl groups, Asn, Gln, and the protonation state of His were optimized to maximize hydrogen bonding. Finally, the ligand-protein complex was refined with a restrained minimization performed by Impref utility that is based on the Impact molecular mechanics engine (Impact Version 5.7, Schrödinger, LLC, NY, 2011) and the OPLS2001 force field, setting a max root mean square deviation of 0.30 (44). Ligand preparation for docking was performed with LigPrep (LigPrep, Version 2.5, Schrödinger, LLC, NY, 2011) application that consists of a series of steps that perform conversions, apply corrections to the structure, generate ionization states and tautomers, and optimize the geometries.

Molecular Dynamic Simulations—For the molecular dynamics simulations, Desmond v.3 software was implemented (Desmond Molecular Dynamics System, Version 3.0, D. E. Shaw Research, NY). The system was prepared by embedding the protein in a 1-palmitoyl-2-oleylphosphatidylcholine (POPC) lipid bilayer, solvating the membrane by TIP4P explicit water, neutralizing with counter ions, and adding 150 mM salt and subsequently following the stepwise equilibration protocol as developed by Desmond for membrane proteins (45). The ori-

entation of the protein in the bilayer was determined by using the Mhp1 template available in OPM database (46). An 18-ns simulation was performed in the NPγT ensemble with Langevin thermostat and barostat and semi-isotropic pressure restraints (supplemental Fig. S1).

Induced Fit Docking—Molecular docking was performed using the induced fit docking (IFD) protocol (Schrödinger Suite 2011 Induced Fit Docking protocol) that is intended to circumvent the inflexible binding site and accounts for the side chain or backbone movements or both upon ligand binding (47). In the first stage of the IFD protocol, softened-potential docking step, 20 poses per ligand were retained. In the second step for each docking pose, a full cycle of protein refinement was performed with Prime 1.6 (Prime, Version 3.0, Schrödinger, LLC, NY, 2011) on all residues having at least one atom within 8.0 Å of an atom in any of the 20 ligand poses. The Prime refinement starts with a conformational search and minimization of the side chains of the selected residues, and after convergence to a low energy solution, an additional minimization of all selected residues (side chain and backbone) is performed with the truncated-Newton algorithm using the OPLS parameter set and a surface Generalized Born implicit solvent model. The obtained complexes are ranked according to Prime calculated energy (molecular mechanics and solvation), and those within 30 kcal/mol of the minimum energy structure are used in the last step of the process, redocking with Glide 5.7 (Glide, Version 5.7, Schrödinger, LLC, NY, 2011) using standard precision and scoring. In the final round, the ligands used in the first docking step are redocked into each of the receptor structures retained from the refinement step. The final ranking of the complexes is done by a composite score that accounts for the receptor-ligand interaction energy (GlideScore) and receptor strain and solvation energies (Prime energy).

Media, Strains, Growth Conditions, and Transformation Genetics—Standard complete and minimal media (MM) for *A. nidulans* were used. Media and supplemented auxotrophies were at the concentrations given online in fgsc. Nitrogen sources were used at the final concentrations, 5 mM urea, 10 mM $NaNO_3$, 0.5 mM purines. *Escherichia coli* was grown on Luria-Bertani medium. Media and chemical reagents were obtained from Sigma or AppliChem (Bioline Scientific SA, Hellas, Greece). Transformations were performed as described previously (48). A single-copy plasmid integration of the FcyB served as a standard wild-type control. A Δ*fcyB*::*argB*; Δ*uapA*; Δ*uapC*::*AFpyrG*; Δ*azgA*; *riboB2*; *pabaA1* mutant strain (13) was the recipient strain in transformations with the wild-type pAN520, the pAN530, the pAN540 (see below), or the mutant FcyB alleles carried on the above vectors. As a negative, transformants with empty pAN vectors were used. These vectors allow selection of transformants based on *p*-aminobenzoic acid (*pabaA1*) or riboflavin (*riboB2*) auxotrophy complementation. Transformants expressing intact *fcyB* or *fcyB-gfp* alleles, through either single-copy or multi-copy plasmid integration events, were identified by PCR and Southern analysis. For the relative subcellular localization of wild-type or mutant N81A FcyB-GFP with mRFP-tagged histone H1 (HhoA-mRFP), appropriate strains were constructed by standard genetic crossing. The strain harboring HhoA-mRFP is reported in Edgerton-Morgan and Oakley (49). Growth tests were performed at 25 and at 37 °C, pH 6.8.

Plasmid Constructions and FcyB Mutations—pAN520 is a modified pBluescript KS(+) vector (Stratagene) based on the plasmid pAN510 (50) that carries the FcyB orf cloned in-frame between the 5′ and 3′ regulatory sequences of the *uapA* gene and also carrying the *pabaA* gene as a selection marker (13). pAN530 and pAN540 are similar to pAN520, but in this case the *fcyB* and the *fcyB* fused C- or N-terminally with the *gfp* together with the 5′ and 3′ regulatory sequences of the *uapA* gene were cloned in the pGEM™-T vector (Promega), also carrying the *riboB* gene from *Aspergillus fumigatus* as a selection marker. Mutations were constructed by site-directed mutagenesis according to the instructions accompanying the QuikChange® site-directed mutagenesis kit (Stratagene) on the above vectors and were confirmed by sequencing. Oligonucleotides used for cloning and site-directed mutagenesis purposes are listed in supplemental Table S1.

Standard Nucleic Acid Manipulations—Genomic DNA extraction from *A. nidulans* was as described in fgsc. Plasmid preparation from *E. coli* strains and DNA bands were purified from agarose gels done with the Nucleospin Plasmid kit and the Nucleospin ExtractII kit according to the manufacturer's instructions (Macherey-Nagel, Lab Supplies Scientific SA, Hellas, Greece). Southern blot analysis was performed as described previously (51). [^{32}P]dCTP-labeled molecules used as *fcyB*-, *riboB*-, or *pabaA*-specific probes were prepared using a random hexanucleotide primer kit following the supplier's instructions (Takara Bio, Lab Supplies Scientific SA) and purified on MicroSpin™ S-200 HR columns following the supplier's instructions (Roche Diagnostics). Labeled [^{32}P]dCTP (3000 Ci/mmol) was purchased from the Institute of Isotopes Co. Ltd, Miklós, Hungary. Restriction enzymes were from Takara Bio. Conventional PCR reactions were done with KAPATaq DNA polymerase (Kapa Biosystems, Lab Supplies Scientific SA). Cloning and amplification of products were done with Pfx Platinum (Invitrogen) or Phusion® Flash High-fidelity PCR MasterMix (New England Biolabs, Lab Supplies Scientific SA).

Membrane Protein Extraction and Western Blot Analysis—Cultures for membrane protein extraction were grown in MM supplemented with urea at 37 °C for 8 h. Membrane protein extraction was performed as previously described (52). Equal sample loading was estimated by Bradford assays. Total proteins (30 μg) were separated by SDS-PAGE (10% w/v polyacrylamide gel) and electroblotted (Mini PROTEAN™ Tetra Cell, Bio-Rad) onto PVDF membranes (Macherey-Nagel, Lab Supplies Scientific SA) for immunodetection. The membrane was treated with 3% (w/v) BSA, and immunodetection was performed with a primary mouse anti-GFP monoclonal antibody (Roche Diagnostics) or a mouse anti-actin monoclonal (C4) antibody (MP Biomedicals Europe, Lab Supplies Scientific SA) and a secondary goat antimouse IgG HRP-linked antibody (Cell Signaling Technology Inc., Bioline Scientific SA). Blots were developed by the chemiluminescent method using the LumiSensor Chemiluminescent HRP Substrate kit (Genscript USA, Lab Supplies Scientific SA) and SuperRX Fuji medical X-Ray films (FujiFILM Europe, Lab Supplies Scientific SA).

Kinetic Analysis—[^{3}H]Hypoxanthine (19.6–33.4 Ci/mmol, Moravek Biochemicals, CA) uptake in MM was assayed in germinating conidiophores of *A. nidulans* concentrated at 10^7 conidiospores/100 μl at 37 °C, pH 6.8, as previously described (40, 53). Initial velocities were measured at 1 min of incubation with concentrations of 0.2–2.0 μM [^{3}H]hypoxanthine at the polarity maintenance stage (3–4 h, 130 rpm). $K_{m/i}$ values were obtained directly by performing and analyzing uptakes (Prism 3.02: GraphPad Software) using labeled hypoxanthine at 0.2–0.5 μM or at various concentrations (0.5–2000 μM) of non-labeled substrates. K_i values were calculated by satisfying the criteria for use of the Cheng and Prusoff equation $K_i = IC_{50}/(1 + (L/K_m))$ in which L is the permeant concentration. IC_{50} values were determined from full dose-response curves, and in all cases the Hill coefficient was close to −1, consistent with the presence of one binding site. Reactions were terminated with the addition of equal volumes of ice-cold MM containing 1000-fold excess of non-radiolabeled substrate. Background uptake values were corrected by subtracting either values measured in the deleted mutants or values obtained in the simultaneous presence of 1000-fold excess of non-radiolabeled substrate. Both approaches led to the same background uptake level, not exceeding 10–15% of the total counts obtained in wild-type strains. All transport assays were carried out in at least three independent experiments, with three replicates for each concentration or time point. S.D. was < 20%.

Epifluorescence Microscopy—Samples for fluorescence microscopy were prepared as previously described (52). In brief, germlings incubated on coverslips in liquid MM supplemented with $NaNO_3$ as the nitrogen source for 12–14 h at 25 °C were observed on an Axioplan Zeiss phase-contrast epifluorescent microscope, and the resulting images were acquired with a Zeiss-MRC5 digital camera using the AxioVs40 V4.40.0 software. Image processing, contrast adjustment, and color combining were made using the Adobe Photoshop CS4 Extended Version 11.0.2 software or the ImageJ software. Images were converted to 8-bit grayscale or RGB and annotated using Photoshop CS4 before being saved to TIFF.

RESULTS

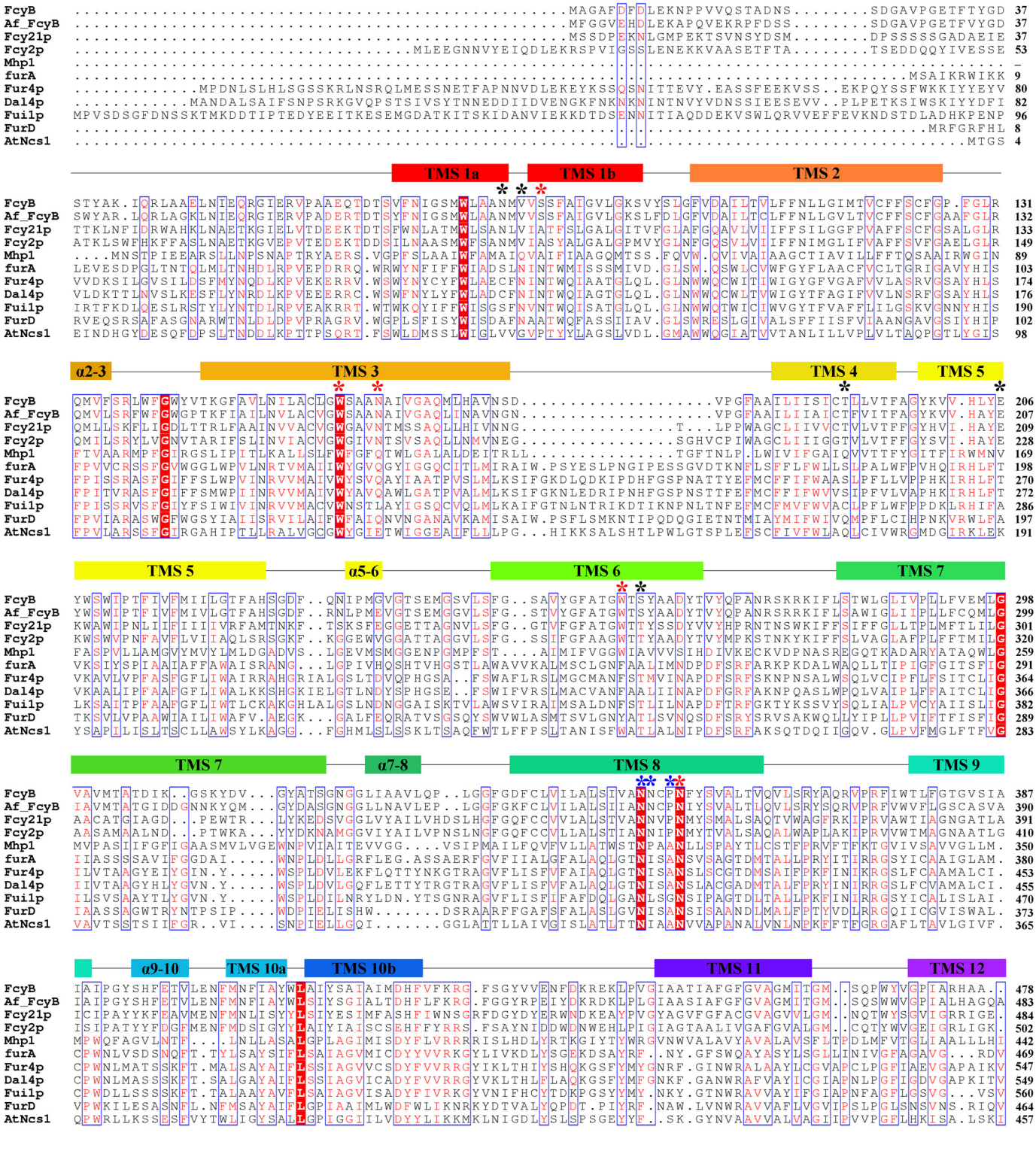
FcyB Structural Model—The construction of a structural model of FcyB was based on the crystal structure of the Mhp1 benzyl-hydantoin permease from *M. liquefaciens* (PDB entry 2JLO (22)). The two proteins share a rather moderate sequence similarity (18% identity), which is, however, adequate for sustaining a theoretical model of FcyB. The model built here was based on the alignment of the two proteins, as shown in a multiple alignment including all NSC1 carriers with known function, which was further modified manually (Fig. 1). Model building was performed using MODELLER software. The loop refinement routine and a slow simulated annealing protocol for model refinement were implemented. As a first validation of the model, the structure with the best spatial restraints score was subjected to a 18-ns molecular dynamics run using Desmond software. The system was prepared by embedding the protein in a POPC lipid bilayer and solvating the membrane by explicit water. The root mean square deviation of the transmembrane helices α-carbons from starting coordinates was monitored throughout the simulation and did not exceed 3 Å, thus indicating the stability of the theoretical model.

The overall three-dimensional structure of the FcyB model (Fig. 2) corresponds to an outward-facing conformer made of 12 TMSs that adopt a mostly helical secondary structure. The architecture of the transporter divides it in two distinct domains, a compact core consisting of segments TMS1–10 and a C-terminal domain made of TMS11–12. The transmembrane helices in the core are connected with rather short loops (<14 amino acid residues), whereas the loop separating the core and TMS11–12 is longer (23 residues). The distribution of the ionized residues on the protein surface is reasonable, as most of them are positioned either at the cytoplasmic and periplasmic sides or along the protein pore in the protein interior. Positive charges are mostly concentrated in the cytoplasm-facing loops. The core of FcyB is subdivided in two inverted repeats made of TMS1–5 and TMS6–10, arranged in a 2-fold pseudosymmetry axis, running parallel with the membrane plane. The two repeat units are completely intertwined, giving two topologically distinct subdomains made of TMSs 1, 2, 6, and 7 and TMSs 3, 4, 8, and 9, respectively, linked with helices TMS5 and TMS10. As will be shown below, the substrate binding site is located in the space between the two subdomains of the core (see Fig. 2*B*).

Substrate Docking in FcyB—Using the FcyB model, aspects of substrate recognition were addressed using docking calculations. The four major physiological substrates, namely hypoxanthine, adenine, guanine, and cytosine, as well as the antifungal 5-FC were docked to the modeled structure of FcyB. Calculations were performed using the IFD protocol as introduced by the Schrödinger 2011 Suite of programs. The IFD protocol is based on an iterative implementation of Glide algorithm for rigid docking and Prime algorithm for protein refinement, resulting in an improved simulation of binding in terms of protein flexibility. Furthermore, because Prime is a modeling tool especially developed for refinement of protein structures derived by homology, its implementation as part of the IFD protocol was considered in the case of FcyB as promising. The best orientation for each substrate was finally selected on the basis of lowest energy, number of intermolecular interactions, and visual inspection.

Fig. 3 shows the modeled interactions of FcyB with its four physiological substrates and the antifungal 5-FC. FcyB interacts with adenine through a bidentate H bond that is formed between Asn-163 and ligand sites C6-NH_2 and N1, an H bond between Ser-85 and purine N7, and π-π stacking of the purine ring between the indole rings of Trp-159 and Trp-259 (Fig. 3*A*). FcyB interacts with hypoxanthine and guanine through the establishment of three H bonds between Asn-163 and N1-H, between Ser-85 and C2=O_2, and between Asn-354 and N3, as well as π-π stacking of the purine ring between Trp-159 and Trp-259 (Fig. 3, *B* and *C*). Finally, FcyB interacts with cytosine and its toxic analog 5-FC by forming three H bonds between Asn-163 and N3, between Asn-354 and ligand C4-NH_2, and between Ser-85 and C2=O_2 as well as π-π stacking with Trp-159 and Trp-259 (Fig. 3, *D* and *E*).

These results establish that Asn-163 and the two invariable Trp residues, Trp-159 and Trp-259, interact with similar positions of the rings of all purines and cytosine, whereas the inter-



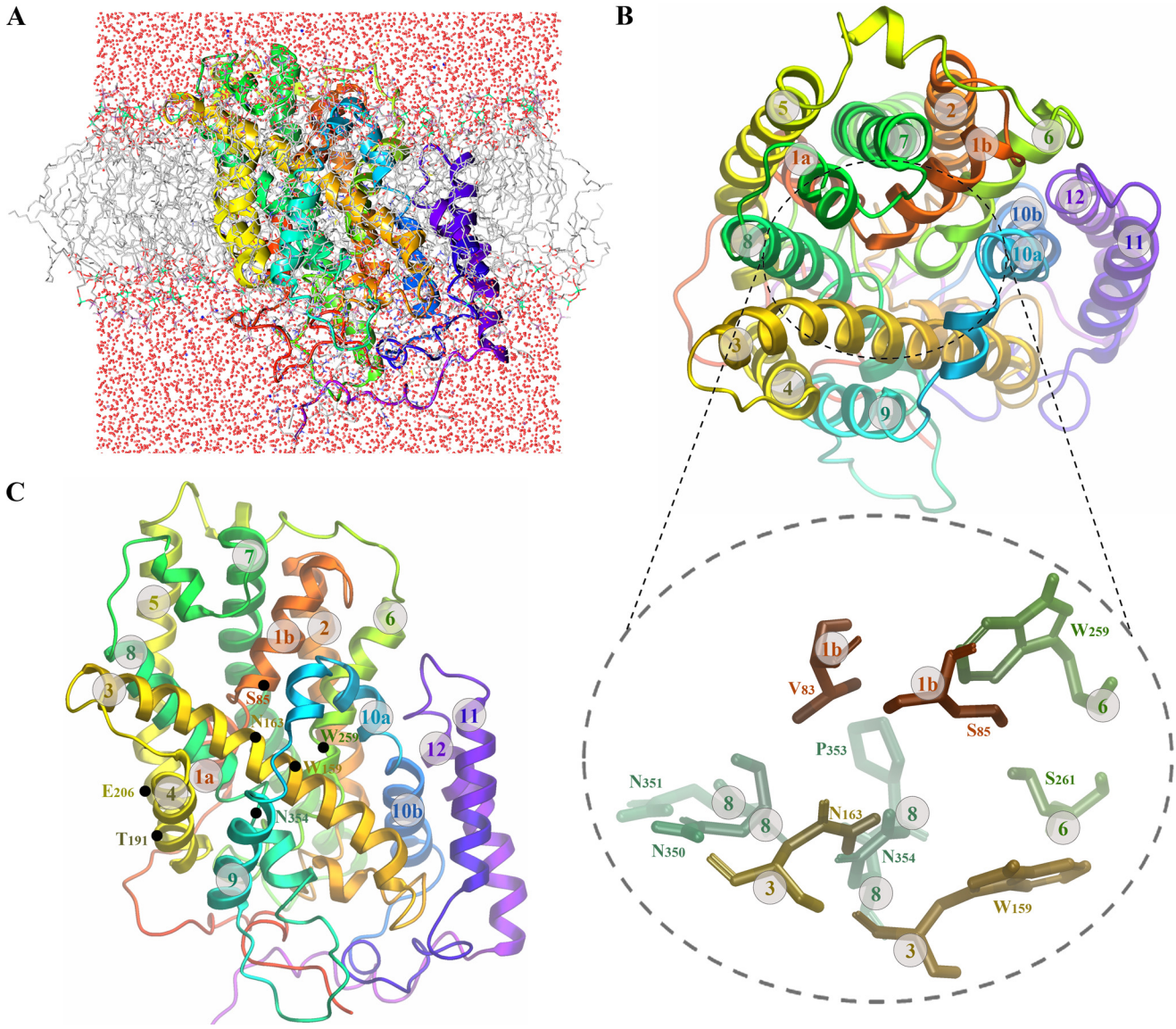


FIGURE 2. **FcyB structural model.** *A*, shown is the modeled three-dimensional structure of FcyB validated with molecular dynamics using Desmond software. *B*, shown is the top view of the FcyB model where the core of the first 10 TMSs is clearly distinguished from the last two TMSs 11 and 12, which probably do not affect transport catalysis *per se*. The core can also be seen as two subdomains, the first made of TMSs 1, 2, 6, and 7 and the second from TMSs 3, 4, 8, and 9, linked with flexible helices TMS5 and TMS10. The substrate binding site is located in the space between the two subdomains of the core. In the *lower panel*, the topology of residues critical for the function of the substrate binding site, Ser-85, Trp-159, Asn-163, Trp-259, Asn-350, Asn-351, Pro-353, and Asn-354, is shown in zoom-out. *C*, shown is the side view of the FcyB structure showing the topology of residues Ser-85, Trp-159, Asn-163, Trp-259, and Asn-354, involved in direct interactions with substrates (see also Fig. 3). Residues Thr-191 and Glu-206 are also shown (see Results). TMS regions are highlighted with colors as in Fig. 1.

actions involving Ser-85 and Asn-354 vary depending on the substrate docked. As will be shown and discussed later, mutational analysis strongly support the docking results. In addition, the model and docking analysis showed that two neighboring Asn residues, Asn-350 and Asn-351, might interact with Asn-354 in a way that probably builds a necessary architecture for substrate binding (see also supplemental Fig. S2). The importance of residues Asn-350 and Asn-351 in substrate binding was rigorously established through mutational analysis in this work (see later). Finally, in the docking models for purine binding, Ser-261 was also within H bond distance from either N3 (adenine) or N9-H (hypoxanthine, guanine). However, in this case, mutational analysis did not support the involvement of Ser-261 in direct substrate binding (see also later).

FIGURE 1. **Multiple sequence alignment of FcyB and NCS1 homologues of known function.** The three-dimensional FcyB model was constructed on the basis of the alignment shown with Mhp1. Putative TMSs of FcyB are denoted in *colored cylinders*. Invariant and highly conserved amino acids are shaded in *red* and *blue-lined boxes*, respectively. Amino acids critical for function and specificity discussed in the Results are highlighted with *asterisks*: *red* for residues interacting with substrates, *blue* for those critical for substrate binding and transport, and *black* for other residues discussed in the Results. The listed NCS1 homologues include: FcyB of *A. nidulans*, GI: 169798762; AfFcyB of *A. fumigatus* (Af_FcyB), GI: 169798764; Fcy21p of *S. cerevisiae*, GI: 392860008; Fcy2p of *S. cerevisiae*, GI: 6320897; Mhp1 of *M. liquefaciens*, GI: 210060746; FurA of *A. nidulans*, GI: 2876438; Fur4p of *S. cerevisiae*, GI: 6319495; Dal4p of *S. cerevisiae*, GI: 392860728; Fui1p of *S. cerevisiae*, GI: 6319429; FurD of *A. nidulans*, GI: 149212441; AtNcs1 of *Arabidopsis thaliana*; GI: 16648836. All Fcy-like proteins are purine-cytosine transporters. Fur4p and FurD are uracil transporters. FurA and Dal4p are allantoin transporters. Fui1p is a uridine transporter. AtNCS1 is reported to be specific for adenine-guanine-uracil transport, but no direct information on its ability to transport other nucleobases exists (15).

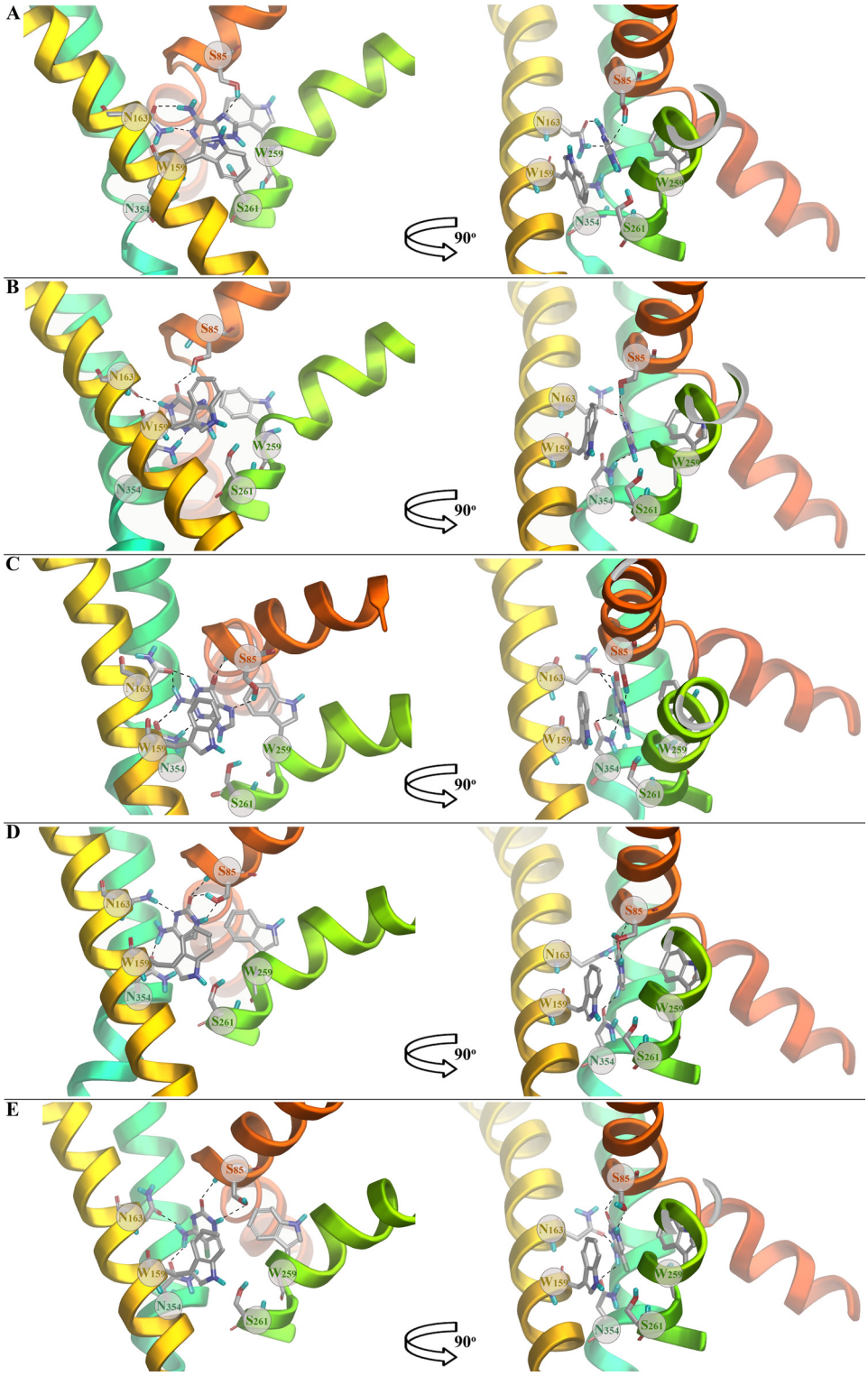


FIGURE 3. **Substrate docking in FcyB.** *A*, adenine. *B*, hypoxanthine. *C*, guanine. *D*, cytosine. *E*, 5-FC. Hydrogen bonds are depicted with *dashed lines*.

Rational Design and Construction of FcyB Mutants—Results derived from docking methodologies suggested that five residues in FcyB might be directly involved in substrate binding. These are Ser-85 in TMS1, Trp-159 and Asn-163 in TMS3, Trp-259 in TMS6, and Asn-354 in TMS8. Two more Asn residues, Asn-350 and Asn-351, might also interact with Asn-354 in a way that is critical for substrate binding. Five of these putatively functional residues, Trp-159, Asn-163, Trp-259, Asn-350, and Asn-354, superimpose with amino acid residues, which have been shown to directly interact with the substrate (benzyl-hydantoin) in Mhp1, namely Trp-117, Gln-121, Trp-220, Asn-314, and Asn-318. Interestingly, these residues are either absolutely (Trp-117, Asn-314, and Asn-318) or highly (Trp-220 and Gln-121) conserved in the NCS1 family (see Fig.

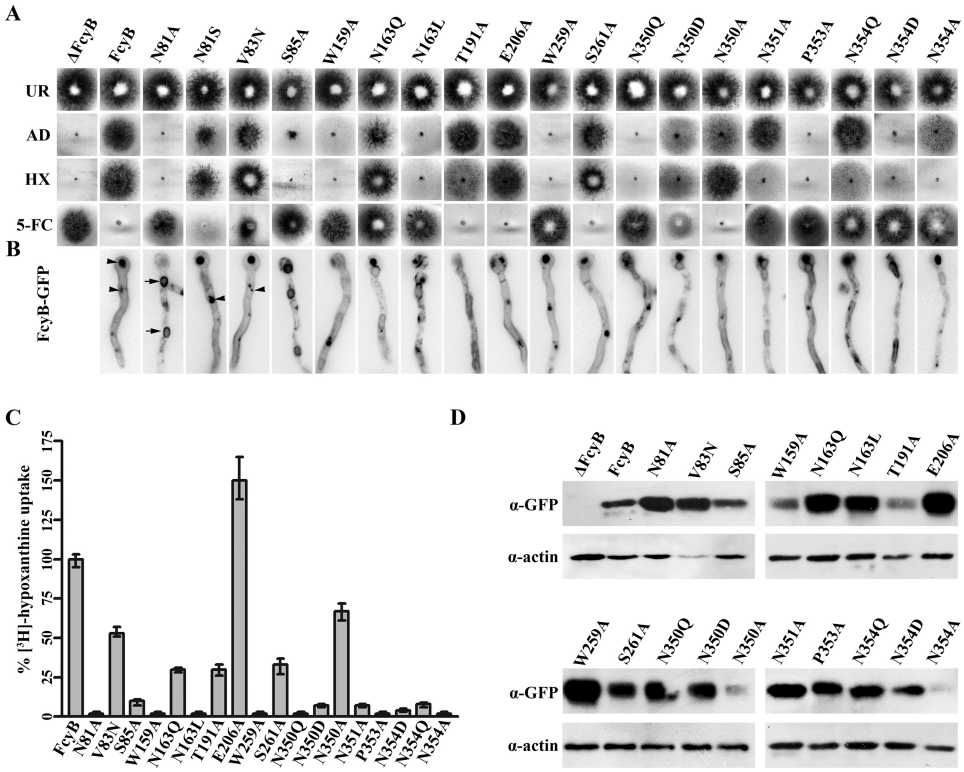


FIGURE 4. **Functional analysis of FcyB mutations.** *A*, shown are growth tests on purines as sole nitrogen sources and resistance/sensitivity test on 5-FC at 25 °C. *AD* is adenine, and *HX* is hypoxanthine. Growth on urea is also shown (*UR*) as a control. Positive (FcyB) and negative (ΔFcyB) isogenic control strains are shown. *B*, epifluorescence microscopy showing *in vivo* subcellular expression of FcyB-GFP mutant alleles and a wild-type control (FcyB) is presented as *dark structures in a grayscale inverted mode*. In selected samples, *arrows* and *arrowheads* depict perinuclear ER membrane rings (verified by DAPI staining; not shown) and vacuoles, respectively. *C*, comparative initial uptake rates of ^{3}H-radiolabeled hypoxanthine in FcyB mutant alleles and a wt control are shown. 100% is the transport rate in the wt (FcyB). *D*, Western blot analysis of total proteins from FcyB-GFP mutants detected with anti-GFP antibody is shown. Antibody against actin was used as an internal marker for equal loading.

1). Given the different substrate specificities of NCS1 members, this suggests that some other residues might also be critical for substrate binding, a prediction supported in this work (see later).

To test the functional role of the residues discussed above, as well as residues differentially conserved in NCS1 members with different specificity profiles or residues positioned close to the proposed binding site, we constructed a series of relevant mutations and analyzed functionally the corresponding mutants. The 19 mutations made were: N81A, N81S, S85A, V83N, W159A, N163Q, N163L, T191A, E206A, W259A, S261A, N350Q, N350D, N350A, N351A, P353A, N354Q, N354D, N354A (See supplemental Table S2). The rationale for the specific changes constructed was either to introduce Ala or highly conserved residues or replace FcyB residues with those found in other NCS1 members with different specificity.

All mutations were made in vectors carrying the FcyB orf expressed from the *uapA* promoter (see "Experimental Procedures"). This was so because the endogenous *fcyB* promoter is very weakly expressed under all conditions tested (13). The *uapA* gene encodes a very well studied uric acid-xanthine transporter (for reviews, see Refs. 2 and 12), and thus we used its promoter, which drives low but significant transcription. Furthermore, the *uapA* promoter can be totally repressed in the presence of ammonium ions or glutamine and can be induced in the presence of uric acid or the gratuitous induced 2-thioxanthine (54). In all constructs *fcyB* transcription is terminated by the 3′-flanking region of *uapA*. In one set of constructs, the FcyB orf was fused in-frame with the *gfp* orf. All plasmids were introduced in a strain lacking all endogenous purine transporters ("Experimental Procedures").

Phenotypic Analysis of FcyB Mutants—We analyzed all selected transformants on different purines as sole nitrogen sources together with positive and negative isogenic control strains. A summary of this analysis of transformants expressing single-copy FcyB plasmids is shown in Fig. 4*A*. The positive control strain expressing wild-type FcyB from the *uapA* promoter can grow on adenine and hypoxanthine as sole nitrogen sources and is sensitive to 5-FC. The negative control strain lacking FcyB cannot grow in any purine and is resistant to 5-FC. Based on our previous experience with *A. nidulans* transporters, any mutation reducing steady state rates of purine uptake to <50% that of the wild-type rate is clearly reflected in relevant growth tests. Transport reductions to levels >50% do not lead to distinguishable growth phenotypes. As expected, all strains showed normal growth on nitrogen sources other than purines (*e.g.* urea). None among the FcyB alleles conferred a capacity for growth on other purines, such as uric acid or xanthine, which are not physiological substrates for the wild-type FcyB (not shown). Within the limits of growth testing on hypoxanthine or adenine and sensitivity to 5-FC, FcyB mutants could be classified into apparent wild-type-like (E206A), total loss-of function (N81A, W159A, N163L, W259A, N350Q, P353A, N354D), or partial loss-of function (V83N, S261A) or altered specificity

showing differential capacities to grow on adenine, hypoxanthine, or 5-FC (T191A, N163Q, N350A, N350D, N351A, N354A, N354Q). In particular, mutants T191A, N351A, N354A, and N354Q grow better on adenine than on hypoxanthine, but in addition, T191A has wild-type like 5-FC sensitivity, whereas the other three mutants are significantly resistant to this toxic analog. In contrast, N163Q and N350A grow better on hypoxanthine than on adenine, but in addition, N163Q is resistant to 5-FC, whereas N350A has wild-type-like 5-FC sensitivity. N350D does not grow on hypoxanthine and adenine but is 5-FC-sensitive. The conclusions from this growth test are: (i) Asn-351 and Asn-354 are very important for hypoxanthine and 5-FC (and probably cytosine) transport but less for adenine, (ii) Asn-163 is more important for 5-FC but less critical for adenine transport and even less for hypoxanthine transport, (iii) Asn-350 is more critical for the transport of adenine and less for hypoxanthine 5-FC, and (iv) T191 seems critical solely for hypoxanthine transport.

Apparent Purine Transport Rates in FcyB Mutants—We performed comparative radiolabeled hypoxanthine uptake measurements in the strains shown in Fig. 4*A*. Our results (Fig. 4*C*) were in line with growth tests. All mutants characterized as loss-of-function had no or extremely low ($V < 6\%$) apparent uptake capacity for hypoxanthine. Partial loss-of-function mutations had 33–53% apparent uptake capacity, and E206A showed 150% transport capacity compared with wild type. Specificity mutants had low rates for hypoxanthine transport (2–15% that of the wild-type) except N350A (67%). These results showed that most amino acid residues tested, except Glu-206 and to a certain degree Val-83 and Ser-261, are critical for FcyB transport activity and substrate specificity.

Expression Levels and Cytoplasmic Localization of Mutant Versions of FcyB—Reduced apparent *V* values could imply FcyB problematic trafficking, high turnover, or *bona fide* reduction in transport activity. To distinguish among these possibilities, we reconstructed relevant FcyB alleles tagged with GFP and performed Western blots and epifluorescence microscopic analyses.

Western blot analysis of total proteins isolated by the mutants probed with anti-GFP showed that the FcyB protein steady state levels are either similar or increased (up to 6-fold) compared with the wild type (Fig. 4*D*). Increased protein steady state levels are in principle due to either increased intrinsic stability, often associated with reduced transport function, or due to reduced vacuolar turnover due to ER retention.

Epifluorescence microscopic analysis of the same mutants confirmed that the FcyB-GFP protein was expressed in all mutants and further revealed that in several mutants (N81A, S85A, N163Q, N163L, T191A, N350D, N351A, N354Q, N354D, N354A) FcyB-GFP had increased ER retention relevant to the wild-type FcyB-GFP protein (Fig. 4*B*). In fungi, the ER membrane appears mostly as very characteristic perinuclear rings, as verified by Hoechst 33242 or DAPI staining, or by fluorescent labeling of an ER resident chaperone (55–57). We have previously shown that transporters with problematic trafficking are mostly retained in these ER perinuclear rings (11, 40). We also verified the identity of the FcyB-GFP-labeled ER perinuclear rings in a relative mutant (N81A), by nuclei staining, and by examining the relative localization of these rings with histone H1 labeled with mRFP (supplemental Fig. S3). Noticeably, in some mutants (*e.g.* N81A) the FcyB-GFP signal intensity does not correspond to the protein steady state levels detected by anti-GFP (see Fig. 4*D*). This discrepancy between higher GFP signals in the cells and low level detection by Western analysis is not uncommon and is probably due to partial misfolding of proteins that might differentially affect protein turnover and GFP visualization. Also, noticeable is the observation that in mutants with more prominent ER retention the protein levels of FcyB-GFP are often higher than in wild type, probably due to reduced vacuolar turnover.

TABLE 1
Kinetic and specificity profile of mutant versions of FcyB

$K_{m/i}$ values (μM) were determined as described under "Experimental Procedures." >500 or >1000 stand for inhibition values close to 10–20% at 0.5 and 1 mM, respectively. Results are averages of at least three independent experiments in triplicate for each concentration point. S.D. was <20%.

	Hypoxanthine	Adenine	Guanine	Cytosine
WT	11	7	17	20
V83N	16	2	7	18
S85A	78	26	16	45
N163Q	47	2	>500	73
T191A	9	11	4	3
E206A	6	7	10	14
S261A	4	6	50	17
N351A	3	1	4	16
N350D	5	8	31	6
N350A	6	>1000	25	5
N354D	3	1	>500	4

Mutants related to residues Trp-159, Trp-259, and Pro-353 show a subcellular localization picture similar to wild type (no significant ER-retention), so that their lack of function should be directly associated to lack of transport activity. In mutants showing problematic FcyB trafficking, such as those concerning polar residues Asn-81, Ser-85, Asn-163, T191, Asn-350, Asn-351, and Asn-354, we cannot rigorously conclude whether the relevant residues are critical for transport activity *per se* or whether they affect transport activity indirectly due to partial misfolding. However, the fact that most of these mutants conserve minimal activity allowed us to perform a kinetic analysis and show that they are indeed critical for substrate binding (see below).

Specificity Profile of FcyB Mutants—We performed direct radiolabeled hypoxanthine uptake competition assays to define the binding constants ($K_{m/i}$s) of FcyB for its physiological substrates in the relevant mutants (for details see "Experimental Procedures"). We could only perform transporter kinetics in mutants that preserve a measurable uptake rate of hypoxanthine. These were S85A, V83N, N163Q, T191A, E206A, S261A, N350D, N350A, N351A, and N354D. Given that $K_{m/i}$ values do not depend on the number of transporter molecules expressed in the plasma membrane, for mutants showing very low hypoxanthine uptake (< 6%), such as N350D, N351A, and N354D, we used multicopy transformants for performing the kinetic analysis. Table 1 summarizes the results obtained. The mutants analyzed could be classified to those little affecting substrate binding (V83N, T191A, E206A, S261A, N350D) and those modifying significantly and, in most cases, differentially the binding of various substrates (N163Q, N350A, N351A, N354D). Among the latter, N163Q leads to a 4-fold reduction in

the binding of hypoxanthine or cytosine, >30-fold reduction in guanine binding, and a 3.5-fold increase in adenine binding. N350A conserves binding affinities for hypoxanthine, guanine, and cytosine close to those of the wild-type (moderately increased binding for hypoxanthine and cytosine) but exhibits a dramatic loss of adenine binding ($K_i > 1000\ \mu M$). N351A has a general tendency to bind with higher affinity (4–7-fold) all purines, but not cytosine. Finally, N354D has a 4-fold increased capacity for hypoxanthine, adenine, and cytosine binding but a dramatic loss of guanine binding ($K_i > 500\ \mu M$). Noticeably, the dramatic differential effect of mutation N350A on the ability of FcyB to bind adenine *versus* hypoxanthine or cytosine binding was nicely reflected in growth tests showing that N350A grows better on hypoxanthine than on adenine and is sensitive to 5-FC (see Fig. 4*A*). In summary, kinetic analysis of mutants showed that residues Asn-163, Asn-350, and Asn-354 are critical for substrate specificity, whereas Asn-163 and Asn-350 are mostly critical for substrate binding affinities.

DISCUSSION

Evidence from structural studies and molecular simulations have strongly supported the proposal that the transporters of the 5HIR superfamily function by an alternating access mechanism in which substrate binding triggers a conformational change from an outward-facing open conformation through an occluded structure to the an inward-facing open state of the protein (32–34, 36, 39). However, many aspects of this mechanism, as for example the exact steps along the substrate and ion translocation pathway or how gating operates in each case, remain to be investigated. Our current work did not intend to approach the alternating mechanism of transport by members of the 5HIR superfamily. It rather focused on the identification of specific residues that determine the function, transport kinetics, and substrate specificity of a eukaryotic member of the NCS1 family.

The first part of our work presents a theoretical structural model of the FcyB purine-cytosine transporter that reveals a number of important aspects concerning how this transporter might bind and transport its substrates. To validate the theoretical conclusions, we proceeded in a systematic mutational analysis of putatively critical residues. Our experimental results fully support the involvement of specific residues in TMS1, TMS3, TMS6, and TMS8 in substrate binding and/or transport. In particular, residues Ser-85, Trp-159, Asn-163, Trp-259, Asn-350, Asn-351, Pro-353, and Asn-354 were shown to be irreplaceable for FcyB-mediated transport. Among these residues, Ser-85, Asn-163, Asn-350, Asn-351, and Asn-354 seem critical for determining the substrate binding affinity and/or specificity of FcyB (see Fig. 4 and Table 1). Among the irreplaceable amino acids, Asn-163 and Asn-351 might also affect FcyB stability or/and turnover as the GFP-tagged versions of these alleles are significantly blocked in the ER compared with the wild-type protein. The other residues analyzed, Asn-81, Val-83, Ser-261, Thr-191, and Glu-206, seem to have a less critical role on protein turnover or FcyB transport activity. Only mutation T191A moderately affects the affinity and substrate specificity of FcyB. Noticeably, E206A, a residue within TMS5 that has been reported to function as a putative inward-facing gate, leads to a reduction in FcyB turnover and an increase in apparent transport capacity (see Fig. 4).

Among the eight residues essential for FcyB function and specificity, Trp-159, Asn-163, Trp-259, Asn-350, and Asn-354 correspond to residues shown to be directly involved in substrate binding by the Mhp1 transporter (Trp-117, Gln-121, Trp-220, Asn-314, and Asn-318). Ser-85 is similarly, but not identically positioned, with Gln-42 in Mhp1, which is also involved in substrate binding. The rest two essential FcyB residues, Asn-351 and Pro-353, are within hydrogen bonding distance with residues proposed to be involved in direct substrate binding such that they may hold the local architecture of the binding site in a position to interact with substrates. Thus, FcyB and Mhp1, despite their important specificity difference and overall low amino acid sequence identity, share very similar residues to build their substrate binding sites; four residues are identical (Trp-159, Trp-259, Asn-350, and Asn-354, numbering of FcyB), one is highly conserved (Asn-163 of FcyB replaced by Gln-121 in Mhp1), and three are not conserved (Ser-85, Asn-351, and Pro-353 in FcyB *versus* Gln-42, Pro-315, and Ala-317 in Mhp1). In fact, the four absolutely conserved residues between FcyB and Mhp1 are also highly conserved in all NCS1 members, whereas the other four residues are differentially conserved among different homologues (see Fig. 1). In particular, Asn-163 is conserved in nearly all purine-cytosine Fcy-like NCS1 members, replaced by a Glu residue in a plant adenine-guanine-uracil transporter (AtNcs1), by a Gln in homologues specific for uracil (Fur4p/FurD) or allantoin (Dal4p/FurA), and by a Leu in a uridine-specific transporter (Fui1p). Asn-351 and Pro-353 are conserved in microbial NCS1 members specific for purine-cytosine (Fcy-like proteins) but are replaced by Ile and Ala in the plant adenine-guanine-uracil transporter (AtNcs1) and the uracil- or allantoin-specific transporters (Fur4p, FurD, FurA, and Dal4p) or by Leu and Gly in a uridine-specific homologue (Fui1p). Interestingly, Ser-85, a major amino acid in FcyB function, is not a conserved residue in the NCS1 family.

None of the single mutations constructed and analyzed in this work, including N163L, N163Q, or P353A, which introduce residues present in uracil, allantoin, hydantoin, or uridine transporters, conferred to FcyB the ability to bind or/and transport substrates other than adenine, guanine, hypoxanthine, and cytosine (results not shown). In fact, mutations N163L, N163Q, or P353A led to loss or highly reduced transport capacity of FcyB. This observation strongly suggests that substrate specificity has a more complex molecular basis and might not even be solely determined by the architecture of the binding site but also by elements acting as gates or selectivity filters or by elements contributing to protein stability. We have arrived at a similar conclusion with respect to the role of interdomain synergy in determining the specificity of another group of purine transporters, the nucleobase-ascorbate transporter (NAT/NCS2) family, which was extensively studied in our laboratory (40–42).

A genetic analysis of Fcy2p, a true orthologue of FcyB in *S. cerevisiae*, supports results presented in this work. Genetically selected and subsequent site-directed mutations have strongly suggested a major role of residues Asn-374, Pro-376, and Asn-377, equivalent to Asn-351, Pro-353, and Asn-354 in

FcyB, in substrate binding (18–19). Noticeably, the effects of mutations in Asn-374 and Asn-377 were partially due to a shift of the pK_a of an ionizable amino acid residue of the unliganded transporter (18). Interestingly, mutation S272L in TMS6, isolated as an allele nonspecific, second-site, suppressor of loss-of-function mutations T213I, N374I, or N377G, partially restored the binding of hypoxanthine and cytosine (19).

Based on the above comparative observations, the docking results with FcyB, and the mutational analysis presented herein, we propose that residues corresponding to Trp-159, Trp-259, Asn-350, Pro-353, and Asn-354 in FcyB are major elements of the substrate binding site in all NCS1 members, whereas residues corresponding to Ser-85, Asn-163, and Asn-351 are mostly specificity determinants. We propose that residues Ser-85, Asn-163, and Asn-354 make direct H-bonds with substrates and Trp-159 and Trp-259 stabilize binding through π-π-stacking interactions with the purine or pyrimidine ring, whereas residues Asn-350, Asn-351, and Pro-353 have an indirect role in substrate binding through local interactions between themselves and the residues binding substrates. The critical role of residues of the four Asn residues, Asn-163, Asn-350, Asn-351, and Asn-354, in FcyB function and specificity is strongly supported by both growth phenotypes and kinetic profiles of the corresponding mutants (Fig. 4*A* and Table 1). For example, FcyB-N350A binds similarly to wild-type FcyB hypoxanthine, guanine, and cytosine but has lost the capacity for adenine binding. FcyB-N350D has increased binding affinity specifically for hypoxanthine and 5-FC, but only the latter is transported efficiently. N354D or N163Q have specifically lost the ability to bind guanine. Finally, N351A, N354A, and N354Q transport adenine much more efficiently than hypoxanthine and 5-FC, as deduced from the growth tests shown in Fig. 4*A* (a similar growth phenotype is also exhibited by N351I; results not shown).

Our work is the first systematic functional analysis of a eukaryotic member of the NCS1 family that, moreover, is combined with theoretical approaches concerning the three-dimensional transporter structure. Given the privilege of the several Mhp1 crystal structures available corresponding to different conformations achieved during transport catalysis, we are now in a position to perform a second round of rational mutagenesis combined with unique genetic screens available in *A. nidulans* to understand in more detail not only how FcyB works and selects its substrates but also how substrate specificity has evolved within the NCS1 family.

Acknowledgment—All molecular dynamic simulations were run on Cy-tera high performance computing (HPC) facility.

REFERENCES

1. de Koning, H., and Diallinas, G. (2000) Nucleobase transporters (review). *Mol. Membr. Biol.* **17,** 75–94
2. Gournas, C., Papageorgiou, I., and Diallinas, G. (2008) The nucleobase-ascorbate transporter (NAT) family. Genomics, evolution, structure-function relationships, and physiological role. *Mol. Biosyst.* **4,** 404–416
3. de Koning, H. P., Bridges, D. J., and Burchmore, R. J. (2005) Purine and pyrimidine transport in pathogenic protozoa. From biology to therapy. *FEMS Microbiol. Rev.* **29,** 987–1020
4. Schultes, N. P., Brutnell, T. P., Allen, A., Dellaporta, S. L., Nelson, T., and Chen, J. (1996) Leaf permease1 gene of maize is required for chloroplast development. *Plant Cell* **8,** 463–475
5. Mourad, G. S., Snook, B. M., Prabhakar, J. T., Mansfield, T. A., and Schultes, N. P. (2006) A fluoroorotic acid-resistant mutant of *Arabidopsis* defective in the uptake of uracil. *J. Exp. Bot.* **57,** 3563–3573
6. Zrenner, R., Stitt, M., Sonnewald, U., and Boldt, R. (2006) Pyrimidine and purine biosynthesis and degradation in plants. *Annu. Rev. Plant Biol.* **57,** 805–836
7. Kraupp, M., and Marz, R. (1995) Membrane transport of nucleobases. Interaction with inhibitors. *Gen. Pharmacol.* **26,** 1185–1190
8. Köse, M., and Schiedel, A. C. (2009) Nucleoside/nucleobase transporters. Drug targets of the future? *Future Med. Chem.* **1,** 303–326
9. Barrett, M. P., and Gilbert, I. H. (2006) Targeting of toxic compounds to the trypanosome's interior. *Adv. Parasitol.* **63,** 125–183
10. Pfaller, M. A. (2012) Antifungal drug resistance. Mechanisms, epidemiology, and consequences for treatment. *Am. J. Med.* **125,** S3–S13
11. Pantazopoulou, A., and Diallinas, G. (2007) Fungal nucleobase transporters. *FEMS Microbiol. Rev.* **31,** 657–675
12. Diallinas G., and Gournas C. (2008) Structure-function relationships in the nucleobase-ascorbate transporter (NAT) family. Lessons from model microbial genetic systems. *Channels* **2,** 363–372
13. Vlanti, A., and Diallinas G. (2008) The *Aspergillus nidulans* FcyB cytosine-purine scavenger is highly expressed during germination and in reproductive compartments and is down-regulated by endocytosis. *Mol. Microbiol.* **68,** 959–977
14. Hamari, Z., Amillis, S., Drevet, C., Apostolaki, A., Vágvölgyi, C., Diallinas, G., and Scazzocchio, C. (2009) Convergent evolution and orphan genes in the Fur4p-like family and characterization of a general nucleoside transporter in *Aspergillus nidulans*. *Mol. Microbiol.* **73,** 43–57
15. Mourad, G. S., Tippmann-Crosby, J., Hunt, K. A., Gicheru, Y., Bade, K., Mansfield, T. A., and Schultes, N. P. (2012) Genetic and molecular characterization reveals a unique nucleobase cation symporter 1 in *Arabidopsis*. *FEBS Lett.* **586,** 1370–1378
16. Weber, E., Rodriguez, C., Chevallier, M. R., and Jund, R. (1990) The purine-cytosine permease gene of *Saccharomyces cerevisiae*. Primary structure and deduced protein sequence of the FCY2 gene product. *Mol. Microbiol.* **4,** 585–596
17. Goudela, S., Tsilivi, H., and Diallinas, G. (2006) Comparative kinetic analysis of AzgA and Fcy21p, prototypes of the two major fungal hypoxanthine-adenine-guanine transporter families. *Mol. Membr. Biol.* **23,** 291–303
18. Ferreira, T., Brèthes, D., Pinson, B., Napias, C., and Chevallier, J. (1997) Functional analysis of mutated purine-cytosine permease from *Saccharomyces cerevisiae*. A possible role of the hydrophilic segment 371–377 in the active carrier conformation. *J. Biol. Chem.* **272,** 9697–9702
19. Ferreira, T., Chevallier, J., Paumard, P., Napias, C., and Brèthes, D. (1999) Screening of an intragenic second-site suppressor of purine-cytosine permease from *Saccharomyces cerevisiae*. Possible role of Ser-272 in the base translocation process. *Eur. J. Biochem.* **260,** 22–30
20. Ferreira, T., Napias, C., Chevallier, J., and Brèthes, D. (1999) Evidence for a dynamic role for proline 376 in the purine-cytosine permease of *Saccharomyces cerevisiae*. *Eur. J. Biochem.* **263,** 57–64
21. Suzuki, S., and Henderson, P. J. (2006) The hydantoin transport protein from *Microbacterium liquefaciens*. *J. Bacteriol.* **188,** 3329–3336
22. Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E. P., Rutherford, N. G., Hadden, J. M., O'Reilly, J., Ma, P., Saidijam, M., Patching, S. G., Hope, R. J., Norbertczak, H. T., Roach, P. C., Iwata, S., Henderson, P. J., and Cameron, A. D. (2008) Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322,** 709–713
23. Yamashita, A., Singh, S. K., Kawate, T., Jin, Y., and Gouaux, E. (2005) Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature* **437,** 215–223
24. Faham, S., Watanabe, A., Besserer, G. M., Cascio, D., Specht, A., Hirayama, B. A., Wright, E. M., and Abramson, J. (2008) The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na^+/sugar symport. *Science* **321,** 810–814

25. Ressl, S., Terwisscha van Scheltinga, A. C., Vonrhein, C., Ott, V., and Ziegler, C. (2009) Molecular basis of transport and regulation in the Na^+/betaine symporter BetP. *Nature* **458,** 47–52
26. Fang, Y., Jayaram, H., Shane, T., Kolmakova-Partensky, L., Wu, F., Williams, C., Xiong, Y., and Miller, C. (2009) Structure of a prokaryotic virtual proton pump at 3.2 Å resolution. *Nature* **460,** 1040–1043
27. Shaffer, P. L., Goehring, A., Shankaranarayanan, A., Gouaux, E. (2009) Structure and mechanism of a Na^+-independent amino acid transporter. *Science* **325,** 1010–1014
28. Chen, N. H., Reith, M. E., and Quick, M. W. (2004) Synaptic uptake and beyond. The sodium- and chloride-dependent neurotransmitter transporter family SLC6. *Pflugers Arch.* **447,** 519–531
29. Wright, E. M., Hirayama, B. A., and Loo, D. F. (2007) Active sugar transport in health and disease. *J. Intern. Med.* **261,** 32–43
30. Yanagida, O., Kanai, Y., Chairoungdua, A., Kim, D. K., Segawa, H., Nii, T., Cha, S. H., Matsuo, H., Fukushima, J., Fukasawa, Y., Tani, Y., Taketani, Y., Uchino, H., Kim, J. Y., Inatomi, J., Okayasu, I., Miyamoto, K., Takeda, E., Goya, T., and Endou, H. (2001) Human L-type amino acid transporter 1 (LAT1). Characterization of function and expression in tumor cell lines. *Biochim. Biophys. Acta* **1514,** 291–302
31. Harada, N., Nagasaki, A., Hata, H., Matsuzaki, H., Matsuno, F., and Mitsuya, H. (2000) Down-regulation of CD98 in melphalan-resistant myeloma cells with reduced drug uptake. *Acta Haematol.* **103,** 144–151
32. Abramson, J., and Wright, E. M. (2009) Structure and function of Na^+ symporters with inverted repeats. *Curr. Opin. Struct. Biol.* **19,** 425–432
33. Krishnamurthy, H., Piscitelli, C. L., and Gouaux, E. (2009) Unlocking the molecular secrets of sodium-coupled transporters. *Nature* **459,** 347–355
34. Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N. G., Hadden, J. M., Sharples, D., Sansom, M. S., Iwata, S., Henderson, P. J., and Cameron, A. D. (2010) Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* **328,** 470–473
35. Singh, S. K., Piscitelli, C. L., Yamashita, A., and Gouaux, E. (2008) A competitive inhibitor traps LeuT in an open-to-out conformation. *Science* **322,** 1655–1661
36. Krishnamurthy, H., and Gouaux, E. (2012) X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature* **481,** 469–474
37. Gao, X., Lu, F., Zhou, L., Dang, S., Sun, L., Li, X., Wang, J., and Shi, Y. (2009) Structure and mechanism of an amino acid antiporter. *Science* **324,** 1565–1568
38. Weyand, S., Shimamura, T., Beckstein, O., Sansom, M. S., Iwata, S., Henderson, P. J., and Cameron, A. D. (2011) The alternating access mechanism of transport as observed in the sodium-hydantoin transporter Mhp1. *J. Synchrotron. Radiat.* **18,** 20–23
39. Adelman, J. L., Dale, A. L., Zwier, M. C., Bhatt, D., Chong, L. T., Zuckerman, D. M., and Grabe, M. (2011) Simulations of the alternating access mechanism of the sodium symporter Mhp1. *Biophys. J.* **101,** 2399–2407
40. Papageorgiou, I., Gournas, C., Vlanti, A., Amillis, S., Pantazopoulou, A., and Diallinas, G. (2008) Specific interdomain synergy in the UapA transporter determines its unique specificity for uric acid among NAT carriers. *J. Mol. Biol.* **382,** 1121–1135
41. Kosti, V., Papageorgiou, I., and Diallinas, G. (2010) Dynamic elements at both cytoplasmically and extracellularly facing sides of the UapA transporter selectively control the accessibility of substrates to their translocation pathway. *J. Mol. Biol.* **397,** 1132–1143
42. Amillis, S., Kosti, V., Pantazopoulou, A., Mikros, E., and Diallinas, G. (2011) Mutational analysis and modeling reveal functionally critical residues in transmembrane segments 1 and 3 of the UapA transporter. *J. Mol. Biol.* **411,** 567–580
43. Sali, A., and Blundell, T. L. (1993) Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815
44. Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118,** 11225–11236
45. Bowers, K., Chow, E., Xu, H., O. Dror, R., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., Salmon, J. K., Shan, Y., and Shaw, D. E. (2006) *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)*, Tampa, Florida, November 11–17, 2006
46. Lomize, M. A., Lomize, A. L., Pogozheva, I. D., and Mosberg, H. I. (2006) OPM. Orientations of proteins in membranes database. *Bioinformatics* **22,** 623–625
47. Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A., Farid, R. (2006) Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **49,** 534–553
48. Koukaki, M., Giannoutsou, E., Karagouni, A., and Diallinas, G. (2003) A novel improved method for *Aspergillus nidulans* transformation. *J. Microbiol. Methods.* **55,** 687–695
49. Edgerton-Morgan, H., and Oakley, B. R. (2012) γ-Tubulin plays a key role in inactivating APC/C^{Cdh1} at the G_1-S boundary. *J. Cell Biol.* **198,** 785–791
50. Diallinas, G., Gorfinkiel, L., Arst, H. N. Jr., Cecchetto, G., and Scazzocchio, C. (1995) Genetic and molecular characterization of a gene encoding a wide specificity purine permease of *Aspergillus nidulans* reveals a novel family of transporters conserved in prokaryotes and eukaryotes. *J. Biol. Chem.* **270,** 8610–8622
51. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
52. Gournas, C., Amillis, S., Vlanti, A., and Diallinas, G. (2010) Transport-dependent endocytosis and turnover of a uric acid-xanthine permease. *Mol. Microbiol.* **75,** 246–260
53. Cecchetto, G., Amillis, S., Diallinas, G., Scazzocchio, C., and Drevet, C. (2004) The AzgA purine transporter of *Aspergillus nidulans*. Characterization of a protein belonging to a new phylogenetic cluster. *J. Biol. Chem.* **279,** 3132–3141
54. Gorfinkiel, L., Diallinas, G., and Scazzocchio, C. (1993) Sequence and regulation of the uapA gene encoding a uric acid-xanthine permease in the fungus *Aspergillus nidulans*. *J. Biol. Chem.* **268,** 23376–23381
55. Amillis, S., Cecchetto, G., Sophianopoulou, V., Koukaki, M., Scazzocchio, C., Diallinas, G. (2004) Transcription of purine transporter genes is activated during the isotropic growth phase of *Aspergillus nidulans* conidia. *Mol. Microbiol.* **52,** 205–216
56. Vlanti, A., Amillis, S., Koukaki, M., Diallinas, G. (2006) A novel-type substrate selectivity filter and ER exit determinants in the UapA purine transporter. *J. Mol. Biol.* **357,** 808–819
57. Erpapazoglou, Z., Kafasla, P., and Sophianopoulou, V. (2006) The product of the SHR3 orthologue of *Aspergillus nidulans* has a restricted range of amino acid transporter targets. *Fungal Genet. Biol.* **43,** 222–233

Fungal Genetics and Biology 63 (2014) 1–8

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Tools and Techniques

Transport assays in filamentous fungi: Kinetic characterization of the UapC purine transporter of *Aspergillus nidulans*

Emilia Krypotou, George Diallinas *

Faculty of Biology, University of Athens, Panepistimioupolis, Athens 15784, Greece

ARTICLE INFO

Article history:
Received 9 October 2013
Accepted 10 December 2013
Available online 16 December 2013

Keywords:
Germinating conidiospores
Uric acid
Xanthine
Uptake
NAT
Specificity

ABSTRACT

Transport assays allow the direct kinetic analysis of a specific transporter by measuring apparent K_m and V_{max} values, and permit the characterization of substrate specificity profiles through competition assays. In this protocol we describe a rapid and easy method for performing uptake assays in the model filamentous ascomycete *Aspergillus nidulans*. Our method makes use of *A. nidulans* germinating conidiospores at a defined morphological stage in which most transporters show maximal expression, avoiding technical difficulties associated with the use of mycelia. In combination with the ease of construction of genetic null mutants in *A. nidulans*, our method allows the rigorous characterization of any transporter in genetic backgrounds that are devoid of other transporters of similar specificity. Here, we use this method to characterize the kinetic parameters and the specificity profile of UapC, a uric acid-xanthine transporter present in all ascomycetes and member of the ubiquitous Nucleobase-Ascorbate Transporter family, in specific genetic backgrounds lacking other relevant transporters.

1. Introduction

Plasma transmembrane (PM) transporters, channels and receptors constitute primary targets of cellular regulatory circuits controlling the communication of cells with their environment, cell nutrition, homeostasis and signaling. Their importance is reflected in the fact that the genome of all organisms contains 10–15% genes encoding such proteins (Lee et al., 2008; http://www.membranetransport.org/). In addition, transporters are related to phenomena of multidrug or antibiotic resistance/sensitivity, whereas their malfunctioning is associated with serious diseases (e.g. neurodegeneration, diabetes, cystic fibrosis; http://www.tcdb.org/disease_explore.php).

Model fungi have contributed significantly in understanding transporter function. Indeed, the first eukaryotic transporters, specific for amino acids, nucleobases or ammonium, were identified genetically in *Saccharomyces cerevisiae* (for a review see Barnett, 2008), *Aspergillus nidulans* (Darlington and Scazzocchio, 1967; Kinghorn and Pateman, 1975; Pateman et al., 1974a, 1974b; Scazzocchio, 1966) and *Neurospora crassa* (Kappy and Metzenberg, 1967) in the 60s or early 70s. Direct evidence for the function and specificity of some of them followed through the development of assays for measuring transport in living fungal cells (Arst and Scazzocchio, 1975; Grenson et al., 1970; Kappy and Metzenberg, 1967; Pall, 1970, 1971; Polak and Grenson, 1973; Scazzocchio and Arst, 1978; Tsao and Marzluf, 1976; Wolfinbarger and Kay, 1973) and through cloning and characterization of the relevant genes in the 80s and 90s (Ahmad and Bussey, 1986; Diallinas and Scazzocchio, 1989; Gorfinkiel et al., 1993; Jauniaux and Grenson, 1990; Sophianopoulou and Scazzocchio, 1989; Unkles et al., 1991). It soon became clear that fungi possess a multitude of transporters of overlapping specificities and a variety of kinetic characteristics.

One of the drawbacks of original efforts to characterize transporter function and specificity by direct uptake assays, especially but not solely before the cloning era, was the multitude of functionally similar transporters in all organisms. In our view, it is impossible to assign rigorously the kinetic characteristics and specificity in a single transporter in cases where other very similar transporters function in parallel. The development of extremely efficient reverse genetic approaches in fungi is however opening the way to study a single transporter in a 'clean' genetic background, where other relevant transporters are knocked-out (see later).

Transport assays can be easily performed, using radiolabeled substrates, with single cell organisms such as bacteria, yeasts or protozoa (de Koning et al., 2005), or with specific types of animal cells such as *Xenopus* oocytes (Sigel, 1990) or erythrocytes (Desai, 2012; Thomas et al., 2011), which can be grown as homogeneous cell suspensions. More difficulties are encountered with multicellular organisms, such as filamentous fungi, and obviously with plant cells or mammalian cell cultures. Early attempts to perform transport assays with filamentous fungi used germinating

* Corresponding author. Fax: +30 (210)7274702.
E-mail address: diallina@biol.uoa.gr (G. Diallinas).

1087-1845/$ - see front matter © 2014 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.fgb.2013.12.004

conidiospores or germlings in *N. crassa* (Tsao and Marzluf, 1976; Wolfinbarger and Kay, 1973), and young mycelia (Arst and Scazzocchio, 1975; Scazzocchio and Arst, 1978) or protoplasts (Elorza et al., 1969) in *A. nidulans*. In most cases, transport activities were expressed per mg of dry fungal mass, involved large amounts of living material, and in general proved quite laborious. An attempt to monitor growth more precisely by amino acid incorporation into proteins was also attempted, but despite its elegance this method had the drawbacks of double radiolabeling and long incubation times (Arst and Scazzocchio, 1980). In addition, early uptake assays were performed in wild type or mutant cells apparently possessing transporters with similar function with the one studied. Thus, despite some success in characterizing a handful of transport activities, these early methods never became very popular.

Work by Uygar Tazebay and one of us (G.D.) in the laboratory of Claudio Scazzocchio in the early 90s established the conditions for performing efficient and rapid transport activity measurements in *A. nidulans* (Diallinas et al., 1995; Tazebay et al., 1995, 1997). The original approach was based on the observation that germinating conidiospores at the end of the isotropic growth phase and prior to germ tube emergence show maximal proline or xanthine transporter-mediated uptake, which moreover did not depend on the constitution of the growth medium. Using this morphological stage where *A. nidulans* conidiospores resemble unicellular yeast cells, transport activities could be expressed per number of living conidiospores, rather than dry mass (Amillis et al., 2004). In addition, uptake assays could be performed with small aliquots of conidiospores distributed in eppendorfs. Importantly, all transporters studied up to date in *A. nidulans* have been shown to follow a maximal, growth medium-independent, expression pattern at this specific morphological stage of conidiospore germination. Over the last 15 years, our method has been optimized and simplified significantly so that today we can easily perform a minimum of 40 transport assays in triplicate in one day. Most importantly, today transport assays can and should be performed in genetic backgrounds lacking similar transporters. Here we give a detailed account of the method as it is presently used and use it to characterize the kinetic parameters and the specificity profile of UapC, a uric acid-xanthine transporter (Diallinas et al., 1995, 1998), in specific genetic backgrounds lacking other relevant transporters.

2. Materials and methods

2.1. Growth media, strains and DNA manipulation

Complete (CM) and minimal media (MM) and supplements for auxotrophies for *A. nidulans* were used at the concentrations given by the Fungal Genetics Stock Centre (http://www.fgsc.net). Growth tests were performed on MM at 25 °C and 37 °C, pH 6.8. Sole nitrogen sources were used at the final concentrations: Urea 5 mM, $NaNO_3$ 10 mM, purines and nucleosides 0.5 mM. In the presence of Urea as sole nitrogen source, Oxypurinol, 5-Fluorouracil and 5-Fluorocytosine were used at final concentrations of 100 μM and 5-Fluorouridine at a final concentration of 10 μM. The principal strains used in this work are listed in Table 1. All single null mutations have been previously described (Amillis et al., 2007; Diallinas et al., 1998; Hamari et al., 2009; Pantazopoulou et al., 2007; Vlanti and Diallinas, 2008). The multiply deleted strains were made by standard genetic crossing using complementary auxotrophic markers. High fidelity and conventional PCR reactions were performed using KAPA-HiFi and KAPA-Taq DNA polymerases respectively (Kapa Biosystems, Lab Supplies Scientific SA, Hellas). Radiolabelled [8-^{3}H]-xanthine (22.8 Ci/mmol) and [2,8-^{3}H]-hypoxanthine (27.7 Ci/mmol) were purchased from Moravek Biochemicals, Brea, CA.

Table 1
A. nidulans strains used in this study. All strains also carry the *veA1* mutation.

Name	Genotype
wt	*pabaA1*
UapC$^-$	*uapC::AfpyrG uapAΔ azgAΔ fcyB::argB furA::AfriboB furD::AfriboB cntA::AfriboB pabaA1 pantoB100*
UapC$^+$	*uapAΔ azgAΔ fcyB::argB furD::AfriboB furA::AfriboB cntA::AfriboB pabaA1 pyroA4*
UapA$^+$	*uapC::AfpyrG azgAΔ pabaA1*
ACA-1	*ACA-1 uapA24 uapC201-401 azgA4 yA2*

2.2. Uptake method

2.2.1. Establishing a morphological stage for transport measurements

Uptake assays were performed with germinating conidiospores at a morphological stage after polarity establishment and just prior to germ tube emergence. This stage, which usually needs 4–5 h of conidiospore incubation at 37 °C or 8–9 h at 25 °C, can be followed and identified microscopically, as previously described (Amillis et al., 2004). It has been previously shown that all genes encoding transporters of nitrogenous compounds tested (UapA, UapC, AzgA, FcyB, FurD, FurA, CntA, PrnB, AgtA, UreA, etc.) are maximally transcribed and expressed at the above morphological stage (Abreu et al., 2010; Amillis et al., 2004, 2007; Hamari et al., 2009; Pantazopoulou et al., 2007; Tazebay et al., 1995; Vlanti and Diallinas, 2008). A similar observation was also obtained for AcpA, an acetate transporter (J. Sá-Pessoa, M. Casal and G. Diallinas, unpublished results). In selected cases (UapA, UapC, AzgA and PrnB), it has been shown that appearance of transport activities is due to *de novo* transcription of the corresponding genes, which starts early during the isotropic growth phase, prior to the first nuclear division. Most interestingly, transcriptional activation of transporters is independent of the constituents of the growth medium. In addition, activation of transporter transcription at this stage is independent of pathway-specific transcriptional factors (e.g. UaY or PrnA) and only partially dependent on general transcription regulators (e.g. AreA). This development control is eventually overridden by regulatory control circuits (e.g. substrate induction, N or C catabolite repression), which start to operate at the end of the germination phase and at the onset of mycelium development (Amillis et al., 2004; Tazebay et al., 1995). We have proposed that, unlike *S. cerevisiae* which uses transporter-like true sensors (Rubio-Texeira et al., 2010), *A. nidulans* seems to employ the transient, developmentally controlled, activation of transporter during germination as a nutrient sensing mechanism (Amillis et al., 2004).

For all transporters studied so far and for the present study of UapC, we used approximately 10^8 conidiospores germinating for 4–4.5 h at 37 °C, at 130–150 rpm, in liquid MM supplemented with 1% w/v Glucose as a carbon source and urea or nitrate as a nitrogen source. Empirically, approximately 10^8 conidiospores can be initially obtained, using a spatula, from a 4 day-grown culture, incubated on CM, at 37 °C. The inoculum, sufficient for 40–50 assays, is transferred in 5 ml of 0.01% v/v Tween80, briefly vortexed in order to separate the conidiospores from vegetative structures, filtered through a nylon net filter (Blutex, pore size $\geqslant$ 50 μm), and used to inoculate a 100 ml Erlenmeyer flask containing 25 ml MM supplemented with appropriate carbon and nitrogen sources and necessary vitamins. The amount of conidiospores can be estimated using a Neubauer counting-chamber slide or by measuring viable conidiospores after standard serial dilutions and plating on CM.

2.2.2. Uptake assays

A stock solution of the radiolabeled substrate of interest is prepared in water or MM, so that for each assay 25 μl of the stock will be used. The stock concentrations should be adapted to the specific activity of the available radiolabelled substrate, and should ideally be 10- to 50-fold lower from the K_m value of the transporter studied (see below). After reaching the optimal germination time, the conidiospores are collected by centrifugation for 5 min at 4000 rpm, at 15–25 °C, resuspended in 4 ml standard MM and distributed in 75 μl aliquots in 1.5 ml Eppendorf tubes. Conidiospore suspensions can be kept at 4 °C for at least 24 h without loss of transport activity. Prior to uptake, aliquots are incubated in a heat block at 37 °C (or 25 °C, if needed) for 5–10 min, before addition of radiolabeled substrate. A crucial parameter for estimating the kinetic parameters of solute uptake for any given transporter is the time of exposure to a relevant radiolabeled substrate. For measuring initial uptake rates, which are necessary for determining $K_{m/i}$ and apparent V_{max} values, the time of incubation is defined through time-course experiments at different concentrations of labeled substrate. Most transporters show linearly increased activities for at least 1 min. Time points usually are 10, 20, 30, 60, 120 and 240 s. For steady state substrate accumulation a period of $\geqslant$5 min is usually used. For each time point, measurements are performed in triplicate. The transport reaction is stopped by adding an equal volume (100 μl) of ice-cold unlabeled substrate at 100- to 1000-fold excess concentration, brief vortex, and direct transfer of the assay/eppendorf tube in an ice bucket. The conidiospores are then collected by centrifugation at 11,000*g* for 3–5 min at 4 °C, the supernatant is discarded through vacuum aspiration using a capillary Pasteur pipette, washed by vortexing in 1 ml ice-cold MM, followed by a second centrifugation at the same conditions as above. The pellet is resuspended in 0.5–1 ml of scintillation fluid and the eppendorf tubes are placed into scintillation vials for analyzing substrate accumulation in a scintillation counter.

For K_m determination of a transporter, different substrate concentrations should be used, for a fixed incubation time, previously determined to reflect initial uptake rates. In most cases this is 1 min. The range of concentrations used is determined at first empirically. In the final experiments, at least 3–4 concentration points below and above the apparent K_m value should be used. For each concentration point measurements are performed in triplicate. The stock solutions are prepared using a mixture of fixed radiolabeled substrate and increasing concentrations of non-radiolabeled substrate. K_i measurements are determined by estimating IC_{50} values (concentration for obtaining 50% inhibition) of a given substrate/inhibitor, using the formula $K_i = IC_{50}/1 + [S]/K_m$, where [S] is the fixed concentration of radiolabeled substrate used. In our assays, K_i values equal IC_{50} values since the [S] is very low (at least 10-fold lower than the K_m value). The method is in fact identical to the one for K_m determination, but stock solutions are prepared using a mixture of fixed radiolabeled substrate and increasing concentrations of non-radiolabeled other putative substrates, ligands or inhibitors. K_m values for a radiolabeled substrate can also be estimated at different concentrations of established inhibitors, in order to test whether inhibition is competitive or not (i.e. single substrate binding site).

Uptake measurements at different pHs or in the presence of proton gradient inhibitors and other energy uncouplers, are also commonly performed in order to define whether a transporter is a facilitator (energy-independent) or a H^+ symporter. For the effect of pH, the assay medium is modified accordingly. The H^+-uncoupler carbonylcyanide chlorophenylhydrazone (CCCP) and the H^+-ATPase inhibitor *N,N′*-dicyclohexylcarbodiimide (DCCD) are commonly added at final concentrations of 30 μM and 100 μM respectively, for 5–10 min, during the pre-incubation period, before the assay, at 37 °C. If needed, the effect of Na^+ or K^+ ions (100 mM) on transporter activity, can also be tested by modifying the uptake medium.

2.2.3. Analysis of transport measurements

Radioactive counts are converted to the amount of moles of substrate taken up by a defined number of germinating conidia over a defined period of time, based on the concentration and specific activity of the stock of radioactive substrate used. The actual number of viable conidiospores in each assay is determined, as described above, by counting individual colonies originating from a series of conidiospore dilutions. Background, transporter-independent, radioactive counts, obtained due to unspecific labeling of cells or in some cases by simple diffusion, are estimated using control assays with an isogenic transporter-deleted strain, or in the simultaneous presence of 100–1000 fold excess non-labeled substrate. Background counts are subtracted from the values obtained in strain expressing the relevant transporter. K_m, K_i and V_{max} determination is carried out using standard Lineweaver–Burk or Eadie–Hofstee equations formulas, or a relevant software (e.g. GraphPad Prism). The analysis is performed through a non-linear regression curve fit and sigmoidal dose response analysis. Quality factors for the analysis result are: R^2 which should be >0.99 and the Hill co-efficient (Weiss, 1997), which should be approximately −1 for a transporter with one binding site.

3. Results and discussion

3.1. Controversy over the specificity of UapC

Growth tests of UapC loss-of-function mutants strongly suggested that UapC is a secondary xanthine and uric acid transporter, but also provided evidence that it might also transport, albeit with very low efficiency, other purines. The main evidence for this last assumption was based on the observation that the leakiness of growth of an *azgA*$^-$ or *azgAΔ* strain on hypoxanthine or adenine is reduced in the presence of *uapC*$^-$ or *uapCΔ* loss-of-function mutations (Scazzocchio and Gorton, 1977). In addition, preliminary efforts to assay the function of UapC, expressed from an integrative plasmid, in a *uapA*$^-$ *azgA*$^-$ background, showed some evidence for UapC-dependent hypoxanthine accumulation (Diallinas et al., 1998). Thus, UapC was tentatively referred as a general purine permease (Diallinas et al., 1995), although no rigorous kinetic and substrate profile has ever been performed. However, a number of more recent preliminary observations have shed doubt on the ability of UapC to transport purines other than uric acid and xanthine. More specifically, overexpression of UapC resulted in significant increase in uric acid or xanthine transport, but to no detectable increase in hypoxanthine or adenine accumulation (G. Diallinas, unpublished results). In addition, evidence has been obtained for a complex functional interplay among purine–pyrimidine transporters, so that, the plasma membrane localization and expression of one can have a moderate, but detectable, indirect dominant effect on the expression or localization of another transporter (E. Krypotou, S. Amillis and G. Diallinas, unpublished observations).

To resolve the above issues concerning the function and specificity of UapC, we used standard reverse genetic approaches (see Materials and methods) to construct two novel, multiply knocked-out mutant strains which allowed rigorous investigation of the function and specificity of UapC. The first strain lacks all known transporters specific for purines–pyrimidines (UapC, UapA, AzgA, FurD and FcyB), allantoin (FurA) or nucleosides (CntA). Although FurA and CntA have been previously shown not to contribute to detectable nucleobase uptake (Hamari et al., 2009), the relevant knock-out alleles were included in the present study

for a broader assignment of the specificity of UapC. The second mutant strain constructed is isogenic to the one described above, but it has an intact *uapC* gene (UapC$^+$ UapAΔ AzgAΔ FurDΔ FurAΔ FcyBΔ CntAΔ). Both mutant strains, which proved fully viable, were tested for growth on purines, nucleosides or allantoin as nitrogen sources, or for their resistance/sensitivity to purine/pyrimidine/nucleoside analogues. Fig. 1 highlights our results; it is clearly observed that the presence of a functional UapC transporter contributes to growth on uric acid and xanthine as sole nitrogen sources, leads to partial sensitivity to oxypurinol, but has no effect on growth on any other purine or analogue tested. The growth on uric acid and xanthine is significantly less than in the control strain expressing UapA, which is the major transporter for these purines. This result strongly suggests that UapC is a low capacity xanthine/uric acid transporter which also contributes to oxypurinol uptake, but has no physiologically significant ability to transport other natural purines or pyrimidines.

3.2. UapC is a H^+ symporter specific for high-affinity xanthine transport and moderate affinity uric acid transport

We used the strain expressing solely UapC to further characterize its kinetic parameters and specificity. Fig. 2A shows that, similarly to all other *A. nidulans* transporters studied, UapC-mediated radiolabeled xanthine transport is linear for 2 min at 37 °C, and reaches a plateau at 4–6 min. At 25 °C, UapC activity is approximately 50% of that measured at 37 °C and increases linearly with time for only 1 min. Fig. 2B shows that UapC activity is fully inhibited in the presence of CCCP, a standard proton gradient uncoupler, strongly suggesting that UapC, as all known *A. nidulans* secondary active transporters, is a H^+ symporter. Using initial uptake rates we established the apparent K_m and V_{max} of UapC for xanthine, as described in Materials and methods. Fig. 2C and D establish that UapC has high-affinity ($K_{m/i}$ 3.3 μM) for xanthine, but only a moderate affinity (K_i 135 μM) for uric acid. The 40-fold difference in the affinity for xanthine versus uric acid contrasts what was found for UapA, which has practically identical high affinities for these purines (10–30 μM; Goudela et al., 2005). In fact, UapC rather resembles its homologues in *A. fumigatus* (Goudela et al., 2008) and *Candida albicans* (Goudela et al., 2006), which also have increased affinity for xanthine compared to uric acid.

3.3. UapC can recognize with very low affinity several nucleobases, but can solely transport oxypurinol

Radiolabelled xanthine transport competition assays were performed, as described in Materials and methods, to further establish the substrate specificity profile of UapC. Fig. 3 shows that, among natural nucleobases, 2000-fold excess (1 mM) of unlabeled adenine, hypoxanthine, guanine or uracil led to 40–60% inhibition of UapC-mediated radiolabelled xanthine transport. A stronger inhibitory effect was observed with some purine analogues, such as oxypurinol, 3-, 7- or 8-methylxanthine (60–75% inhibition) and especially with 1-methylxanthine or 2-thioxanthine (~90% inhibition). In contrast, excess allopurinol, purine, cytosine, thymine, L-ascorbic acid (Fig. 3), or nucleosides (not shown) did not, or little, inhibited UapC activity. Finally, UapC showed a very moderate increase in activity in the presence of allopurinol. This increase is however much less than that seen previously for UapA in the presence of allopurinol (Diallinas, 2013). In conclusion, the competition profile is clearly distinct from that established previously for UapA (Goudela et al., 2005).

Fig. 1 shows that the presence of UapC clearly improved growth on uric acid and xanthine, which is compatible with the kinetic analysis of UapC. However, despite the low affinity binding of UapC measured for purines other than xanthine and uric acid, the utilization of hypoxanthine or adenine was not improved in the presence of UapC. This observation supported the notion that UapC can bind purines other than xanthine and uric acid, but cannot transport them. To further confirm this, we also performed direct uptake assays of radiolabelled hypoxanthine of the strain expressing solely UapC and the experiments showed that UapC cannot mediate detectable hypoxanthine uptake (data not shown). Thus our results collectively show that UapC, despite showing a more promiscuous substrate *recognition* ability compared to UapA, is unable to transport any purine other than xanthine, uric acid or oxypurinol.

3.4. A chimeric UapA–UapC molecule including TMS8-TMS11 of UapC shows a kinetic profile similar to UapC, but a specificity profile similar to UapA: further evidence of independent specificity domains in the NAT family

Previous genetic, biochemical and structural modeling analyses have led to the identification of four residues in UapA interacting directly with substrates (xanthine) (Amillis et al., 2011; Kosti et al., 2012). These residues correspond to F155 in transmembrane segment 3 (TMS3), E356 in TMS8, and A407 and Q408 in TMS10. Of these four, E356 (TMS8) and Q408 (TMS10), play the most important role as they contribute to strong H bonds through their polar side chains, whereas F155 and A407 are only involved in weak H bonds through their backbones. These residues are highly conserved in the bacterial homologues XanQ, XanP and YgfU, where it has also been shown that they directly interact with substrates (Frillingos, 2012; Karatza and Frillingos, 2005; Papakostas and Frillingos, 2012). In fact, these residues are highly conserved in all known members of the NAT family that are specific for nucleobases, including the plant homologue Lpe1 or the rat homologue rSNBT1. In contrast, members of the NAT family specific for L-ascorbate uptake, rather than for nucleobases, found only in primates, conserve E356 in TMS8, but lack the second critical polar residue (Glu or Gln) in TMS10, which is replaced by a Pro residue

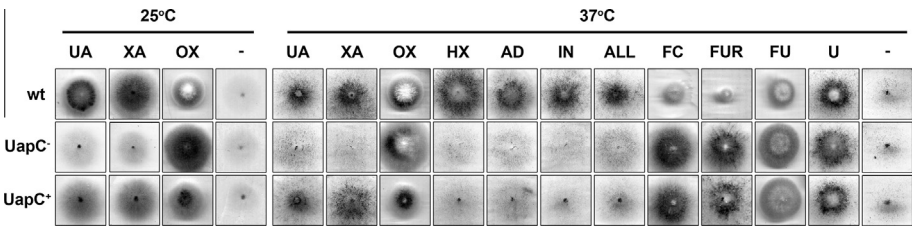


Fig. 1. UapC-mediated growth phenotypes. Growth and sensitivity/resistance tests of a standard wild type strain (wt), a strain carrying null mutations in genes encoding all known purine/pyrimidine/nucleosides transporters (UapC$^-$), and an isogenic strain expressing solely UapC (UapC$^+$). Tests shown are carried out, at 37 °C or 25 °C, in MM containing purines (UA: uric acid, XA: xanthine, HX: hypoxanthine, AD: adenine), allantoin (ALL) or nucleosides (IN: inosine) at 0.5 mM as sole nitrogen sources, or absence of any nitrogen source (–), or relevant toxic analogues at 100 μM (OX: oxypurinol, FC: 5-fluorocytosine, FU: 5-fluorouracil) or 10 μM (FUR: 5-fluorouridine), in the presence of urea (U) (5 mM) as the nitrogen source.

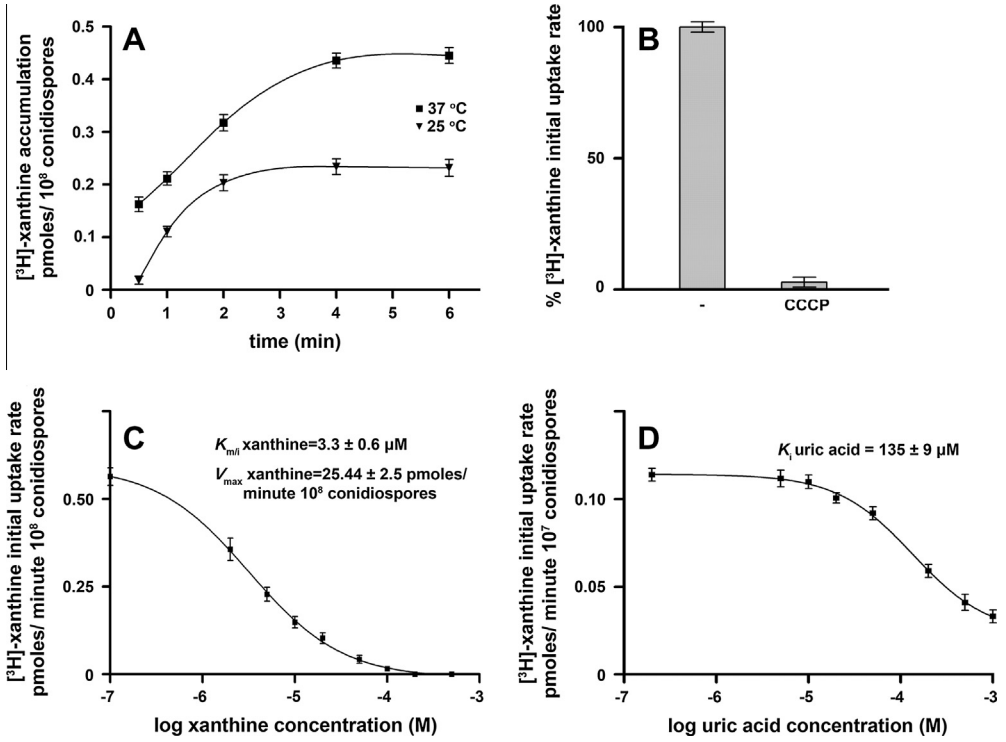


Fig. 2. Kinetic characterization of UapC. (A) Time course of the uptake of [^{3}H]-xanthine (0.2 μM) at 25 °C and 37 °C. (B) Effect of the proton gradient inhibitor CCCP (30 μM) on the initial uptake rate of [^{3}H]-xanthine (0.2 μM). (C and D) Dose response curves of UapC for xanthine and uric acid. Initial uptake rates of [^{3}H]-xanthine were measured in the simultaneous presence of increasing concentrations of non-labeled xanthine and uric acid. Apparent $K_{m/i}$ and V_{max} values were determined as described in Materials and methods. Graphs show averages of at least two independent experiments, in which triplicate assays were performed for each concentration point. Standard deviation was <20%, in all cases.

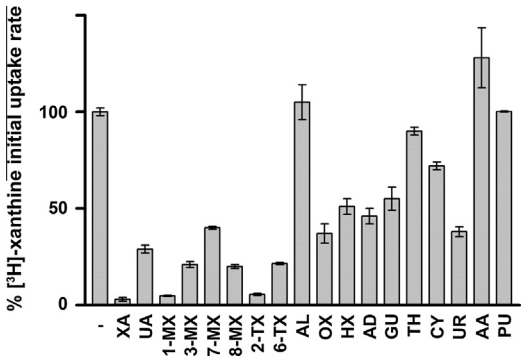


Fig. 3. Determination of the specificity profile of UapC. Initial uptake rates of [^{3}H]-xanthine (0.5 μM) were measured in the simultaneous presence of excess (1 mM) of various putative substrates or ligands. XA: xanthine, UA: uric acid, 1-MX: 1-methylxanthine, 3-MX: 3-methylxanthine, 7-MX: 7-methylxanthine, 8-MX: 8-methylxanthine, 2-TX: 2-thioxanthine, 6-TX: 6-thioxanthine, AL: allopurinol, OX: oxypurinol, HX: hypoxanthine, AD: adenine, GU: guanine, TH: thymine, CY: cytosine, UR: uracil, AA: ascorbic acid, PU: purine. Graph shows averages of two independent experiments, in which triplicate assays were performed for each putative substrate/ligand. In all cases, standard deviation was <20%.

(Fig. 4) (Diallinas and Gournas, 2008; Gournas et al., 2008; Kosti et al., 2012).

Interestingly however, while nearly all NATs specific for nucleobases seem to use practically identical residues for substrate binding, they seem to have very distinct affinities and specificities for different nucleobases. For example, bacterial or fungal NATs which are able to bind and transport xanthine with high affinity (4–10 μM) differ significantly in their binding affinity and transport rate for uric acid. More specifically, the K_m values for uric acid of UapA, UapC, Xut1 and YgfU are 10, 135, 50 and 500 μM, respectively, while XanQ and XanP do not bind uric acid at all (Diallinas and Gournas, 2008; Frillingos, 2012; Gournas et al., 2008). Furthermore, bacterial NATs specific for uracil do not recognize any other nucleobase (Lu et al., 2011). In addition rSNBT1, which can also bind and transport xanthine, cannot recognize uric acid but can transport hypoxanthine, guanine or uracil (Yamamoto et al., 2010). This observation suggests that other elements along the substrate translocation trajectory can affect the affinity for a substrate, and thus also its specificity profile. This assumption gained significant support as far as it concerns UapA, where specific amino acid residues in TMS12, TMS13 and TMS14, or in the loop linking TMS1 with TMS2, that is, at regions not involved in the formation of the major binding site, were shown to affect both its affinity for uric acid and its overall specificity (Amillis et al., 2011; Kosti et al., 2010, 2012; Papageorgiou et al., 2008).

In order to investigate the molecular basis of the difference in UapC and UapA specificity, we characterized the kinetic parameters of a functional chimeric transporter. This chimera, called ACA-1, includes the TMS8-TMS11 region of UapC (from residue 344 to 436), and TMS1–TMS7 and TMS12–TMS14 from UapA (Diallinas et al., 1998; Vlanti et al., 2006). In other words, this chimera includes the region containing the main molecular elements of the putative substrate binding site of UapC, in an otherwise UapA molecule. Results in Table 2 show that the ACA-1 chimera has affinities for xanthine and uric acid nearly identical with those of UapC, rather than of UapA. We also established by competition assays the substrate profile of the ACA-1 chimera. Fig. 5 shows that ACA-1 is highly specific for xanthine and uric acid and shows only very weak affinity for some other purines or pyrimidines. In that sense, the ACA-1 chimera resembles UapA, rather than UapC. In conclusion, the ACA-1 chimera shows a kinetic profile similar to UapC, but a substrate specificity profile similar to UapA. This provides further strong evidence for the existence of independent

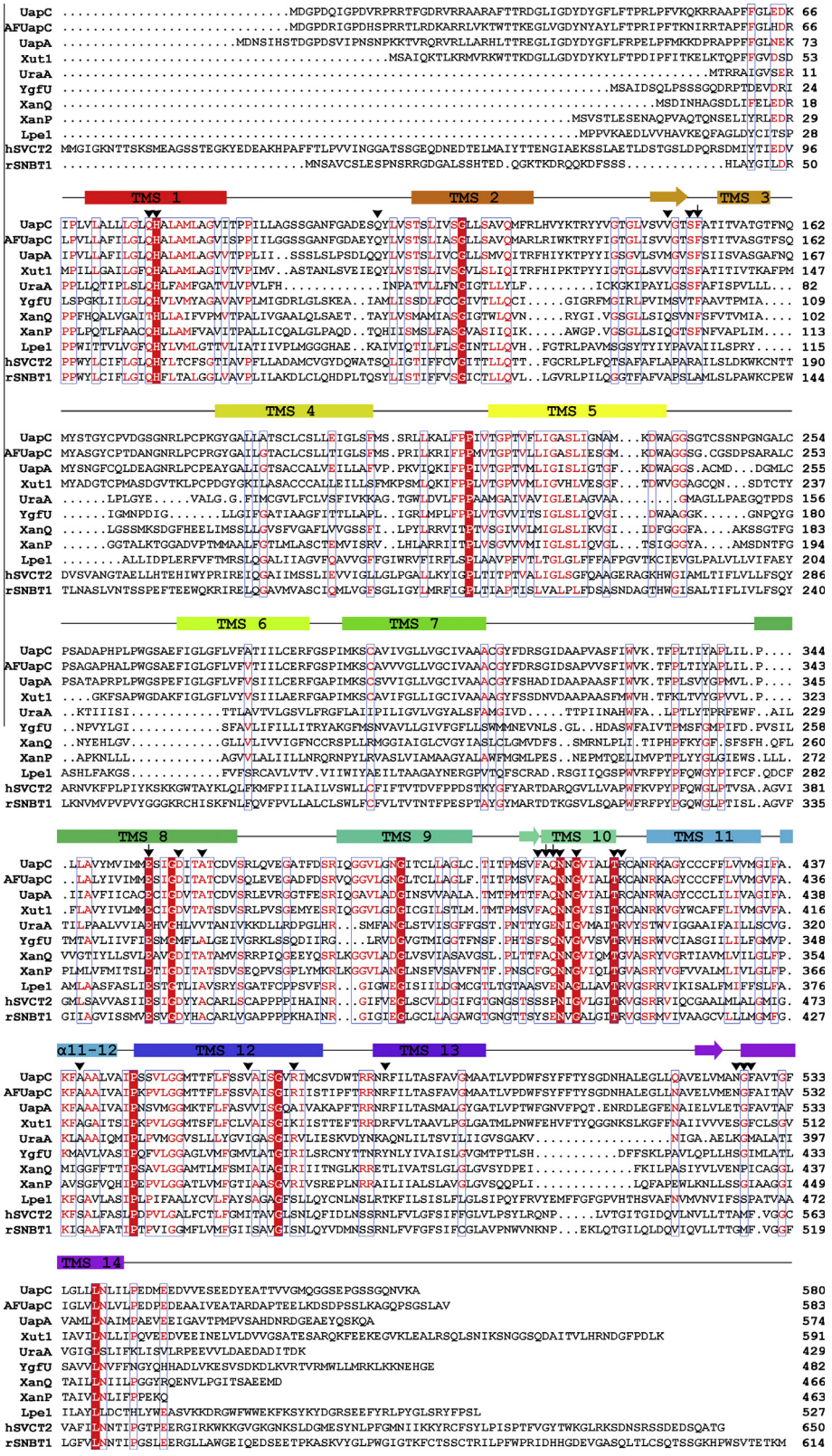


Fig. 4. Multiple alignment of UapC with selected homologues of well established specificity. Putative α-helical transmembrane segments (TMS) are indicated in colored rectangles. TMSs forming short β-sheets are shown with arrows. Invariant and highly conserved amino acids are shaded in red and blue-lined boxes, respectively. Amino acid residues of UapA shown to be involved in direct substrate binding are highlighted with arrows. Amino acid residues found to be critical for UapA function and specificity are marked with arrow heads. Transmembrane regions (TMS) are also indicated. Transporter sequences used, their specificity, and their origin are the following: UapC (xanthine-uric acid, *A. nidulans*, GI: 790973), UapA (uric acid-xanthine, *A. nidulans*, GI: 88984992), *Af*UapC (xanthine-uric acid, *A. fumigatus*, GI:74668719), Xut1 (xanthine-uric acid, *C. albicans*, GI: 60502227), XanQ (xanthine, *Escherichia coli*, GI: 161784262), XanP (xanthine, *E. coli*, GI: 84028014), YgfU (uric acid-xanthine, *E. coli*, GI: 85675700), UraA (uracil, *E. coli*, GI: 187775829), Lpe1 (uric acid-xanthine, *Zea mays*, GI: 162462794), rSNBT1 (xanthine–hypoxanthine–guanine–uracil–thymine, *Rattus norvegicus*, GI: 284010030), hSVCT2 (L-ascorbate, *Human*, GI: 6048257). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Comparison of affinities for xanthine and uric acid.

	K_m Xanthine (μM)	K_i Uric acid (μM)
UapC	3.3	135
UapA	10	30
ACA-1	2.8	130

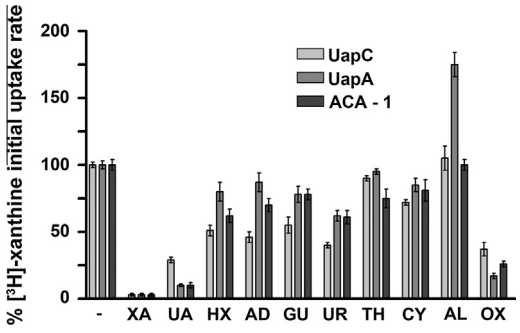


Fig. 5. Kinetic characterization and specificity profile of a UapA–UapC chimera. Relative comparison of [^{3}H]-xanthine initial uptake rates in strains expressing UapC, UapA or the chimeric protein ACA-1 (see text), in the presence of excess putative substrates/ligands. XA: xanthine, UA: uric acid, HX: hypoxanthine, AD: adenine, GU: guanine, UR: uracil, TH: thymine, CY: cytosine, AL: allopurinol, OX: oxypurinol. The results shown represent an average of two independent experiments, in which triplicate assays were performed for each substrate/ligand. Standard deviation was <20% in all cases.

elements, outside the major substrate binding site, that affect UapA or UapC function and specificity.

We have previously proposed that outward- and cytoplasmic-facing dynamic gating domains, mostly located in TMS12, TMS13 and TMS14, in the NAT family, permit, block or partially block the access of substrates to a *sensu strictu* binding site (Amillis et al., 2011; Kosti et al., 2010, 2012; Papageorgiou et al., 2008). Subtle synergistic interactions of the gating domains with other transmembrane regions and mostly those hosting the substrate binding site determine the overall function and specificity of NATs. It seems that partial genetic or evolutionary modification of such interactions has a more dramatic effect on uric acid transport than on xanthine transport. We do not know the mechanistic basis for this, but the observation that uric acid is a larger and more protonated purine, compared to xanthine, might be the basis for the difference in xanthine or uric acid transport by UapA and UapC.

4. Conclusion

We have described an updated method for the functional analysis of transporters in *A. nidulans* using ^{3}H- or ^{14}C-labeled substrate accumulation. The method utilizes germinating conidiospores at a defined morphological stage in which all transporters tested show maximal, growth medium-independent, expression, ideal for measuring apparent K_m and V_{max} values, and defining the specificity profile of a transporter. The method is rapid, efficient and rigorous, allowing the full characterization of any transporter in less than a week. The method described can be modified and adapted for most filamentous fungi that produce asexual spores. In fact, we have used the present method successfully to perform uptake assays in *A. fumigatus* (Goudela et al., 2008), but also in *C. albicans* (Goudela et al., 2006) and *S. cerevisiae* (Leung et al., 2010).

We further propose that any given transporter should, ideally, be functionally characterized in a genetic background lacking all other similar transporters, that is, in mutants where similar transporters are knocked-out, as described herein. In most cases, easy growth tests might be employed for initial testing of knock-out strains and for concluding whether a substrate is taken up by one or more transporters. Analogues of substrates should also be tested both in initial growth tests and in competition assays for defining an extended specificity profile of a transporter.

We used this method to characterize the kinetic parameters and the specificity profile of UapC, a member of the ubiquitous Nucleobase-Ascorbate Transporter family. As suggested above, the kinetic characterization of UapC was performed in appropriate genetic backgrounds lacking other relevant nucleobase transporters. UapC proved to be a secondary xanthine-uric acid/H^+ symporter, with 40-fold higher affinity for xanthine than uric acid. In addition we showed that UapC functions as a moderate affinity transporter for the drug oxypurinol, and can also recognize, but is unable to transport, other purine-related solutes with very low, probably physiologically non-significant, affinities. In conclusion, UapC compared to UapA, recognizes xanthine equally well, but shows distinct moderate to low affinities for other purines and analogues, and in particular recognizes uric acid with significantly reduced affinity. Interestingly, it is UapC that is ubiquitously present in all Aspergilli, rather than UapA. Curiously however, UapA homologues seem to exist not only in *A. nidulans* and *A. zonatus*, but also in some phylogenetic out-groups such as *N. crassa* and members of *Pezizomycotina*. Our results are expected to contribute further in understanding the evolution and the basis of structure–function-specificity relationships in the NAT family.

Acknowledgments

We are grateful to Helen Tsilivi, undergraduate ex-student, for the initial uptakes concerning UapC, to Sotiris Amillis and also to all undergraduate and Ph.D ex-students who since 2000 helped improving the uptake method.

References

Abreu, C., Sanguinetti, M., Amillis, S., Ramon, A., 2010. UreA, the major urea/H^+ symporter in *Aspergillus nidulans*. Fungal Genet. Biol. 47, 1023–1033.
Ahmad, M., Bussey, H., 1986. Yeast arginine permease: nucleotide sequence of the *CAN1* gene. Curr. Genet. 10, 587–592.
Amillis, S., Cecchetto, G., Sophianopoulou, V., Koukaki, M., Scazzocchio, C., Diallinas, G., 2004. Transcription of purine transporter genes is activated during the isotropic growth phase of *Aspergillus nidulans* conidia. Mol. Microbiol. 52, 205–216.
Amillis, S., Hamari, Z., Roumelioti, K., Scazzocchio, C., Diallinas, G., 2007. Regulation of expression and kinetic modeling of substrate interactions of a uracil transporter in *Aspergillus nidulans*. Mol. Membr. Biol. 24, 206–214.
Amillis, S., Kosti, V., Pantazopoulou, A., Mikros, E., Diallinas, G., 2011. Mutational analysis and modeling reveal functionally critical residues in transmembrane segments 1 and 3 of the UapA transporter. J. Mol. Biol. 411, 567–580.
Arst Jr., H.N., Scazzocchio, C., 1975. Initiator constitutive mutation with an 'up-promoter' effect in *Aspergillus nidulans*. Nature 254, 31–34.
Arst Jr., H.N., Scazzocchio, C., 1980. A method for monitoring macromolecular synthesis and its use to study control of nuclear acid synthesis in *Aspergillus nidulans*. J. Gen. Microbiol. 118, 443–454.
Barnett, J.A., 2008. A history of research on yeasts 13. Active transport and the uptake of various metabolites. Yeast 25, 689–731.
Darlington, A.J., Scazzocchio, C., 1967. Use of analogues and the substrate-sensitivity of mutants in analysis of purine uptake and breakdown in *Aspergillus nidulans*. J. Bacteriol. 93, 937–940.
de Koning, H.P., Bridges, D.J., Burchmore, R.J., 2005. Purine and pyrimidine transport in pathogenic protozoa: from biology to therapy. FEMS Microbiol. Rev. 29, 987–1020.
Desai, S.A., 2012. Ion and nutrient uptake by malaria parasite-infected erythrocytes. Cell Microbiol. 14, 1003–1009.
Diallinas, G., 2013. Allopurinol and xanthine use different translocation mechanisms and trajectories in the fungal UapA transporter. Biochimie 95, 1755–1764.
Diallinas, G., Gournas, C., 2008. Structure-function relationships in the nucleobase-ascorbate transporter (NAT) family: lessons from model microbial genetic systems. Channels (Austin) 2, 363–372.
Diallinas, G., Scazzocchio, C., 1989. A gene coding for the uric acid-xanthine permease of *Aspergillus nidulans*: inactivational cloning, characterization, and sequence of a cis-acting mutation. Genetics 122, 341–350.
Diallinas, G., Gorfinkiel, L., Arst Jr., H.N., Cecchetto, G., Scazzocchio, C., 1995. Genetic and molecular characterization of a gene encoding a wide specificity purine permease of *Aspergillus nidulans* reveals a novel family of transporters conserved in prokaryotes and eukaryotes. J. Biol. Chem. 270, 8610–8622.

Diallinas, G., Valdez, J., Sophianopoulou, V., Rosa, A., Scazzocchio, C., 1998. Chimeric purine transporters of *Aspergillus nidulans* define a domain critical for function and specificity conserved in bacterial, plant and metazoan homologues. EMBO J. 17, 3827–3837.

Elorza, M.V., Arst Jr., H.N., Cove, D.J., Scazzocchio, C., 1969. Permeability properties of *Aspergillus nidulans* protoplasts. J. Bacteriol. 99, 113–115.

Frillingos, S., 2012. Insights to the evolution of Nucleobase-Ascorbate Transporters (NAT/NCS2 family) from the Cys-scanning analysis of xanthine permease XanQ. Int. J. Biochem. Mol. Biol. 3, 250–272.

Gorfinkiel, L., Diallinas, G., Scazzocchio, C., 1993. Sequence and regulation of the *uapA* gene encoding a uric acid-xanthine permease in the fungus *Aspergillus nidulans*. J. Biol. Chem. 268, 23376–23381.

Goudela, S., Karatza, P., Koukaki, M., Frillingos, S., Diallinas, G., 2005. Comparative substrate recognition by bacterial and fungal purine transporters of the NAT/NCS2 family. Mol. Membr. Biol. 22, 263–275.

Goudela, S., Tsilivi, H., Diallinas, G., 2006. Comparative kinetic analysis of AzgA and Fcy21p, prototypes of the two major fungal hypoxanthine–adenine–guanine transporter families. Mol. Membr. Biol. 23, 291–303.

Goudela, S., Reichard, U., Amillis, S., Diallinas, G., 2008. Characterization and kinetics of the major purine transporters in *Aspergillus fumigatus*. Fungal Genet. Biol. 45, 459–472.

Gournas, C., Papageorgiou, I., Diallinas, G., 2008. The nucleobase-ascorbate transporter (NAT) family: genomics, evolution, structure-function relationships and physiological role. Mol. Biosyst. 4, 404–416.

Grenson, M., Hou, C., Crabeel, M., 1970. Multiplicity of the amino acid permeases in *Saccharomyces cerevisiae*. IV. Evidence for a general amino acid permease. J. Bacteriol. 103, 770–777.

Hamari, Z., Amillis, S., Drevet, C., Apostolaki, A., Vágvölgyi, C., Diallinas, G., Scazzocchio, C., 2009. Convergent evolution and orphan genes in the Fur4p-like family and characterization of a general nucleoside transporter in *Aspergillus nidulans*. Mol. Microbiol. 73, 43–57.

Jauniaux, J.C., Grenson, M., 1990. *GAP1*, the general amino acid permease gene of *Saccharomyces cerevisiae*. Nucleotide sequence, protein similarity with the other bakers yeast amino acid permeases, and nitrogen catabolite repression. Eur. J. Biochem. 190, 39–44.

Kappy, M.S., Metzenberg, R.L., 1967. Multiple alterations in metabolite uptake in a mutant of *Neurospora crassa*. J. Bacteriol. 94, 1629–1637.

Karatza, P., Frillingos, S., 2005. Cloning and functional characterization of two bacterial members of the NAT/NCS2 family in *Escherichia coli*. Mol. Membr. Biol. 22, 251–261.

Kinghorn, J.R., Pateman, J.A., 1975. Mutations which affect amino acid transport in *Aspergillus nidulans*. J. Gen. Microbiol. 86, 174–184.

Kosti, V., Papageorgiou, I., Diallinas, G., 2010. Dynamic elements at both cytoplasmically and extracellularly facing sides of the UapA transporter selectively control the accessibility of substrates to their translocation pathway. J. Mol. Biol. 397, 1132–1143.

Kosti, V., Lambrinidis, G., Myrianthopoulos, V., Diallinas, G., Mikros, E., 2012. Identification of the substrate recognition and transport pathway in a eukaryotic member of the nucleobase-ascorbate transporter (NAT) family. PLoS ONE 7, e41939.

Lee, T.J., Paulsen, I., Karp, P., 2008. Annotation-based inference of transporter function. Bioinformatics 24, 259–267.

Leung, J., Karachaliou, M., Alves, C., Diallinas, G., Byrne, B., 2010. Expression and purification of a functional uric acid-xanthine transporter (UapA). Protein Expr. Purif. 72, 139–146.

Lu, F., Li, S., Jiang, Y., Jiang, J., Fan, H., Lu, G., Deng, D., Dang, S., Zhang, X., Wang, J., Yan, N., 2011. Structure and mechanism of the uracil transporter UraA. Nature 472, 243–246.

Pall, M.L., 1970. Amino acid transport in *Neurospora crassa*. 3. Acidic amino acid transport. Biochim. Biophys. Acta 211, 513–520.

Pall, M.L., 1971. Amino acid transport in *Neurospora crassa*. IV. Properties and regulation of a methionine transport system. Biochim. Biophys. Acta 233, 201–214.

Pantazopoulou, A., Lemuh, N.D., Hatzinikolaou, D.G., Drevet, C., Cecchetto, G., Scazzocchio, C., Diallinas, G., 2007. Differential physiological and developmental expression of the UapA and AzgA purine transporters in *Aspergillus nidulans*. Fungal Genet. Biol. 44, 627–640.

Papageorgiou, I., Gournas, C., Vlanti, A., Amillis, S., Pantazopoulou, A., Diallinas, G., 2008. Specific interdomain synergy in the UapA transporter determines its unique specificity for uric acid among NAT carriers. J. Mol. Biol. 382, 1121–1135.

Papakostas, K., Frillingos, S., 2012. Substrate selectivity of YgfU, a uric acid transporter from *Escherichia coli*. J. Biol. Chem. 287, 15684–15695.

Pateman, J.A., Dunn, E., Kinghorn, J.R., Forbes, E.C., 1974a. The transport of ammonium and methylammonium in wild type and mutant cells of *Aspergillus nidulans*. Mol. Gen. Genet. 133, 225–236.

Pateman, J.A., Kinghorn, J.R., Dunn, E., 1974b. Regulatory aspects of L-glutamate transport in *Aspergillus nidulans*. J. Bacteriol. 119, 534–542.

Polak, A., Grenson, M., 1973. Evidence for a common transport system for cytosine, adenine and hypoxanthine in *Saccharomyces cerevisiae* and *Candida albicans*. Eur. J. Biochem. 32, 276–282.

Rubio-Texeira, M., Van Zeebroeck, G., Voordeckers, K., Thevelein, J.M., 2010. *Saccharomyces cerevisiae* plasma membrane nutrient sensors and their role in PKA signaling. FEMS Yeast Res. 10, 134–149.

Scazzocchio, C., 1966. Studies on the Genetic Control of Purine Oxidation in *Aspergillus nidulans*. Ph.D. Thesis, University of Cambridge, Cambridge, UK.

Scazzocchio, C., Arst Jr., H.N., 1978. The nature of an initiator constitutive mutation in *Aspergillus nidulans*. Nature 274, 177–179.

Scazzocchio, C., Gorton, D.J., 1977. The regulation of purine breakdown. In: Smith, J.E., Pateman, J.A. (Eds.), Genetics and Physiology of *Aspergillus*. Academic Press, New York, NY, pp. 255–265.

Sigel, E., 1990. Use of *Xenopus* oocytes for the functional expression of plasma membrane proteins. J. Membr. Biol. 117, 201–221.

Sophianopoulou, V., Scazzocchio, C., 1989. The proline transport protein of *Aspergillus nidulans* is very similar to amino acid transporters of *Saccharomyces cerevisiae*. Mol. Microbiol. 3, 705–714.

Tazebay, U.H., Sophianopoulou, V., Cubero, B., Scazzocchio, C., Diallinas, G., 1995. Post-transcriptional control and kinetic characterization of proline transport in germinating conidiospores of *Aspergillus nidulans*. FEMS Microbiol. Lett. 132, 27–37.

Tazebay, U.H., Sophianopoulou, V., Scazzocchio, C., Diallinas, G., 1997. The gene encoding the major proline transporter of *Aspergillus nidulans* is upregulated during conidiospore germination and in response to proline induction and amino acid starvation. Mol. Microbiol. 24, 105–117.

Thomas, S.L., Bouyer, G., Cueff, A., Egée, S., Glogowska, E., Ollivaux, C., 2011. Ion channels in human red blood cell membrane: actors or relics? Blood Cells Mol. Dis. 46, 261–265.

Tsao, T.F., Marzluf, G.A., 1976. Genetic and metabolic regulation of purine base transport in *Neurospora crassa*. Mol. Gen. Genet. 149, 347–355.

Unkles, S.E., Hawker, K.L., Grieve, C., Campbell, E.I., Montague, P., Kinghorn, J.R., 1991. *CrnA* encodes a nitrate transporter in *Aspergillus nidulans*. Proc. Natl. Acad. Sci. USA 88, 204–208.

Vlanti, A., Diallinas, G., 2008. The *Aspergillus nidulans* FcyB cytosine–purine scavenger is highly expressed during germination and in reproductive compartments and is downregulated by endocytosis. Mol. Microbiol. 68, 959–977.

Vlanti, A., Amillis, S., Koukaki, M., Diallinas, G., 2006. A novel-type substrate-selectivity filter and ER-exit determinants in the UapA purine transporter. J. Mol. Biol. 357, 808–819.

Weiss, J.N., 1997. The Hill equation revisited: uses and misuses. FASEB J. 11, 835–841.

Wolfinbarger Jr., L., Kay, W.W., 1973. Acidic amino acid transport in *Neurospora crassa* mycelia. Biochim. Biophys. Acta 330, 335–343.

Yamamoto, S., Inoue, K., Murata, T., Kamigaso, S., Yasujima, T., Maeda, J.Y., Yoshida, Y., Ohta, K.Y., Yuasa, H., 2010. Identification and functional characterization of the first nucleobase transporter in mammals: implication in the species difference in the intestinal absorption mechanism of nucleobases and their analogs between higher primates and other mammals. J. Biol. Chem. 285, 6522–6531.

Molecular Microbiology (2014) ■ doi:10.1111/mmi.12646

Modelling, substrate docking and mutational analysis identify residues essential for function and specificity of the major fungal purine transporter AzgA

Emilia Krypotou,[1] George Lambrinidis,[2†] Thomas Evangelidis,[2†] Emmanuel Mikros[2] and George Diallinas[1*]**
[1]*Faculty of Biology, University of Athens, Panepistimiopolis, Athens 15784, Greece.*
[2]*Faculty of Pharmacy, University of Athens, Panepistimiopolis, Athens 15771, Greece.*

Summary

The AzgA purine/H$^+$ symporter of *Aspergillus nidulans* is the founding member of a functionally and phylogenetically distinct transporter family present in fungi, bacteria and plants. Here a valid AzgA topological model is built based on the crystal structure of the *Escherichia coli* uracil transporter UraA, a member of the nucleobase-ascorbate transporter (NAT/NCS2) family. The model consists of 14 transmembrane, mostly α-helical, segments (TMSs) and cytoplasmic N- and C-tails. A distinct compact core of 8 TMSs, made of two intertwined inverted repeats (TMSs 1–4 and 8–11), is topologically distinct from a flexible domain (TMSs 5–7 and 12–14). A putative substrate binding cavity is visible between the core and the gate domains. Substrate docking, molecular dynamics and mutational analysis identified several residues critical for purine binding and/or transport in TMS3, TMS8 and TMS10. Among these, Asn131 (TMS3), Asp339 (TMS8) and Glu394 (TMS10) are proposed to directly interact with substrates, while Asp342 (TMS8) might be involved in subsequent substrate translocation, through H$^+$ binding and symport. Thus, AzgA and other NAT transporters use topologically similar TMSs and amino acid residues for substrate binding and transport, which in turn implies that AzgA-like proteins constitute a distant subgroup of the ubiquitous NAT family.

Accepted 8 May, 2014. For correspondence. *E-mail diallina@biol.uoa.gr; Tel. (+30) 210 7274649; Fax (+30) 210 7274702. **E-mail mikros@pharm.uoa.gr; Tel. (+30) 210 7274813; Fax (+30) 210 7274747. †Equal contribution.

Introduction

Purines and pyrimidines, besides being the precursors of nucleotide and nucleic acid biosynthesis, are also involved in cell signalling, nutrition, response to stress and cell homeostasis. Most cells in all domains of life possess specific nucleobase transporters, usually distinct from those involved in the uptake of nucleosides (De Koning and Diallinas, 2000; Gournas *et al*., 2008). Purine transporters are important in bacterial, protozoan and plant growth and development and in the susceptibility of human, protozoan, bacterial or fungal cells to purine-related drugs (Kraupp and Marz, 1995; Schultes *et al*., 1996; De Koning *et al*., 2005; Barrett and Gilbert, 2006; Mourad *et al*., 2006; Zrenner *et al*., 2006; Köse and Schiedel, 2009; Frillingos, 2012; Pfaller, 2012; Young *et al*., 2013). Fungi possess three structurally, functionally and phylogenetically distinct families of H$^+$ gradient-dependent symporters for purines and/or pyrimidines (Pantazopoulou and Diallinas, 2007; Diallinas and Gournas, 2008; Gournas *et al*., 2008). These are the NCS1 (nucleobase cation symporter 1) family, the NAT (nucleobase-ascorbate transporter) family, also known as the NCS2 family, and the AzgA-like family (see also http://www.tcdb.org/). NAT/NCS2 members are found ubiquitously in all domains of life, whereas NCS1 and AzgA-like transporters are present in prokaryotes, fungi and plants.

The AzgA family has been historically defined by the homonymous AzgA permease of the filamentous fungus *Aspergillus nidulans*. *AzgA*-specific mutations were isolated in the 60s and very early on the specificity AzgA for salvageable purines was assigned (Darlington *et al*., 1965; Darlington and Scazzocchio, 1967). The gene was cloned and characterized by a transposon methodology (Cecchetto *et al*., 2004) and soon after the AzgA transport activity and specificity were characterized in great detail (Goudela *et al*., 2006). AzgA is now known to be a high-affinity, high-capacity, H$^+$ gradient-dependent transporter, specific for adenine-hypoxanthine-guanine. AzgA also recognizes with moderate affinity specific purine analogues, some of which are cytotoxic, such as, purine or 8-azaguanine, and guanosine. On the other hand, AzgA

cannot recognize the oxidized-purines xanthine and uric acid, pyrimidines, or major purine derivatives, such as allantoin or caffeine. AzgA expression is regulated at the level of transcription, being induced in the presence of purines and in response to conidiospore germination (Amillis *et al*., 2004; Cecchetto *et al*., 2004; Pantazopoulou *et al*., 2007). A very similar gene/protein has also been studied in *Aspergillus fumigatus* (Goudela *et al*., 2008). Interestingly, while AzgA-like proteins exist in most fungi, prominent losses seem to exist in *Saccharomyces cerevisiae* and other post-genome duplication yeasts. Two plant AzgA homologues have also been identified and characterized and shown to be functionally similar to fungal AzgAs (Mansfield *et al*., 2009). Recently, two paralogous AzgA-like proteins have also been characterized in *Borrelia burgdorferi*, the tick-borne bacterium that causes the multistage inflammatory Lyme disease (Jain *et al*., 2012). Most interestingly, *B. burgdorferi* lacks the enzymes required for *de novo* synthesis of purines for RNA and DNA synthesis and therefore is dependent upon the tick vector and mammalian host environments for salvage of purine bases for nucleic acid biosynthesis. In the same work, *B. burgdorferi* lacking AzgA-like transporters was shown to be non-infectious in mice, suggesting that purine transport systems may serve as novel antimicrobial targets for the treatment of Lyme disease.

AzgA-like proteins show little primary amino acid sequence similarity (< 18% identity) with members of other nucleobase or functionally related transporter families. In addition, AzgA-like proteins do not possess the conserved motifs shown, in the NAT (Kosti *et al*., 2012) or NCS1 (Krypotou *et al*., 2012) families, to be involved in substrate binding and transport. Furthermore, earlier phylogenetic analyses placed AzgA proteins in a cluster clearly separated from the NAT or NCS1 families. However, some recent statistical methods, based on structural and sequence similarity, proposed that the AzgA family might be phylogenetically related to the amino acid-polyamine-organocation (APC) superfamily, which also includes the NAT and NCS1 families (Wong *et al*., 2012). Importantly for the present work, in a preliminary homology threading search we have observed that AzgA-like proteins are very probably structurally related to members of the NAT family.

Here, we construct and validate a 3D model of the *A. nidulans* AzgA transporter based on the inward facing, substrate-occluded, crystal structure of the uracil transporter UraA from *Escherichia coli* (Lu *et al*., 2011). Subsequently, we use this model to perform docking approaches, molecular dynamics and mutational analysis, which led to the identification of residues that are involved in substrate binding and/or transport. Our results are discussed in relationship to the phylogenetic origin of AzgA.

Results

AzgA is phylogenetically distinct from the NAT and NCS1 families

We ran a standard blastp analysis of AzgA towards available databases and subsequently analysed by phylogenetic approaches the sequences selected. We confirmed that there is no AzgA-like sequences in protozoa and metazoa. One single apparent exception is an annotated *Ceratitis capitata* homologue, which however, is very probably a microbial contaminant or a product of horizontal transfer, as it is 93% identical to AzgA-like sequences from enterobacteria. Another AzgA-like sequence of 435 amino acids appears in the annotation of the human genome (http://blast.ncbi.nlm.nih.gov), but this is also a bacterial contaminant, as it is 90% identical to a sequence from *Leuconostoc lactis*. Fungal and plant AzgA sequences cluster closer to each other (35% average identities) than to bacterial and archaeal sequences (25–30% identities). Bacterial AzgA homologues are significantly shorter (430–480 amino acids on average) than fungal and plant proteins (530–580 amino acids), mostly due to much shorter N and C termini. Within fungi, AzgA sequences are ubiquitously conserved in filamentous molds and *Schizosaccharomycetes* sp., but absent in *Saccharomycotina*, including *Candida* species. Ascomycete and Basidiomycete homologues are 57–90% and 50–67% identical, respectively, with *A. nidulans* AzgA. All 21 known genomes of *Aspergillus* sp. include a single gene encoding highly similar AzgA-like orthologues (533–581 amino acid long, 87–94% identical). This contrasts what is observed for the two other fungal families of purine transporters, the NAT and especially the NCS1, which include several and more divergent paralogues in all *Aspergillus* species. Finally, in 17 of the *Aspergillus* sp. genomes, *azgA*-like genes include 3 introns with conserved topologies. *A. nidulans*, *A. niger* and *A. aculeatus* have 4 introns, *Neosartoria fischeri* has 2, and a shorter *A. zonatus* gene has none.

Interestingly, all known AzgA-like purine transporters and homologous sequences (Fig. S1) are clustered in a closely related group, clearly distinct from representative NAT or NCS1 members (Fig. 1). The AzgA clade is however closer to the NAT clade compared with that of the NCS1 family, suggesting that there is a distant evolutionary relationship between AzgA and NAT carriers, which, as will be shown in the present work, is reflected in similar 3D topology in respect to transmembrane segments and, importantly, residues critical for substrate binding and transport.

AzgA structural model

We constructed a structural model of AzgA based on the crystal structure of the UraA uracil transporter of *E. coli*

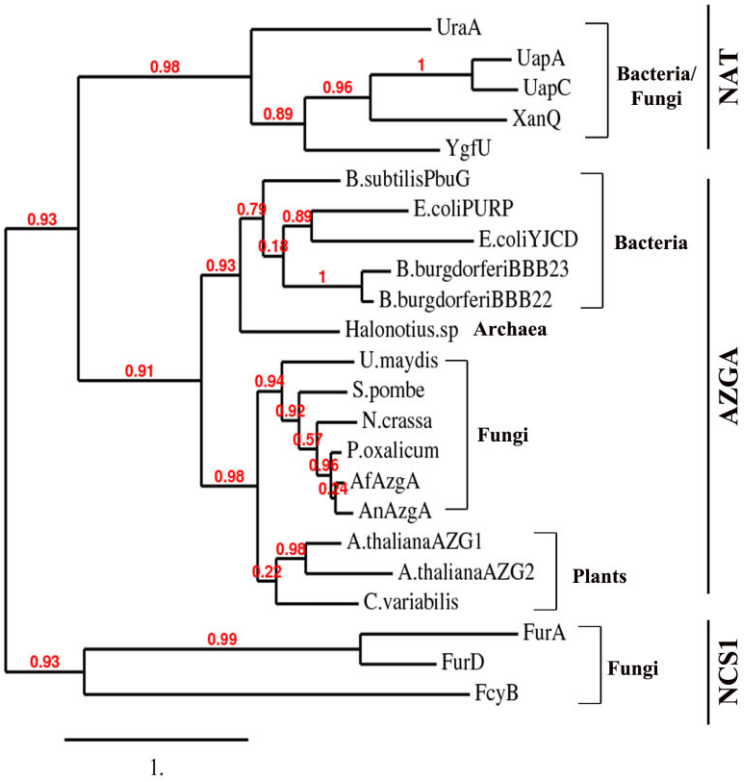


Fig. 1. Phylogenetic analysis of AzgA. A robust phylogenetic tree was constructed as described in *Experimental procedures* using AzgA-like sequences of known function, plus some selected representatives as out-groups, and selected NAT and NCS1 proteins of known function and substrate specificity. The sequences shown are: AzgA (CAE00849.1) and AzgA-like sequences from *Aspergillus fumigatus* (AFUA_5G09750), *Penicillium oxalicum* (EPS25670.1), *Neurospora crassa* (XP_965513.2), *Schizosaccharomyces pombe* (NP_596491.1), *Ustilago maydis* (XP_757903.1), *Escherichia coli* (PurP/P31466.2 and YjcD/AGX33403.1), *Bacillus subtilis* (PbuG/O34987.1), *Borellia burgdorferi* (BBB22/NP_047008.1 and BBB23/NP_047009), *Halonotius* sp. (WP_021040151.1), *Chlorella variabilis* (EFN58296.1) and *Arabidopsis thaliana* (NP_566384.1 and NP_199841.2), as well as NAT sequences from *A. nidulans* (UapA/CAA50681.2, UapC/CAA56190.1) and *E. coli* (UraA/YP_490725.1 XanQ/NP_417358.2, YgfU/YP_491089.1), and NCS1 sequences from *A. nidulans* (FurD/ABR22526.1, FurA/AN0660, FcyB/ACA81792.1).

(Lu *et al.*, 2011). The two proteins share a rather moderate sequence similarity (~ 20% identity in 418 amino acid overlap), which is however adequate for sustaining a theoretical model of AzgA. The final AzgA-UraA alignment (Fig. S2) used for building the AzgA model was based on the binary alignment with UraA, manually modified on the basis of a multiple alignment of selected AzgA-like sequences (see *Experimental procedures*).

Using the membrane comparative modelling protocol of Rosetta 3.4 we constructed 9230 topologically correct models of AzgA, which were clustered on the binding site. The top 89 cluster medoids were selected for further analysis, as described in *Experimental procedures*. All 12 substrates of AzgA with measurable K_i constants (Goudela *et al.*, 2006) were docked into the putative active site and the top-down concordance coefficient between the predicted and experimental ranking was calculated for the 89 cluster representatives using 47 different scoring schemes. In this respect, we selected the medoid of the 70th largest cluster as the best structural model of the substrate binding cavity of AzgA. That model was included placed in the top-10 list yielded by Vina, NNscore1 and NNscore2, but not DSX scoring algorithm (Durrant and McCammon, 2010; 2011a,b). The respective top-down concordance coefficients were 0.253 ($P = 0.2$) from Vina, 0.723 ($P = 0.012$) from network 9 of NNscore1 and 0.823 ($P = 0.003$) from network 17 of NNscore2 software.

The overall 3D structure of the AzgA model (Figs 2 and S3) corresponds to a cytoplasmic-facing occluded conformation, in which the substrate is not directly accessible to the cytosol. The structure consists of 14 transmembrane segments (TMSs), mostly helical in nature, that also retain the characteristic two beta strands at the core of the protein. TMSs 1–4 and 8–11 form a rather compact core domain, whereas TMSs 5–7 and 12–14 form a more flexible gate domain, similar to other NAT members modelled by similar approaches (Amillis *et al.*, 2011; Karena and Frillingos, 2011; Kosti *et al.*, 2012). As proposed for the UraA structure (Lu *et al.*, 2011), these two domains possess distinct topology and function. The transmembrane helices are connected by large loops, most of which are the result of lengthy insertions that render the prediction of accurate 3D co-ordinates very difficult for those regions. In line with UraA and UapA (Lu *et al.*, 2011; Kosti *et al.*, 2012), a mild negative electrostatic potential distribution is accumulated on the outer surface of the transporter, while the inner surface possesses strong positive charge within a region below helices 13–14, at the inner entrance of the transporter (Fig. S4). The surface of the intermediate region of the transporter is neutral, consistent with its natural position inside the membrane bilayer.

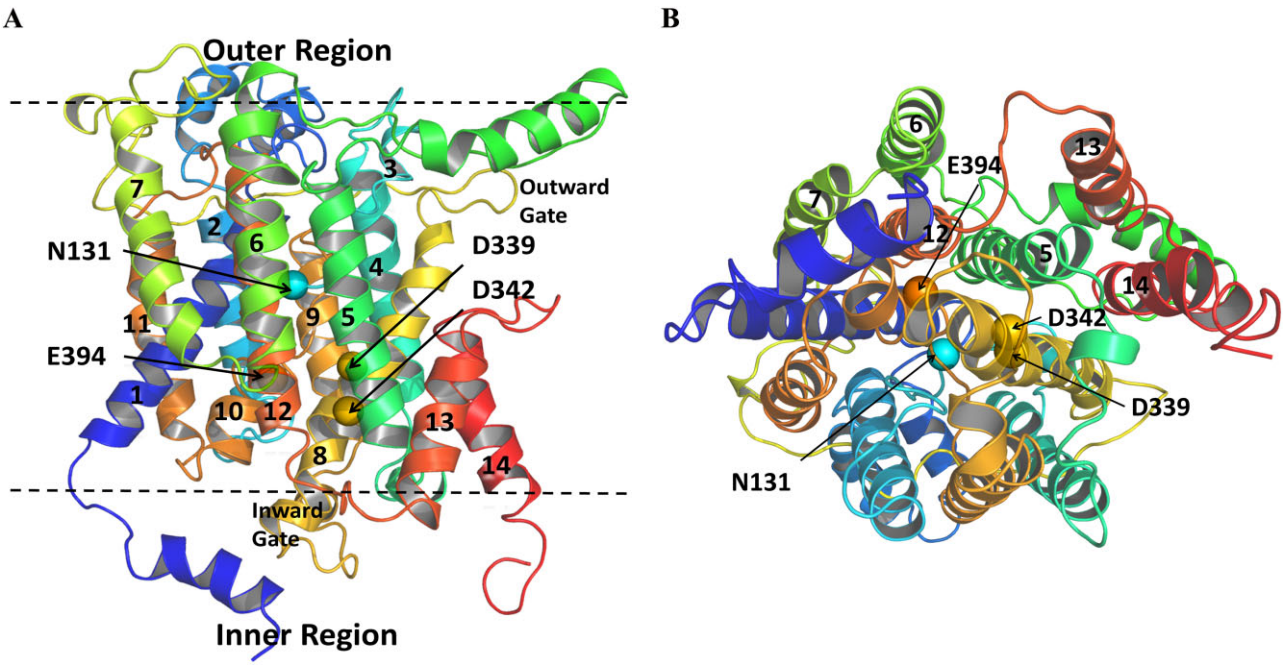


Fig. 2. Theoretical model of AzgA. The 3D AzgA model was constructed on the basis of the alignment with UraA shown in Fig. S2.
A. Side-view of the best Rosetta model (medoid of cluster 72) highlighting the topology of residues selected as crucial for the function. The transmembrane helices are designated with numbers (core domain: 1–4, 8–11; gate domain: 5–7, 12–14) and the planes of the membrane with dashed lines.
B. Bottom view of the same model. The substrate binding site is located in the space between the core and the gate domains. The topology of residues critical for the function of the substrate binding site are also shown (see also Fig. S3).

Substrate docking in AzgA

Using the AzgA model, aspects of substrate recognition were addressed using docking calculations. The three major physiological substrates, namely hypoxanthine, adenine, and guanine were docked to the modelled structure of AzgA. Starting from the structure of AzgA in complex with hypoxanthine or adenine or guanine, as derived from Autodock Vina calculations, 1000 steps of Low-Mode/Monte Carlo docking calculations were run as implemented on Macromodel v9.9 (Suite 2012: MacroModel, version 9.9, Schrödinger, LLC, New York, NY, 2012). Figure 3 shows the modelled interactions of AzgA with its three physiological substrates. A single substrate orientation is proposed for hypoxanthine (Fig. 3A) and guanine (Fig. 3B), but two different substrate positions are proposed for adenine showing 3.5 kcal mol^{-1} energy difference (Fig. 3C and D). AzgA interacts with all its physiological substrates through H bonding mediated by the polar side-chains of residues Asn131, Asp339 and Glu394, and the amide group of the backbone of Val393. This rationalizes the very similar $K_{m/i}$ values measured for hypoxanthine, adenine and guanine. According to the models proposed these residues interact with N1(H) and C6=O (hypoxanthine, guanine) or C6-NH_2 (adenine) of the pyrimidine moiety, and N7(H) and N9(H) of the imidazole moiety. In addition, the C2-NH_2 group in guanine, interacts with Asn131 (backbone) and Thr390 (side-chain) through a double hydrogen bond (Fig. 3B). N3 might only be involved in AzgA interactions with adenine, if we consider the first orientation proposed (see Fig. 3C). Finally, a weak π–π stacking interaction with Phe392 might also stabilize hypoxanthine and guanine, but not adenine, binding.

Molecular dynamics simulations of AzgA in complex with hypoxanthine, adenine and guanine

The MD trajectories of hypoxanthine, adenine and guanine capture the transition of the inward-facing occluded state to an inward-facing open conformation. This transition happens very early in all 3 trajectories, when transmembrane helices (TM) 13 and 14 of the gate domain move upwards, thus opening a putative inward-facing gate (see Fig. 2). A similar gate was also described in the uric-acid-xanthine/H^+ symporter UapA (Kosti *et al.*, 2012). At the same time, the movement of TM helices 13 and 14, as well as bending of TM helix 5, elicits a flood of the substrate binding cavity with cytosolic water. Although water is present in the cavity for almost 190 ns (see Video S1), hypoxanthine remains constantly bound to the side-chains of Asn131, Asp339, Glu394 and the backbone of Val393

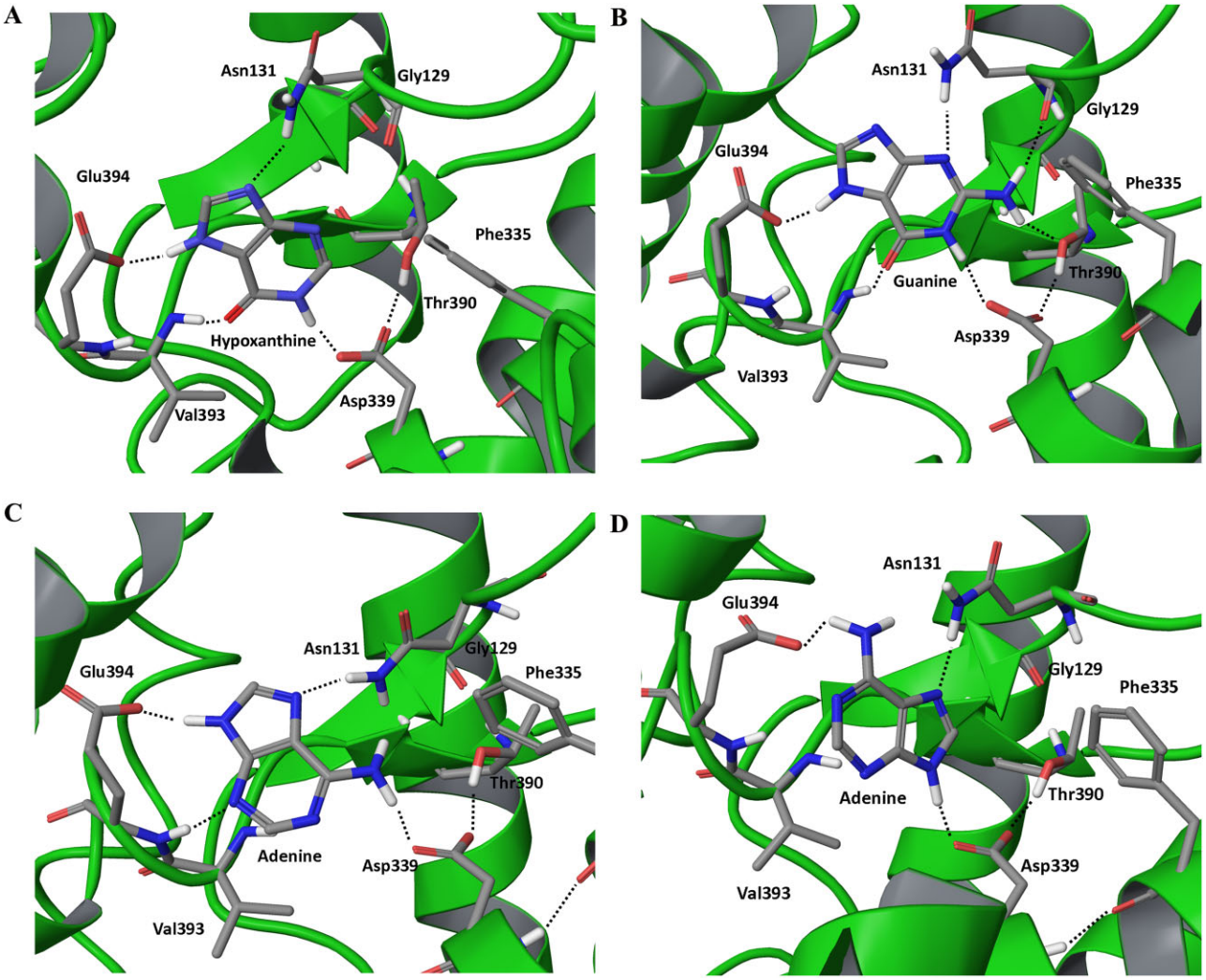


Fig. 3. Substrate docking in AzgA. (A) Hypoxanthine, (B) guanine, (C and D) adenine. Amino acid residues implicated in substrate binding are shown. Hydrogen bonds are depicted with dashed lines.

and Asn131, through strong H-bonds (Table S1). We should note that a hydrogen bond between the side-chains of Thr390 and Asp339 is always present in the course of 200ns of simulation time, stabilizing the conformation of Asp339 to face always the binding pocket. The putative conformation transition from the inward-facing occluded conformation to an open conformation has been captured in Video S2, but is not validated experimentally. In contrast to hypoxanthine, adenine and guanine are displaced from their initial position in the course of the trajectory. Adenine is initially H-bonded to Asn131, Asp339, Thr390 and Glu394 side-chains, as well as to Asn131, Ala132 and Tyr133 backbone (Table S2), but approximately in the middle of the trajectory, it moves towards TM helices 5, 8 and 14 of the gate domain, where it is immobilized via H-bonds with Asp339, Glu394 and Asp478 side-chains. Similarly, at the first 150 ns guanine is constantly at the major putative substrate binding pocket through stabilizing H-bonds with Asn131, Asp339, Thr390, Glu394 side-chains and the backbone of Asn131 (Table S3). In the last 50 ns the ligand moves lower between TM helices 5, 8 and 14, and stays there through H-bonds with Asp339, Asp342 side-chains and Phe335 backbone.

Apart from the H-bonds implicated in ligand binding we also identified a few H-bonds that appear to stabilize the tertiary structure of the transporter (Table S4). These include H-bonds between Ser443 side-chain of TMS12 and Val263 main chain of TMSs 6–7 connecting loop, Phe392 main chain from the beta strand of TMS10 and Gly127 from the beta strand of TMS3, Asn60 side-chain from TM helix 1 and the Asn131 side-chain from TMS3, Asn62 side-chain from TM helix 1 and Phe426 main chain of TMSs 11–12 connecting loop, Ser65 side-chain of TMSs 1–2 connecting loop and Tyr202 of TMS5, and finally,

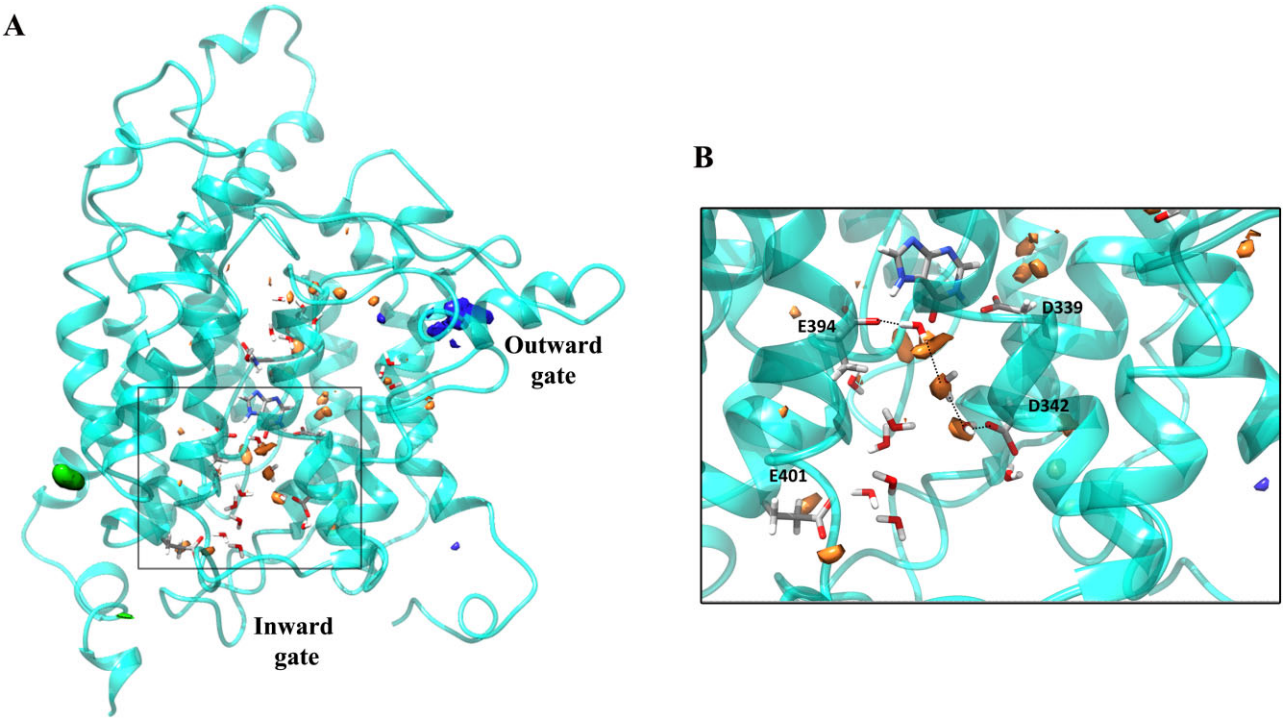


Fig. 4. Computational solvation analysis and the H^+ Trajectory.
A. Overview of the average co-ordinates of AzgA in complex with hypoxanthine.
B. Detailed view of the H-bond network on the substrate binding region. Small blobs (orange coloured in the online version): water oxygen density; Na^+ and Cl^- densities are also shown (green and blue coloured blobs in the online version) close to the start of TMS1 and in the putative outward gate respectively. Only low free energy water molecules are present as calculated from SZAMP software. The putative trajectory of the H^+ has been marked with black dotted lines.

Arg32 side-chain of N-terminus loop and Ala351 of TM helix 8. The importance of these residues in AzgA turnover is not examined by mutational and functional assays in this work.

Water molecule analysis and H^+ trajectory

The binding mode of a substrate can be examined using docking calculations and molecular dynamics simulations, but the H^+ trajectory is, in principle, very difficult to be simulated. Previous hypotheses assume that residues like Asp, Glu and His are employed in interactions with H^+ (Abramson *et al.*, 2003; 2004). We considered that water molecules inside the transmembrane surface of the transporter might be also involved in the H^+ translocation. Consequently, mapping the solvation of the transporter interior would result to models of a hypothetical proton pathway through the protonation of both amino acid residue side-chains and low-mobility water molecules.

To evaluate our hypothesis, we used two different approaches to analyse the existence of water molecules inside the transporter. First water oxygen, Na^+ and Cl^- densities were calculated from all trajectories. Second, we calculated the average structure from the simulation run and we used the SZMAP application (OpenEye Inc.) to fully investigate possible water accessible surfaces inside the AzgA structure. From the resulted calculations we selected to examine all water molecules having free energy estimations less than -1.5 kcal mol^{-1}, near the binding cavity (within 12 Å). The results from this consensus water molecule analysis are depicted in Fig. 4A, B. In Fig. 4B it is clear that both approaches pointed three regions near the binding cavity, with high water density, creating an H-bond network connecting Asp342 with Glu394 through three water molecules. Similar calculations were run for adenine and guanine substrates depicting hydrogen bond network among Asp339-Glu394-Asp342 (Fig. S5).

Rational design and construction of AzgA mutants

A relative comparison of the modelled topology of AzgA with the crystal structure of UraA bound with uracil (Lu *et al.*, 2011) and the modelled structures of other NATs (UapA and XanQ) (Karena and Frillingos, 2011; Kosti *et al.*, 2012), as well as, the substrate docking positions shown in Fig. 3, led us to the identification of a limited number of AzgA residues that might be critical for purine binding and transport. The functional role of these residues, as well as, of selected conserved residues in AzgA sequence alignments, was tested through a rational mutational analysis.

The 24 mutations constructed concerned 14 residues located in TMS1 (Thr49), TMS3 (Gly129, Asn131), TMS8 (Phe335, Asp339, Asp342, Thr346), TMS10 (Thr390, Glu394, Ser395, Gly398, Gly402, Lys404) and in the outward-facing limit of TMS14 (Asp478). The rationale for the specific substitutions made was to introduce either Ala residues or isofunctional/isostructural residues. All mutations were constructed in vectors carrying the *azgA* gene, fused in frame with the green fluorescent protein orf, and expressed from the *azgA* native promoter. We have previously shown that C-terminal translational fusion of AzgA with GFP is fully functional and shows normal subcellular trafficking and plasma membrane localization (Pantazopoulou *et al*., 2007). All plasmids were introduced in the mutant strain *azgA*Δ *uapA*Δ *uapC*Δ, which carries total deletions in the genes encoding all major endogenous purine transporters (for more details see *Experimental procedures*).

Phenotypic analysis of AzgA mutants

We analysed and selected transformants, arising from single-copy integration events, on different purines as sole nitrogen sources, together with positive (*azgA*$^{+}$) or negative (*azgA*Δ) isogenic control strains. A summary of the analysis of the selected transformants is shown in Fig. 5A. The positive control strain expressing wild-type AzgA–GFP (*azgA*$^{+}$) grows on adenine or hypoxanthine as sole nitrogen sources, and is sensitive to the toxic analogue 8-azaguanine. The negative control strain *azgA*Δ *uapA*Δ *uapC*Δ (*azgA*Δ) shows leaky growth on adenine or hypoxanthine, due to the presence of the minor purine transporter FcyB (Vlanti and Diallinas, 2008; Krypotou *et al*., 2012), and is resistant to 8-azaguanine. As expected, all strains showed normal growth on nitrogen sources unrelated to purines (e.g. nitrate). None among the AzgA mutant alleles conferred a capacity for growth on purines, such as uric acid or xanthine, or purine metabolic derivatives (e.g. allantoin), which are not physiological substrates for the wild-type AzgA (not shown). Several of the mutants showed reduced growth phenotypes on adenine or hypoxanthine or increased resistance to toxic 8-azaguanine, compatible with diminished AzgA transport capacity. In fact, AzgA mutants could be classified into two major groups; those showing growth profiles similar to the control strain expressing wild-type AzgA (T49A, T49S, G129A, N131A, N131D, F335A, D342E, T346A, E394D, S395A, S395N, G398A, K404R, D487A) and those with significantly reduced or apparently lost AzgA transport activity (G129F, D339A, D339N, D339E, D342A, D342N, T390A, E394A, E394Q, G402A). Some AzgA mutants seemed to grow slightly better on hypoxanthine than adenine (D342E or S395A) or *vice versa* (E394Q or G402A). In summary, growth testing suggested that at least residues Gly129 in TMS3, Asp339 and Asp342 in TMS8, and Thr390, Glu394 and Gly402 in TMS10, are critical for apparent AzgA activity. In fact, residue Asp339 seemed irreplaceable for function, as all substitutions made, including its replacement with Asn or Glu, led to lack of growth on purines.

Apparent purine transport rates in AzgA mutants

We performed comparative radiolabelled hypoxanthine uptake measurements (*V*, initial transport rates) in mutant strains shown in Fig. 5B. Our results were in good agreement with growth tests shown in Fig. 5A. Most AzgA mutants characterized by growth tests as reduced or loss-of-function mutants, had reduced or no AzgA-mediated transport rates for hypoxanthine ($V < 20\%$ compared with wild-type AzgA), whereas most AzgA mutants with no significant defect for growth on adenine or hypoxanthine, had hypoxanthine transport rates of > 30% of the wild-type AzgA. This observation is in line with previous experience with *A. nidulans* transporters, where only when a mutation reduces transport rates to < 30–40% of the wild-type, the defect is reflected in relevant growth tests (Papageorgiou *et al*., 2008; Krypotou *et al*., 2012). The only apparent disagreement between growth tests and hypoxanthine uptake rates concerned mutants N131A and N131D, which had significantly reduced transport rates, but could still grow well on hypoxanthine. But even in this case, when we measured steady state accumulation of hypoxanthine, rather than initial uptake rate, we found out that these mutants were eventually capable of 30–60% hypoxanthine accumulation compared with wild-type, explaining why they grow on this purine as sole nitrogen source (results not shown).

Expression levels and plasma membrane localization of mutant versions of AzgA

Apparently reduced AzgA uptake rates could imply problematic trafficking towards the plasma membrane, high turnover or *bona fidae* reduction in AzgA transport activity. To distinguish among these possibilities we performed western blots and epifluorescence microscopic analyses of all AzgA mutants. Figure 5C (see also Table 2) shows that in most mutants AzgA protein steady levels are similar or little modified (up to ~ 2-fold) compared with the wild-type control strain. A few mutants show more significant increases (3- to 5-fold) in protein levels (D339A, D339E and E394D). Significantly reduced AzgA protein levels (from > 4-fold and up to the limit of detection) were detected in T49A, T49S, G129A, N131A, F335A and T390A. Epifluorescence microscopic analysis (see *Experimental procedures*) further revealed that several of the AzgA mutants (N131D, D342A, D342N, E394A, E394Q and G402A) exhibited partial ER-retention or association with vacuoles

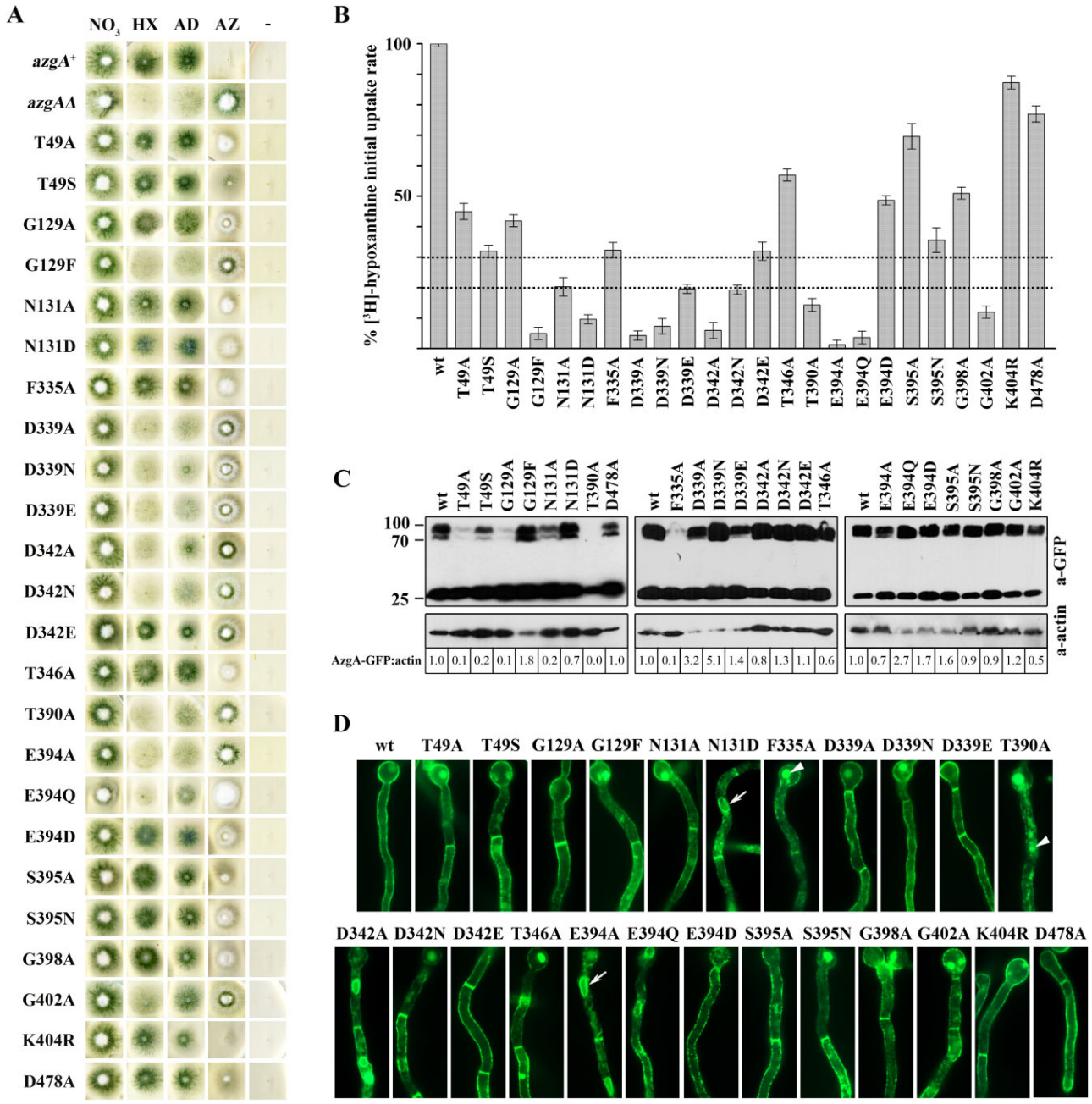


Fig. 5. Functional analysis of AzgA mutants.

A. Phenotypic analysis of AzgA mutants. Growth tests on nitrate (NO_3^-), hypoxanthine (HX) or adenine (AD) as sole nitrogen sources, resistance/sensitivity test on 8-azaguanine (AZ) and without any nitrogen source (–), at 32°C. Positive (*azgA*$^+$) and negative (*azgA*Δ) isogenic strains are shown as isogenic controls.

B. Comparative initial uptake rates of [^{3}H]-radiolabelled hypoxanthine in AzgA mutant alleles and a wild type (wt) (*azgA*$^+$) control. 100% is the transport rate in the wt.

C. Western blot analysis of total proteins from AzgA–GFP mutants and a wt control detected with anti-GFP antibody. Antibody against actin was used as an internal marker for equal loading. Molecular weight markers in kDa are shown. Relative amounts of AzgA–GFP, normalized by using actin protein levels, are shown below the blots (see *Experimental procedures*).

D. Epifluorescence microscopy showing *in vivo* subcellular expression of AzgA–GFP mutant alleles and a wt control. In selected samples, arrows and arrow heads depict perinuclear ER membrane and vacuoles respectively. The identity of such perinuclear rings as ER membranes has been previously rigorously established, and is a picture often obtained when partially misfolded versions of GFP-tagged transporters are expressed (Vlanti *et al.*, 2006; Gournas *et al.*, 2010; Krypotou *et al.*, 2012).

Table 1. Kinetic profile of AzgA mutants.

	K_m HX (μM)	K_i AD (μM)	K_i GU (μM)
Wt	2.8	3	1.3
T49A	1.8	2	2.5
T49S	1.6	0.7	0.4
G129A	0.3	0.3	0.4
G129F	n.m.	n.m.	n.m.
N131A	1	16	2
N131D	2	1	4
F335A	1.6	0.7	1.1
D339A	0.3	0.2	0.5
D339N	0.2	< 0.2	< 0.2
D339E	< 0.2	0.3	0.2
D342A	3.1	3.9	1
D342N	0.8	1.5	3
D342E	3.7	16	30
T346A	2	3.2	1.3
T390A	3.7	1.3	2.8
E394A	n.m.	n.m.	n.m.
E394Q	n.m.	n.m.	n.m.
E394D	1.5	16	3.6
S395A	2.3	3.4	2.8
S395N	1	1	0.4
G398A	2.1	3.3	1.4
G402A	3.1	1.5	0.7
K404R	4.2	2.0	2.1
D478A	4	3.2	4.4

$K_{m/i}$ values (μM) were determined as described in *Experimental procedures*. Results are averages of at least three independent experiments in triplicates for each concentration point. Standard deviation was < 20% (n.m., non-measurable). HX, AD, GU stand for hypoxanthine, adenine and guanine respectively.

and other cytoplasmic foci (Fig. 5D). In all cases showing significant ER-retention or increased turnover, which are probably due to partial misfolding, we cannot conclude whether the relevant mutation has also a direct effect on transport activity. On the other hand, most mutants showed AzgA protein levels and plasma membrane topology similar to wild-type and thus could provide more rigorous information on the role of the relevant residues mutated (see below).

Kinetic and specificity profile of AzgA mutants

K_m values for hypoxanthine and K_i values for adenine and guanine were measured for all mutants showing detectable AzgA transport activities, as described in *Experimental procedures*. Table 1 summarizes the results obtained. Most mutations did not alter significantly the K_m or K_i values for hypoxanthine, adenine or guanine, and thus are considered as mutations that do not grossly affect AzgA folding or the ability to bind substrates. Reduced hypoxanthine transport rates or growth on purines in several of these mutants, was most often associated with reduced plasma membrane localization and/or increased turnover (see Fig. 5). Thus, no further conclusion concerning the functional importance of the relevant residues could be drawn from these specific mutants.

On the other hand, prominent modifications of $K_{m/i}$ values or loss of transport activity were observed for some other AzgA mutants, which showed normal AzgA expression and localization in the plasma membrane. The first concerns mutation G129F which leads to loss of transport activity, despite increased protein levels and normal plasma membrane localization. In addition, mutation G129A, which despite its very low protein expression allows measurement of $K_{m/i}$ values, shows a very significant increase (up to 10-fold) in substrate affinities for all AzgA substrates (see Fig. 5 and Table 1). Thus, we can conclude that Gly129 is a residue essential for substrate binding and transport. A second functionally critical residue is Asp339, where all substitutions tested do not affect expression in the plasma membrane but lead to total loss of transport activity and, surprisingly, significantly increased $K_{m/i}$ values for all substrates. The simplest explanation of these results is that Asp339 is a residue involved directly in interactions with substrates and that its replacement traps substrates in the binding pocket. Similar results were obtained with conserved mutations replacing Glu356, a residue shown to bind substrates, in UapA (Papageorgiou *et al.*, 2008; Kosti *et al.*, 2012). A third AzgA residue critical for substrate binding and transport is Glu394. Although mutants E394A and E394Q show loss of apparent transport activity, most probably associated with misfolding and dramatically reduced localization in the plasma membrane (see Fig. 5), the conserved substitution E394D is functional and leads to distinctly modified binding affinities for different purines. More specifically, hypoxanthine and guanine binding is similar to wild-type, while adenine binding is 5.3-fold reduced. Finally, specific mutations in two other residues, Asn131 and Asp342, also seem to affect differentially the binding of purines. Mutants N131A and D342E, which are both normally expressed in the plasma membrane (see Fig. 5), reduce the binding of adenine 5.3-fold, while D342E also reduces guanine binding 23-fold.

The above observations suggest, independently from modelling and docking approaches, that residues Gly129 and Asn131 in TMS3, Asp339 and Asp342 in TMS8, and Glu394 in TMS10, are involved in substrate binding/or transport (see Table 2). Importantly, modelling and docking approaches strengthened the conclusions originating from functional studies, as discussed below.

Discussion

In the first part of our work we presented a theoretical structural model of the AzgA purine transporter, which shows 14 transmembrane, mostly helical, segments (TMSs), arranged as a compact core (TMSs 1–4 and 8–11) and a flexible gate (TMSs 5–7 and 12–14) domain. A putative binding pocket was located between the core and the gate domains. This model and independent docking

Table 2. Summary of AzgA mutants functional analysis.

Allele	TMS	Hypoxanthine growth	V (%)	Effect on substrate/H^+ binding and translocation	ER retention/ Traffic defect	Protein (%)
azgA$^+$		+++	100		No	100
T49A	1	++	46	–	No	10
T49S	1	++	32	–	No	20
G129A	3	++	44	+*	No	10
G129F	3	–	3	+*	No	180
N131A	3	+	23	+	No	20
N131D	3	+	10	+	Yes	70
F335A	8	++	35	–	No	10
D339A	8	–	3	+	No	320
D339N	8	–	5	+	No	510
D339E	8	–	21	+	No	140
D342A	8	–	7	+	Yes	80
D342N	8	–	19	+	Yes	130
D342E	8	++	32	+	No	110
T346A	8	+++	57	–	No	60
T390A	10	–	16	–	No	0
E394A	10	–	0	+	Yes	70
E394Q	10	–	0	+	Yes	270
E394D	10	++	49	+	No	170
S395A	10	+++	71	–	No	160
S395N	10	++	40	–	No	90
G398A	10	++	53	–	No	90
G402A	10	–	12	–	Yes	120
K404R	10	+++	88	–	No	50
D478A	14	+++	75	–	No	100

The TMS location of each mutant is indicated. V is initial uptake rate. Protein levels are estimated by ImageJ using the relative intensities of AzgA- and Actin-specific bands shown in Fig. 5C. The asterisk ' *' stands for an indirect effect. For more details see text.

approaches and MD simulations revealed a number of important aspects concerning how this transporter might bind and transport its substrates. In particular, residues Asn131 (TMS3), Asp339 (TMS8), Glu394 (TMS10) and the amide group of the backbone of Val393 (TMS10), were proposed to be the major residues interacting, through H bonds with all substrates. These interactions implicate N1(H) and C6=O (hypoxanthine, guanine) or C6-NH_2 (adenine) of the pyrimidine moiety and N7(H) and N9(H) of the imidazole moiety. In addition, MD approaches suggested that a hydrogen bond network among Asp339-Glu394-Asp342 might be involved in H^+ binding and substrate translocation, thus adding one more residue, Asp342, in the set of amino acids absolutely critical for AzgA function.

In the second part of the work we experimentally tested the functional importance of several residues by systematic mutational analysis. The residues selected were chosen on two basic criteria: evolutionary conservation in AzgA sequences and topological position in the modelled structure of AzgA in respect to substrate and/or H^+ binding. Our results showed that Asn131, Asp339, Asp342 and Glu394, which are the four major residues proposed to be involved in direct interactions with substrates and/or H^+, were found to be irreplaceable for physiological AzgA transport activity. More importantly, specific highly conserved substitutions of these residues lead to significantly modified affinity constants for substrates. Specific substitutions in three other residues, Gly129, Thr390 and Gly402 were also found to dramatically reduce or abolish AzgA-mediated purine transport. All these residues are in close proximity (H bonding distance) to the residues proposed to be involved in direct substrate and/or H^+ binding. Importantly, our conclusions in respect to the functional role of these residues are based on mutations that do not modify the expression and localization of AzgA in the plasma membrane. A summary of the functional analysis of AzgA mutants is presented in Table 2.

The present structural models are in very good agreement with previous transport competition experiments using purine analogues with variably modified positions of the purine ring (Goudela *et al.*, 2006). For example, the importance of positions N1(H) and C6=O/C6-NH_2 in binding is supported by the significantly reduced affinity of AzgA for purine (K_i 99) and especially 1-deazapurine (K_i 589). The importance of positions N7 and N9 is also supported by the non-binding of pyrimidines and 7-deazaguanine, or the reduced affinity (K_i 339) for

9-methylguanine. In all three cases protein interacts with the substrate through a formation of 3 hydrogen bonds: Glu394 side-chain as an acceptor, Val393 amide backbone as a donor and Asp339 side-chain as an acceptor. A similarity between guanine and hypoxanthine is apparent from the docking results where the corresponding scheme for hydrogen bonding is N9(H) as a donor, C6=O as an acceptor and N1(H) as a donor. Adenine shows a complete different scheme with N(9)H as a donor, N3 as an acceptor and NH_2 at position 6 as a donor. Moreover Asn131 appears to be always a donor to N7 or N3; this should be important and detrimental for the affinity against xanthine and uric acid, as in the relevant position the nitrogen is always protonated. Finally the non-binding of 3-methyladenine, supports the model of adenine binding depicted in Fig. 3C, rather than the one shown in Fig. 3D.

The MD simulations suggest a possible mechanism of purine translocation. Although the substrate binding cavity remains open throughout most of the trajectory, the strong H-bond network that keeps hypoxanthine/adenine/guanine at that position is unperturbed, which further suggests that a more drastic change is necessary for triggering the release of substrates into the cytosol. The driving force may be a protonation of the side-chain of one of the Asp339 or/and Glu394, which form constant H-bonds with all three ligands. The importance of Asp339 was also supported by the replacement of Thr390 by alanine, as MD simulations revealed that Thr390 is stabilizing the conformation of Asp339 side-chain towards the binding pocket. In the T390A mutant $K_{m/i}$ values were not changed although the apparent initial uptake rate of ^{3}H-radiolabelled hypoxanthine dropped to 16% of the wild-type levels, although this seems to be due to an increase in the turnover of the protein. Another interesting observation made from the MD trajectories is the increased water oxygen density at contiguous positions, which traces a pathway that starts from a putative outward-facing secondary gate and ends at the bottom of the transporter. These 'structural' water molecules may serve as a bridge to the proton that is believed to be symported with the ligand (Johansson *et al*., 2011).

Interestingly, increased Na^+ ion density in the hypoxanthine trajectory was only detected at the brink of a putative outward-facing gate. In that case, the ligand did not move from its initial position and no cations were constantly present in the interior of the transporter. In contrast, in adenine and guanine trajectories, where the ligands moved downwards, increased Na^+ density was observed between the Asp339 and Glu394, directly beneath the ligand. From the solvation analysis, the position and the orientation of water molecules inside the binding cavity could indicate a hypothetical trajectory of H^+. A theoretical trajectory pathway can be visualized in Fig. 4B. Three low-mobility water molecules connecting Asp342 and Glu394, are placed beneath the binding pocket and could be considered as transition states of H^+ trajectory, which triggers the final re-orientation of the transporter in order to facilitate the movement of the substrate inside the cell. In that case the role of H^+ is to neutralize progressively Asp342 and Glu394 in order to facilitate the movement of the substrate inside the cytosol. Mutations in Asp342 and Glu394 affect AzgA-mediated transport and reduce cell growth on purines, a phenomenon probably related to substrate/H^+ flow in agreement with our solvation analysis.

When this manuscript was in preparation, K. Papakostas, M. Botou and S. Frillingos published an article on the functional identification of four AzgA-like purine transporters in *E. coli*, namely PurP, YicO, YjcD and YgfQ (Papakostas *et al*., 2013). The first two were shown to be specific for adenine, whereas the other two for hypoxanthine and guanine. The work of Papakostas *et al*. (2013) also shows that residues irreplaceable for purine uptake activity or crucial for substrate selectivity in the *E. coli* AzgA-like transporters are identical or very similar to the functionally residues identified herein for *A. nidulans* AzgA (Fig. S1). In particular, they have found that an Asp and a Glu residue in TMS8 and TMS10, respectively, are absolutely essential for the function of the *E. coli* AzgA-like transporters, as we also showed for Asp339 and Glu394 in AzgA. On the other hand, the observation that *E. coli* AzgA-like homologues show a specificity preference for adenine or hypoxanthine-guanine, not shown by fungal or plant homologues could be instructing on which amino acid residues are important for conferring substrate promiscuity *versus* increased specificity. For example, among the residues shown to be critical for substrate binding by AzgA, Gly129, Asn131, Asp339 and Val393 are differentially conserved in the *E. coli* transporters, which recognize either adenine or hypoxanthine-guanine. More specifically, PurP has an Asn and Ile residues instead of Asp339 and Val393, whereas YjcD has Ser, Thr, Ala and Ile residues instead of Gly129, Asn131, Asp339 and Val393 respectively (see Fig. S1). The ease of direct genetic screens for altering the purine specificity of AzgA, as it was achieved for NATs (Amillis *et al*., 2001; Papageorgiou *et al*., 2008; Kosti *et al*., 2010), might constitute a powerful tool in uncovering the molecular elements affecting AzgA substrate/ligand recognition.

Interestingly, functionally important residues in bacterial and fungal AzgA transporters are located at topologically similar positions with residues which are also critical for the function of NAT homologues, such as UapA or XanQ. Thus, AzgA-like transporters were shown not only to fold similarly to NAT proteins, but also to use topologically similar TMSs and similar amino acid side-chains to dictate their function and specificity preferences. Does this observation by itself suggests that AzgA-like and NAT proteins, despite their very low overall primary amino acid sequence similarity, are phylogenetically related? Although the phylogenetic

analysis performed here does not give a rigorous answer to this question, the AzgA-like proteins clearly form a distinct transporter group which seems more related to the NAT group, compared with other purine transporter families (e.g. NCS1). In agreement with the idea that AzgA forms a functionally and evolutionary distinct group of transporters, all functionally critical residues of AzgA, identified in the present work and in Papakostas *et al.* (2013), are nearly absolutely conserved in the AzgA family, but none of these residues is conserved in the NAT or NCS1 families (Fig. S1). We can thus conclude that, although the AzgA and NAT proteins seem to share a common molecular ancestry, as reflected in similar overall and local topologies, they also seem to have significantly diverged in respect to their specificities by employing different polar side-chains for substrate binding and transport.

Experimental procedures

Model construction

Multiple threading alignments were created with LOMETS metaserver (Wu and Zhang, 2007) using the full-length sequence of AzgA from *A. nidulans* (Uniprot ID: Q7Z8R3). All alignments of AzgA with the sequence of the *E. coli* uracil/H^+ symporter UraA (PDB ID: 3QE7), as well as the consensus alignments of their combinations created with T-Coffee (Notredame *et al.*, 2000), were tested on the basis of model construction with Rosetta 3.4 (Leaver-Fay *et al.*, 2011). The final alignment used for production of AzgA structural models, shown in Fig. S2, was obtained by HHsearch (Söding, 2005) and modified manually on the basis of the multiple alignment of selected AzgA-like sequences shown in Fig. S3. The full-length sequence of AzgA from *A. nidulans* was submitted to TOPCONS meta-server (Bernsel *et al.*, 2009) for consensus membrane protein topology prediction. A total of 10257 models of AzgA residues 21–502 were constructed in total using the Rosetta threading protocol with the membrane specific scoring function terms (Yarov-Yarovoy *et al.*, 2006; Barth *et al.*, 2007; 2009). The secondary structure of the protein, as predicted by SAM server (Karplus, 2009), as well as the consensus membrane protein topology prediction made by TOPCONS meta-server, were also taken into account in the form of restraints during model construction. 3-mer and 9-mer peptide fragments for modelling were created locally using the SAM-predicted secondary structure of AzgA. Unfolded and broken structures were excluded by retaining only the 90% top-scores models (9230). A random model among them was superimposed on the crystal structure of UraA using the combinatorial extension (CE) algorithm (Shindyalov and Bourne, 1998), and the putative substrate binding cavity of AzgA was determined by selecting every residue that was poised within 8 Å from uracil ligand. A trajectory was assembled from the 9230 models and was clustered with GROMACS Tools v4.6.1 (Pronk *et al.*, 2013) using the backbone heavy atoms of the putative substrate binding pocket. The algorithm used for clustering was 'gromos' and the cut-off used was 1 Å, which produced 2095 clusters in total. Only 89 clusters encompassed 10 or more models and those were retained for further analysis.

Docking and rescoring

The medoid structures of the 89 largest clusters were superimposed on UraA X-ray structure using the backbone heavy atoms of the putative substrate binding cavity. Then a search volume box for docking with Vina (Trott and Olson, 2010) was defined centred at uracil with dimensions $17 \times 21 \times 21$ in Angstroms. Side-chains of residues within 7 Å of uracil were left flexible during docking of the 12 substrates for which accurate K_i values existed (Goudela *et al.*, 2006). The whole setup was conducted in Pymol with the Computer-aided drug design platform (Lill and Danielson, 2011). Subsequently all the poses generated with Vina were rescored with NNscore1 (Durrant and McCammon, 2010) and NNscore2 (Durrant and McCammon, 2011b) as well as DSX (Neudert and Klebe, 2011) programs. That procedure yielded 47 different scoring schemes for each of the 89 cluster medoid structures. Approximate free energies of binding were calculated from the experimental K_i values using the following formula:

$$E = -R * T * \ln\left(\frac{10^6}{Ki}\right)$$

where $R*T$ was set to 0.5957632602.

The 12 substrates were ranked according to their binding affinity using their docking scores, therefore 47 different rankings were produced (one for each scoring scheme) for each of the 89 cluster medoid structures. For each ranking we calculated the statistic termed top-down concordance (Iman and Conover, 1987; Teles, 2012), which quantifies the agreement between raters (in this case a particular scoring scheme and the experimental measurements), by emphasizing more on the lowest ranks (namely the most potent substrates).

$$C_T = \frac{1}{M^2 * (n - S_1)} * \left(\sum_{i=1}^{n} R_i^2 - M^2 * n\right)$$

where M is the number of groups of ratings (2 in our case); n the number of items ranked (12 in our case) and R_i is the sum of the Savage scores (Savage, 1956), across the M groups, at rank position i; S_1 is Savage score 1. For a given sample size n, the i^{th} Savage score is:

$$S_i = \sum_{j=1}^{n} \frac{1}{j}$$

Thus, for example, if $n = 4$, then S1 = 1/1 + 1/2 + 1/3 + 1/4 = 2.083, S2 = 1/2 + 1/3 + 1/4 = 1.083, S3 = 1/3 + 1/4 = 0.583, and S4 = 1/4 = 0.250. If there are tied ranks (two or more substrates with identical docking scores or estimated free energies), then we used the mean of the Savage scores for the positions of the tied data. For example, if $n = 4$ and ranks 2 and 3 were tied then we use (1.083 + 0.583)/2 = 0.833 for both S2 and S3.

The top-down concordance coefficient ranges between 0 and 1 (the higher the better) and its significance (P-value) can be found in the table for $N = 12$ (p. 23) in Iman (1987).

The validity of the cluster medoids was assessed according to their highest top-down concordance values and by visual

inspection of the docking poses with Binana (Durrant and McCammon, 2011a). The selected best docking poses of hypoxanthine, adenine and guanine were further subjected to fully flexible receptor docking calculation performed using the Low-Mode/Monte Carlo (LMCS) protocol implemented in Macromodel 9.9 (Suite 2012: MacroModel, version 9.9, Schrödinger, LLC, New York, NY, 2012). The LMCS protocol encompasses 1000 step search of the mixed Monte Carlo/Low Mode (MC/LMOD) search algorithm with a ratio of 0.3 and OPLS-2005* force field. A distance-dependent dielectric 'constant' of 4*r* was used. All residues within 6.0 Å from the ligand were allowed to move freely, all residues within 4.0 Å from the freely moving atoms were treated as fixed using a 100 kJ mol^{-1} $Å^{-2}$ force constant and the remaining residues were treated as 'frozen atoms'. After each successful run the complex was minimized using the TNCG algorithm (rmsG < 0.01 kJ mol^{-1} $Å^{-1}$). Unique conformations were stored only if they were within the lowest 50 kJ mol^{-1}.

Molecular dynamic simulations and trajectory analysis

For the MD simulations Desmond v.3 software was implemented (Desmond Molecular Dynamics System, version 3.0, D. E. Shaw Research, NY). The system was prepared by embedding the protein in a POPC lipid bilayer, solvating the membrane by TIP3P explicit water, neutralizing with counter ions and adding 150 mM salt and subsequently following the stepwise equilibration protocol as developed by Desmond for membrane proteins. A 200 ns simulation was performed in the NPT ensemble with Langevin thermostat and barostat and semi isotropic pressure restraints for each of the three substrates studied (adenine, hypoxanthine and guanine). The H-bond frequencies were calculated with VMD 1.9 (Humphrey *et al*., 1996) using a distance cut-off of 3.5 Å and angle cut-off of 30°. The electrostatic potential distribution was calculated with APBSmem (Baker *et al*., 2001; Callenberg *et al*., 2010), upon Gasteiger charge assignment with PDB2PQR (Dolinsky *et al*., 2004; 2007). The dielectric constants for the protein and the lipid tails were set to 2, whereas for the solvent and lipid headgroups to 80. Images were created with PyMOL [The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC] and UCSF Chimera 1.8 (Pettersen *et al*., 2004), and the videos with UCSF Chimera and VMD 1.9.

Computational solvent mapping

Possible hydration sites of the protein–ligand complex were identified using both the water densities derived from the MD trajectories, as well as the SZMAP application. Average water, Na^+, Cl^- and ligand densities were calculated from the 200 ns trajectories with Amber Tools 13 (Case *et al*., 2012) using a grid spacing of 0.5 Å. The SZMAP protocol was the following: first *pch* application was run to assign partial charges and radii and split protein and ligand files. After, SZMAP application was run using stabilization protocol: calculations for *complex*, *apo* and *ligand* grids were run and then further processed to yield (*complex-apo-ligand)* stabilization grids, describing where water is stabilized or destabilized the binding reaction. For every grid point s*zmap* places one explicit water, then treats the rest as a continuum and samples the water orientations.

Media, strains, growth conditions and transformation genetics

Standard complete and minimal media (MM) for *A. nidulans* were used. Media and supplemented auxotrophies were at the concentrations given in http://www.fgsc.net. Nitrogen sources were used at the final concentrations: urea 5 mM, $NaNO_3$ 10 mM, purines 0.5 mM. *E. coli* was grown on Luria-Bertani medium. Media and chemical reagents were obtained from Sigma-Aldrich (Life Science Chemilab SA, Hellas) or AppliChem (Bioline Scientific SA, Hellas). Transformations were performed as described previously (Koukaki *et al*., 2003). A single-copy plasmid integration of the AzgA served as a standard wild type control. A *azgA*Δ *uapA*Δ *uapC*Δ*::AfpyrG argB2 pabaA1* mutant strain (Vlanti *et al*., 2006) was the recipient strain in transformations with the plasmids carrying the mutant AzgA alleles. As a negative control a transformant with an empty vector was used. These vectors allow selection of transformants based on arginine (*argB2*) auxotrophy complementation. Transformants expressing intact *azgA–gfp* alleles, through either single-copy or multicopy plasmid integration events, were identified by PCR and Southern analysis. Single copy transformants were selected for further work, and when necessary multicopy transformants were used for the kinetic analysis of the mutants with very low uptake rate. Growth tests were performed at 32°C, pH 6.8.

Plasmid constructions and AzgA mutations

The plasmid used is a modified pBluescript KS(+) vector (Stratagene) carrying the *azgA* gene fused C-terminally with the *gfp* gene, as well as, the *argB* gene as a selection marker. The *argB* gene was cloned in the SalI site of the pBluescript multicloning site and the *azgA* gene (orf and regulatory sequences) was amplified from *A. nidulans* genome locus and cloned in the BamHI position of the multicloning site. PCR mutagenesis was used to destroy the XbaI site of the multicloning site and replace the *azgA* stop codon with a XbaI restriction site. Finally, *gfp* was amplified with XbaI extensions and cloned in frame at the end of the *azgA* orf. This plasmid was used for the creation of the mutant alleles of *azgA*. Mutations were constructed by site-directed mutagenesis according to the instructions accompanying the Quik-Change® Site-Directed Mutagenesis Kit (Stratagene) on the above vectors and were confirmed by sequencing. Oligonucleotides used for cloning and site-directed mutagenesis purposes are listed in Table S5.

Standard nucleic acid manipulations

Genomic DNA extraction from *A. nidulans* was as described in http://www.fgsc.net. Plasmid preparation from *E. coli* strains and DNA bands were purified from agarose gels were done with the Nucleospin Plasmid kit and the Nucleospin ExtractII kit according to the manufacturer's instructions (Macherey-Nagel, Lab Supplies Scientific SA, Hellas). Southern blot analysis was performed as described in Sambrook *et al.* (1989). ^{32}P-dCTP labelled molecules used as *azgA or argB* specific probes were prepared using a random hexanucleotide primer kit following the supplier's instructions (Takara Bio, Lab

Supplies Scientific SA, Hellas) and purified on MicroSpin™ S-200 HR columns, following the supplier's instructions (Roche Diagnostics, Hellas). Labelled ^{32}P-dCTP (3000 Ci mmol^{-1}) was purchased from the Institute of Isotops Co. Ltd, Miklós, Hungary. Restriction enzymes were from Takara Bio (Lab Supplies Scientific SA, Hellas). Conventional PCR reactions were done with KAPATaq DNA polymerase (Kapa Biosystems, Lab Supplies Scientific SA, Hellas). Cloning and amplification of products were done with Kapa HiFi (Kapa Biosystems, Lab Supplies Scientific SA, Hellas.

Total protein extraction and western blot analysis

Cultures for total protein extraction were grown in MM supplemented with nitrate at 32°C for 14 h. Total protein extraction was performed as previously described (Karachaliou *et al*., 2013). Equal sample loading was estimated by Bradford assays and Coomasie staining. Total proteins (30 μg) were separated by SDS-PAGE (10% w/v polyacrylamide gel) and electroblotted (Mini PROTEAN™ Tetra Cell, BIORAD) onto PVDF membranes (Macherey-Nagel, Lab Supplies Scientific SA, Hellas) for immunodetection. The membrane was treated with 2% (w/v) non fat dry milk and immunodetection was performed with a primary mouse anti-GFP monoclonal antibody (Roche Diagnostics, Hellas) or a mouse anti-actin monoclonal (C4) antibody (MP Biomedicals Europe, Lab Supplies Scientific SA, Hellas) and a secondary goat anti-mouse IgG HRP-linked antibody (Cell Signaling Technology Inc., Bioline Scientific SA, Hellas). Blots were developed by the chemiluminescent method using the LumiSensor Chemiluminescent HRP Substrate kit (Genscript USA, Lab Supplies Scientific SA, Hellas) and SuperRX Fuji medical X-Ray films (FujiFILM Europe, Lab Supplies Scientific SA, Hellas). Densitometry analysis of AzgA–GFP and actin-specific bands was performed using the ImageJ software. The relative density of the AzgA–GFP double band was quantified and normalized by dividing with the respective actin-band density.

Kinetic analysis

[^{3}H]-hypoxanthine (19.6–33.4 Ci mmol^{-1}, Moravek Biochemicals, CA, USA) uptake in MM was assayed in germinating conidiospores of *A. nidulans* concentrated at 10^7 conidiospores/100 μl, at 32°C, pH 6.8, as previously described (Cecchetto *et al*., 2004; Papageorgiou *et al*., 2008). Initial velocities were measured at 1 min of incubation with concentrations of 0.2–2.0 μM of [^{3}H]-hypoxanthine at the polarity maintenance stage (3–4 h, 130 rpm). $K_{m/i}$ values were obtained directly by performing and analysing uptakes (Prism 3.02: GraphPad Software), using labelled hypoxanthine at 0.2–0.5 μM, or at various concentrations (0.5–2000 μM) of non-labelled substrates. K_i values were calculated by satisfying the criteria for use of the Cheng and Prusoff equation $K_i = IC_{50}/(1 + (L/K_m))$, in which L is the permeant concentration. IC_{50} values were determined from full dose–response curves and in all cases the Hill coefficient was close to –1, consistent with the presence of one binding site. Reactions were terminated with addition of equal volumes of ice-cold MM containing 1000-fold excess of non-radiolabelled substrate. Background uptake values were corrected by subtracting either values measured in the deleted mutants or values obtained in the simultaneous presence of 1000-fold excess of non-radiolabelled substrate. Both approaches led to the same background uptake level, not exceeding 10–15% of the total counts obtained in wild-type strains. All transport assays were carried out in at least three independent experiments, with three replicates for each concentration or time point. Standard deviation was < 20%.

Epifluorescence microscopy

Samples for fluorescence microscopy were prepared as previously described (Karachaliou *et al*., 2013). In brief, germlings grown in 35 mm Petri dish (WillCo-dish®, Series GWSt-3512) in liquid MM supplemented with $NaNO_3$ as nitrogen source for 12–14 h at 32°C, were observed on a Zeiss Axio Observer Z1 inverted epi-fluorescent microscope and the resulting images were acquired with an AxioCamMR3 digital camera using the Zen 2012 software. Image processing, contrast adjustment and colour combining were made using the Adobe Photoshop CS4 Extended version 11.0.2 software or the ImageJ software. Images were converted to RGB and annotated using Photoshop CS4 before being saved to TIFF.

Phylogenetic analysis

The AzgA protein sequence was used as an *in silico* probe to search, using standard blastp, for similar sequences in the non-redundant protein sequence database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and in specialized fungal databases (http://www.aspgd.org/; http://genome.jgi.doe.gov/programs/fungi/index.jsf). Multiple sequence alignments were built at either http://multalin.toulouse.inra.fr/multalin/ or http://www.ebi.ac.uk/Tools/msa/muscle/. Phylogenetic analysis and evolutionary tree construction was made at http://www.phylogeny.fr/, using PhyML 3.0 and TreeDyn 198.3 respectively (Dereeper *et al*., 2008).

Acknowledgements

We thank Prof. S. Efthimiopoulos for microscopic facilities and Dr S. Amillis for the original AzgA plasmid. T.E. acknowledges funding from Bodossaki Foundation. The authors acknowledge financial support from the European Commission under the Seventh Framework Program by means of the grant agreement for the Integrated Infrastructure Initiative No. 262348 European Soft Matter Infrastructure (ESMI). This work was also supported by the Cy-Tera Project (NEA YPODOMI/STRATI/0308/31), which is co-funded by the European Regional Development Fund and the Republic of Cyprus through the Research Promotion Foundation.

References

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., and Iwata, S. (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301:** 610–615.

Abramson, J., Iwata, S., and Kaback, H.R. (2004) Lactose permease as a paradigm for membrane transport proteins (Review). *Mol Membr Biol* **21:** 227–236.

Amillis, S., Koukaki, M., and Diallinas, G. (2001) Substitution F569S converts UapA, a specific uric acid-xanthine transporter, into a broad specificity transporter for purine-related solutes. *J Mol Biol* **313:** 765–774.

Amillis, S., Cecchetto, G., Sophianopoulou, V., Koukaki, M., Scazzocchio, C., and Diallinas, G. (2004) Transcription of purine transporter genes is activated during the isotropic growth phase of *Aspergillus nidulans* conidia. *Mol Microbiol* **52:** 205–216.

Amillis, S., Kosti, V., Pantazopoulou, A., Mikros, E., and Diallinas, G. (2011) Mutational analysis and modeling reveal functionally critical residues in transmembrane segments 1 and 3 of the UapA transporter. *J Mol Biol* **411:** 567–580.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* **98:** 10037–10041.

Barrett, M.P., and Gilbert, I.H. (2006) Targeting of toxic compounds to the trypanosome's interior. *Adv Parasitol* **63:** 125–183.

Barth, P., Schonbrun, J., and Baker, D. (2007) Toward high-resolution prediction and design of transmembrane helical protein structures. *Proc Natl Acad Sci USA* **104:** 15682–15687.

Barth, P., Wallner, B., and Baker, D. (2009) Prediction of membrane protein structures with complex topologies using limited constraints. *Proc Natl Acad Sci USA* **106:** 1409–1414.

Bernsel, A., Viklund, H., Hennerdal, A., and Elofsson, A. (2009) TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Res* **37:** W465–W468.

Callenberg, K.M., Choudhary, O.P., Forest, G.L., de Gohara, D.W., Baker, N.A., and Grabe, M. (2010) APBSmem: a graphical interface for electrostatic calculations at the membrane. *PLoS ONE* **5:** e12722.

Case, D.A., Darden, T.A., Cheatham, T.E., Simmerling, C.L., Wang, J., Duke, R.E., *et al.* (2012) *AMBER 12*. San Francisco, CA: University of California.

Cecchetto, G., Amillis, S., Diallinas, G., Scazzocchio, C., and Drevet, C. (2004) The AzgA purine transporter of *Aspergillus nidulans*. Characterization of a protein belonging to a new phylogenetic cluster. *J Biol Chem* **279:** 3132–3141.

Darlington, A.J., and Scazzocchio, C. (1967) Use of analogues and the substrate-sensitivity of mutants in analysis of purine uptake and breakdown in *Aspergillus nidulans*. *J Bacteriol* **93:** 937–940.

Darlington, A.J., Scazzocchio, C., and Pateman, J.A. (1965) Biochemical and genetical studies of purine breakdown in *Aspergillus*. *Nature* **206:** 599–600.

De Koning, H., and Diallinas, G. (2000) Nucleobase transporters (Review). *Mol Membr Biol* **17:** 75–94.

De Koning, H.P., Bridges, D.J., and Burchmore, R.J. (2005) Purine and pyrimidine transport in pathogenic protozoa: from biology to therapy. *FEMS Microbiol Rev* **29:** 987–1020.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., *et al.* (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36:** W465–W469.

Diallinas, G., and Gournas, C. (2008) Structure-function relationships in the nucleobase-ascorbate transporter (NAT) family: lessons from model microbial genetic systems. *Channels* **2:** 363–372.

Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Res* **32:** W665–W667.

Dolinsky, T.J., Czodrowski, P., Li, H., Nielsen, J.E., Jensen, J.H., Klebe, G., and Baker, N.A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res* **35:** W522–W525.

Durrant, J.D., and McCammon, J.A. (2010) NNScore: a neural-network-based scoring function for the characterization of protein–ligand complexes. *J Chem Inf Model* **50:** 1865–1871.

Durrant, J.D., and McCammon, J.A. (2011a) BINANA: a novel algorithm for ligand-binding characterization. *J Mol Graph Model* **29:** 888–893.

Durrant, J.D., and McCammon, J.A. (2011b) NNScore 2.0: a neural-network receptor-ligand scoring function. *J Chem Inf Model* **51:** 2897–2903.

Frillingos, S. (2012) Insights to the evolution of nucleobase-ascorbate transporters (NAT/NCS2 family) from the Cys-scanning analysis of xanthine permease XanQ. *Int J Biochem Mol Biol* **3:** 250–272.

Goudela, S., Tsilivi, H., and Diallinas, G. (2006) Comparative kinetic analysis of AzgA and Fcy21p, prototypes of the two major fungal hypoxanthine-adenine-guanine transporter families. *Mol Membr Biol* **23:** 291–303.

Goudela, S., Reichard, U., Amillis, S., and Diallinas, G. (2008) Characterization and kinetics of the major purine transporters in *Aspergillus fumigatus*. *Fungal Genet Biol* **45:** 459–472.

Gournas, C., Papageorgiou, I., and Diallinas, G. (2008) The nucleobase-ascorbate transporter (NAT) family: genomics, evolution, structure-function relationships and physiological role. *Mol Biosyst* **4:** 404–416.

Gournas, C., Amillis, S., Vlanti, A., and Diallinas, G. (2010) Transport-dependent endocytosis and turnover of a uric acid-xanthine permease. *Mol Microbiol* **75:** 246–260.

Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph* **14:** 33–38, 27–28.

Iman, R.L. (1987) Tables of the exact quantiles of the top-down correlation coefficient for n – 3(1)14. *Commun Stat Theory Methods* **16:** 1513–1540.

Iman, R.L., and Conover, W.J. (1987) A measure of top-down correlation. *Technometrics* **29:** 351.

Jain, S., Sutchu, S., Rosa, P.A., Byram, R., and Jewett, M.W. (2012) *Borrelia burgdorferi* harbors a transport system essential for purine salvage and mammalian infection. *Infect Immun* **80:** 3086–3093.

Johansson, A.L., Chakrabarty, S., Berthold, C.L., Högbom, M., Warshel, A., and Brzezinsk, P. (2011) Proton-transport mechanisms in cytochrome c oxidase revealed by studies of kinetic isotope effects. *Biochim Biophys Acta* **1807:** 1083–1094.

Karachaliou, M., Amillis, S., Evangelinos, M., Kokotos, A.C.,

Yalelis, V., and Diallinas, G. (2013) The arrestin-like protein ArtA is essential for ubiquitination and endocytosis of the UapA transporter in response to both broad-range and specific signals. *Mol Microbiol* **88:** 301–317.
Karena, E., and Frillingos, S. (2011) The role of transmembrane segment TM3 in the xanthine permease XanQ of *Escherichia coli*. *J Biol Chem* **286:** 39595–39605.
Karplus, K. (2009) SAM-T08, HMM-based protein structure prediction. *Nucleic Acids Res* **37:** W492–W497.
Kosti, V., Papageorgiou, I., and Diallinas, G. (2010) Dynamic elements at both cytoplasmically and extracellularly facing sides of the UapA transporter selectively control the accessibility of substrates to their translocation pathway. *J Mol Biol* **397:** 1132–1143.
Kosti, V., Lambrinidis, G., Myrianthopoulos, V., Diallinas, G., and Mikros, E. (2012) Identification of the substrate recognition and transport pathway in a eukaryotic member of the nucleobase-ascorbate transporter (NAT) family. *PLoS ONE* **7:** e41939.
Koukaki, M., Giannoutsou, E., Karagouni, A., and Diallinas, G. (2003) A novel improved method for *Aspergillus nidulans* transformation. *J Microbiol Methods* **55:** 687–695.
Köse, M., and Schiedel, A.C. (2009) Nucleoside/nucleobase transporters: drug targets of the future? *Future Med Chem* **1:** 303–326.
Kraupp, M., and Marz, R. (1995) Membrane transport of nucleobases: interaction with inhibitors. *Gen Pharmacol* **26:** 1185–1190.
Krypotou, E., Kosti, V., Amillis, S., Myrianthopoulos, V., Mikros, E., and Diallinas, G. (2012) Modeling, substrate docking, and mutational analysis identify residues essential for the function and specificity of a eukaryotic purine-cytosine NCS1 transporter. *J Biol Chem* **287:** 36792–36803.
Leaver-Fay, A., Tyka, M., Lewis, S.M., Lange, O.F., Thompson, J., Jacak, R., *et al*. (2011) ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* **487:** 545–574.
Lill, M.A., and Danielson, M.L. (2011) Computer-aided drug design platform using PyMOL. *J Comput Aided Mol Des* **25:** 13–19.
Lu, F., Li, S., Jiang, Y., Jiang, J., Fan, H., Lu, G., *et al*. (2011) Structure and mechanism of the uracil transporter UraA. *Nature* **472:** 243–246.
Mansfield, T.A., Schultes, N.P., and Mourad, G.S. (2009) AtAzg1 and AtAzg2 comprise a novel family of purine transporters in *Arabidopsis*. *FEBS Lett* **583:** 481–486.
Mourad, G.S., Snook, B.M., Prabhakar, J.T., Mansfield, T.A., and Schultes, N.P. (2006) A fluoroorotic acid-resistant mutant of *Arabidopsis* defective in the uptake of uracil. *J Exp Bot* **57:** 3563–3573.
Neudert, G., and Klebe, G. (2011) DSX: a knowledge-based scoring function for the assessment of protein–ligand complexes. *J Chem Inf Model* **51:** 2731–2745.
Notredame, C., Higgins, D.G., and Heringa, J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302:** 205–217.
Pantazopoulou, A., and Diallinas, G. (2007) Fungal nucleobase transporters. *FEMS Microbiol Rev* **31:** 657–675.
Pantazopoulou, A., Lemuh, N.D., Hatzinikolaou, D.G., Drevet, C., Cecchetto, G., Scazzocchio, C., and Diallinas, G. (2007) Differential physiological and developmental expression of the UapA and AzgA purine transporters in *Aspergillus nidulans*. *Fungal Genet Biol* **44:** 627–640.
Papageorgiou, I., Gournas, C., Vlanti, A., Amillis, S., Pantazopoulou, A., and Diallinas, G. (2008) Specific interdomain synergy in the UapA transporter determines its unique specificity for uric acid among NAT carriers. *J Mol Biol* **382:** 1121–1135.
Papakostas, K., Botou, M., and Frillingos, S. (2013) Functional identification of the hypoxanthine/guanine transporters YjcD and YgfQ and the adenine transporters PurP and YicO of *Escherichia coli* K-12. *J Biol Chem* **288:** 36827–36840. doi: 10.1074/jbc.M113.523340
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **25:** 1605–1612.
Pfaller, M.A. (2012) Antifungal drug resistance: mechanisms, epidemiology, and consequences for treatment. *Am J Med* **125:** S3–S13.
Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., *et al*. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinforma Oxf Engl* **29:** 845–854.
Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
Savage, I.R. (1956) Contributions to the theory of rank order statistics – the two-sample case. *Ann Math Stat* **27:** 590–615.
Schultes, N.P., Brutnell, T.P., Allen, A., Dellaporta, S.L., Nelson, T., and Chen, J. (1996) Leaf permease1 gene of maize is required for chloroplast development. *Plant Cell* **8:** 463–475.
Shindyalov, I.N., and Bourne, P.E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* **11:** 739–747.
Söding, J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinforma Oxf Engl* **21:** 951–960.
Teles, J. (2012) Concordance coefficients to measure the agreement among several sets of ranks. *J Appl Stat* **39:** 1749–1764.
Trott, O., and Olson, A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31:** 455–461.
Vlanti, A., and Diallinas, G. (2008) The *Aspergillus nidulans* AzgA cytosine-purine scavenger is highly expressed during germination and in reproductive compartments and is downregulated by endocytosis. *Mol Microbiol* **68:** 959–977.
Vlanti, A., Amillis, S., Koukaki, M., and Diallinas, G. (2006) A novel-type substrate-selectivity filter and ER-exit determinants in the UapA purine transporter. *J Mol Biol* **357:** 808–819.
Wong, F.H., Chen, J.S., Reddy, V., Day, J.L., Shlykov, M.A., Wakabayashi, S.T., and Saier, M.H., Jr (2012) The amino acid-polyamine-organocation superfamily. *J Mol Microbiol Biotechnol* **22:** 105–113.
Wu, S., and Zhang, Y. (2007) LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic Acids Res* **35:** 3375–3382.

Yarov-Yarovoy, V., Schonbrun, J., and Baker, D. (2006) Multipass membrane protein structure prediction using Rosetta. *Proteins* **62:** 1010–1025.

Young, J.D., Yao, S.Y., Baldwin, J.M., Cass, C.E., and Baldwin, S.A. (2013) The human concentrative and equilibrative nucleoside transporter families, SLC28 and SLC29. *Mol Aspects Med* **34:** 529–547.

Zrenner, R., Stitt, M., Sonnewald, U., and Boldt, R. (2006) Pyrimidine and purine biosynthesis and degradation in plants. *Annu Rev Plant Biol* **57:** 805–836.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Fungal Genetics and Biology 75 (2015) 56–63

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Functional characterization of NAT/NCS2 proteins of *Aspergillus brasiliensis* reveals a genuine xanthine–uric acid transporter and an intrinsically misfolded polypeptide

Emilia Krypotou [a], Claudio Scazzocchio [b,c], George Diallinas [a,*]

[a] Faculty of Biology, University of Athens, Panepistimioupolis, Athens 15784, Greece
[b] Department of Microbiology, Imperial College, London SW7 2AZ, United Kingdom
[c] Institut de Génétique et Microbiologie, Université Paris-Sud, France

ARTICLE INFO

Article history:
Received 21 November 2014
Accepted 21 January 2015
Available online 29 January 2015

Keywords:
Purine
Fungi
Phylogeny
Pseudogene
Specificity

ABSTRACT

The Nucleobase-Ascorbate Transporter (NAT) family includes members in nearly all domains of life. Functionally characterized NAT transporters from bacteria, fungi, plants and mammals are ion-coupled symporters specific for the uptake of purines, pyrimidines and related analogues. The characterized mammalian NATs are specific for the uptake of L-ascorbic acid. In this work we identify *in silico* a group of fungal putative transporters, named UapD-like proteins, which represent a novel NAT subfamily. To understand the function and specificity of UapD proteins, we cloned and functionally characterized the two *Aspergillus brasiliensis* NAT members (named AbUapC and AbUapD) by heterologous expression in *Aspergillus nidulans*. AbUapC represents canonical NATs (UapC or UapA), while AbUapD represents the new subfamily. AbUapC is a high-affinity, high-capacity, H^+/xanthine–uric acid transporter, which can also recognize other purines with very low affinity. No apparent transport function could be detected for AbUapD. GFP-tagging showed that, unlike AbUapC which is localized in the plasma membrane, AbUapD is ER-retained and degraded in the vacuoles, a characteristic of misfolded proteins. Chimeric UapA/AbUapD molecules are also turned-over in the vacuole, suggesting that UapD includes intrinsic peptidic sequences leading to misfolding. The possible evolutionary implication of such conserved, but inactive proteins is discussed.

1. Introduction

The ubiquitous Nucleobase-Ascorbate Transporter (NAT) family, also known as the Nucleobase-Cation Symporter 2 (NCS2) family, includes thousands of putative members in nearly all organisms in almost every phylum (Diallinas and Gournas, 2008; Papageorgiou et al., 2008; Frillingos, 2012). NAT proteins share 14 α-helical hydrophobic transmembrane segments (TMS) and cytoplasmic N- and C-terminal regions, as well as, significant overall amino acid sequence identity (>24% among bacteria, fungi, plants or mammals) and conserved motifs. The two most conserved motifs, located in TMS1 (Gln-His) and TMS10 (Gln/Glu/Pro-Asn-X-Gly-X_4-Thr-Arg/Lys/Gly), serve as additional diagnostic criteria for NAT transporters. These motifs are part of the substrate-binding site of NATs (Amillis et al., 2011; Karena and Frillingos, 2011; Kosti et al., 2012)

Approximately 20 members of the NAT family have been functionally characterized to date (see Frillingos, 2012). Most of these are H^+ (bacteria, fungi, plants) or Na^+ (mammals) coupled symporters specific for the cellular uptake of xanthine, uric acid or uracil, and related analogues or drugs. Interestingly, in primates, NATs are solely specific for L-ascorbic acid (vitamin C) transport. Structure–function relationships of two NAT proteins, UapA from the ascomycete *Aspergillus nidulans* (Diallinas et al., 1998; Meintanis et al., 2000; Koukaki et al., 2005; Vlanti et al., 2006; Pantazopoulou and Diallinas, 2007; Papageorgiou et al., 2008; Kosti et al., 2010; Amillis et al., 2011) and XanQ from *Escherichia coli* (Georgopoulou et al., 2010; Mermelekas et al., 2010; Frillingos, 2012) have been studied in great detail. These studies have led to the identification of amino acids critical for substrate recognition and transport. The publication of the first crystal structure of a bacterial NAT member in 2011, the uracil transporter UraA from *E. coli* (Lu et al., 2011), together with genetic and biochemical studies in UapA and XanQ, led to a coherent view concerning structure–function relationships in NATs. An important novel concept arising from the studies on UapA is that

* Corresponding author. Fax: +30 (210)7274702.
E-mail address: diallina@biol.uoa.gr (G. Diallinas).

http://dx.doi.org/10.1016/j.fgb.2015.01.009
1087-1845/

substrate specificity is not solely determined by the major substrate binding site, but also by the functioning of channel-like gating domains (Diallinas, 2008, 2014). These gates are present at the outward- and inward-facing sides of the transporters and their apparent role is to open and close upon substrate binding and release (Papageorgiou et al., 2008; Kosti et al., 2010). The opening and closing of the gates is probably substrate-selective, but this issue remains to be shown (Diallinas, 2014). The possible *a priori* prediction of transporter specificity by *in silico* analyses is a vexing question. The work on NAT proteins has provided some valuable answers to this problem. However, as gating domains are not very conserved and function with different mechanisms in different transporters (Kazmier et al., 2014; Penmatsa and Gouaux, 2014; Simmons et al., 2014), predictions on substrate specificity are not straightforward. In addition, specificities can also arise through convergent evolution, which complicates predictions based on primary amino acid sequences and phylogenetic analyses. This is particularly evident in another nucleobase transporter family, NCS1 (Hamari et al., 2009).

In this work we describe the *in silico* identification and study of a sub-family of NATs showing a patchy distribution in fungi. This sub-family, called UapD, shows 45–47% identity with canonical fungal NATs, which have significantly higher similarity among themselves (>55%). Most importantly, the UapD-like proteins do not conserve the specific amino acid residues involved in purine or pyrimidine binding and transport, present in all other NATs. To understand the function and substrate specificity of the UapD group, we cloned and functionally characterized by heterologous expression in *A. nidulans* the two *Aspergillus brasiliensis* NAT proteins, one of which is similar to canonical NATs (named AbUapC) while the other represents the UapD group (AbUapD). While one of the homologues functions as a canonical NAT transporter, AbUapD is an intrinsically misfolded non-functional protein. The possible evolutionary implication of such conserved but inactive proteins will be discussed.

2. Materials and methods

2.1. Phylogenetic analysis

Fungal genomes currently present at NCBI, JGI (Nordberg et al., 2014; Grigoriev et al., 2012) and AspGD (Cerqueira et al., 2013) were searched for NAT sequences using blastp and tblastn. The *A. nidulans* UapC sequence (Uniprot ID: P48777) was used as the *in silico* probe. Selected NATs from all major taxa of dikarya and NATs of known function from all organisms were used to construct a phylogenetic tree (the species included in the phylogenetic trees are shown in Supplementary Table S1). Multiple sequence alignment was performed with MAFFT (Katoh and Standley, 2013), curation of the alignment with BMGE (Criscuolo and Gribaldo, 2010) and phylogenetic analysis with PhyML (Guindon et al., 2010). Trees drawn with FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

2.2. Media, strains, plasmids, growth conditions and A. nidulans transformation

Standard complete and minimal media (MM) for *A. nidulans* were used. Media and supplemented auxotrophies were at the concentrations given in http://www.fgsc.net. Nitrogen sources and analogues were used at the final concentrations: $NaNO_3$ (NO_3) 10 mM, purines 0.5 mM, xanthine (XA) 1 mM and oxypurinol (OX) 100 μM. Media and chemical reagents were obtained from Sigma–Aldrich (Life Science Chemilab SA, Hellas) or AppliChem (Bioline Scientific SA, Hellas). *A. nidulans* transformation was performed as in Koukaki et al., 2003. A *ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2* mutant, named ΔΑCZ (Pantazopoulou et al., 2007), was the recipient strain in transformations with plasmids carrying the *AbuapC* and *AbuapD* genes. Selection was based on complementation of arginine auxotrophy *argB2*. *AbuapC* and *AbuapD* were amplified from *A. brasiliensis* genomic DNA and cloned in a modified pBluescript vector containing 5′ and 3′ *uapA* regulatory regions and the *argB* gene as a genetic marker. The vector also contained the *gfp* open reading frame cloned in frame at the C-terminal region of *AbuapC* and *AbuapD* genes. Similar vectors without the *gfp* orf or with the *gfp* orf cloned in frame at the N-terminal region of *AbuapD* or chimeric versions with *uapA* and *AbuapD*, were also used. Strains and primers used in this study are shown in Tables 1 and 2, respectively. Transformants expressing intact alleles, through either single-copy or multi-copy plasmid integration events, were identified by PCR and Southern analysis. All shown strains are single-copy transformants. Growth tests were performed at 37 °C or 25 °C, at pH 6.8, for both *A. nidulans* and *A. brasiliensis*.

2.3. Standard nucleic acid manipulations

A. nidulans and *A. brasiliensis* DNA extraction was as described in http://www.fgsc.net. Plasmid preparation and purification, and purification of DNA fragments from agarose gels, was performed with the Nucleospin Plasmid kit and the Nucleospin ExtractII kit (Macherey–Nagel, Lab Supplies Scientific SA, Hellas). Southern blot analysis was performed as described in Sambrook et al., 1989. Conventional PCR reactions were performed with KAPATaq DNA polymerase (Kapa Biosystems, Lab Supplies Scientific SA, Hellas). Cloning and amplification of products were performed with Kapa HiFi (Kapa Biosystems, Lab Supplies Scientific SA, Hellas).

2.4. Transport assays and Epifluorescence microscopy

[^{3}H]-xanthine (22.8 Ci mmol^{-1}, Moravek Biochemicals, CA, USA) assays were performed at 37 °C and pH 6.8, in germinating conidiospores of *A. nidulans* as recently described in detail in Krypotou and Diallinas (2014). Epifluorescence microscopy was as previously described (Gournas et al., 2010; Karachaliou et al., 2013).

3. Results and discussion

3.1. Phylogenetic analysis revealed a previously unknown NAT sub-family: the UapD group

Currently known fungal NATs are named UapA or UapC, based on their phylogenetic relationship to the *A. nidulans* UapA and UapC transporters, which have been rigorously analyzed with respect to their function and specificity (Goudela et al., 2005; Krypotou and Diallinas, 2014). A UapC protein from *Aspergillus fumigatus* has also been kinetically characterized (Goudela et al., 2008). The major distinction between UapA- and UapC-like proteins is that the former are specific for the high-affinity binding and transport of both uric acid and xanthine, whereas the latter are more promiscuous, recognizing xanthine with high affinity, uric acid with moderate affinity, and several other purines with very low affinity.

We used the *A. nidulans* UapC transporter sequence as an *in silico* probe to search for all fungal NAT sequences. NAT sequences were ubiquitously found in dikarya, but showed a limited and patchy distribution in more primitive fungi (data not shown). Fig. 1 shows an evolutionary tree in a cartoon form of selected proteins showing significant similarity to known NATs (UapC/UapA) of

Table 1
Strains used in this study.

Strain name	Genotype (Reference)
wt	*pabaA1* (FGSC)
ΔACZ	*ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2* (Pantazopoulou et al., 2007)
UapA-GFP	*uapA-GFP-argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1* (Pantazopoulou et al., 2007)
UapC-GFP	*uapC-GFP-argB ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2* (Pantazopoulou et al., 2007)
AbUapD	*AbuapD-argB** *ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2* (This study)
AbUapC-GFP	*AbuapC-GFP-argB** *ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2* (This study)
AbUapD-GFP	*AbuapD-GFP-argB** *ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2* (This study)
GFP-AbUapD	*GFP-AbuapD-argB** *ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2* (This study)
AbuapD-A-GFP	*AbuapD-A-GFP-argB** *ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2* (This study)
ADA-GFP	*ADA-GFP-argB** *ΔuapA ΔuapC::AfpyrG ΔazgA pabaA1 argB2* (This study)
CBS101740	*Aspergillus brasiliensis* (A gift from the Fungal Biodiversity Centre Institute of the Royal Netherlands Academy of Arts and Sciences (KNAW))

* For plasmid constructions see Section 2.

Table 2
Oligonucleotides used in this study. Restriction sites are in bold.

Primer name	Sequence 5–3′
AbuapC FOR	GCG**AGATCT**ACCTGACCAGATAGGCCCG
AbuapC REV	GCG**ACTAGT**GCACTCGAGGGCTCGC
AbuapD FOR	GCG**CTGCAG**ATGGAGTCCAAGTCTAGCCC
AbuapD REV	GCG**TCTAGA**TCAAGGCAACACTGTATCCACCTTC
AbuapD NS REV	GCG**TCTAGA**GGCAACACTGTATCCACCTTC
AbuapD-NO-CT REV	CGC**CTGCAG**GCAACAACAAATTCAACAACACCCC
uapAct FOR	GCG**CTGCAG**AAGTCGAAGAGATCGGTG
AbuapD ADA FOR	GCGC**CTGCAG**CAACAGGCTACTGGTCTCGC
AbuapD ADA REV	GCGC**CTGCAG**CGAATTTACCGAAGATGCCC
GFP FOR	CGC**CTGCAG**ATGGTGAGCAAGGGCGAGG
GFP NS REV	CGC**CTGCAG**CTTGTACAGCTCGTCCATGCC

dikaryotic fungi. All NAT sequences found and used for the phylogenetic analysis are shown in Supplementary Fig. S1. The tree can be divided in two major clades, one including the canonical NATs (UapA- and UapC-like and similar sequences) and a clade not previously identified, which we called UapD. The term canonical is used here to describe NAT sequences fulfilling the following criteria: (i) >55% identity with the functionally characterized UapA and UapC transporters of *A. nidulans* and *A. fumigatus*, (ii) the presence of conserved motifs and specific amino acid residues shown to be necessary for purine or pyrimidine binding and transport, as identified through genetic, biochemical, structural and modeling studies on UapA and bacterial homologues, and (iii) conservation, throughout dikarya (with the exception that post-whole genome duplication members of the Saccharomycetales have lost NATs altogether). The UapD sequences do not satisfy any of these criteria.

The major group of canonical NATs includes the UapC-like homologues (with identities of 74–91.4%), present in nearly all fungi and duplicated in some. *A. nidulans* has both UapA and UapC transporters. A UapA-like paralogue is present in *Aspergillus zonatus*, but absent from both *Aspergillus versicolor* and *Aspergillus sydowii*, which are closely related to *A. nidulans*. *Neurospora crassa* has a UapA-like transporter, as evidenced both by the present phylogenetic analysis and older biochemical assays (Tsao and Marzluf, 1976). The functional analysis of the *Candida albicans* NAT transporter (Xut1) shows intermediate characteristics of UapA and UapC (Goudela et al., 2005), which is in line with its phylogenetic position.

UapD proteins show patchy conservation in ascomycetes and are absent from basidiomycetes, suggestive of significant evolutionary losses. UapDs are also absent from non-dikaryotic fungi and non-fungal organisms. UapD sequences show 45.5–47.8% identity with UapC and UapA, but most importantly do not conserve residues necessary for purine or pyrimidine binding (see below). UapD-like proteins share 60–96% identity among themselves, depending on the evolutionary distance of the relevant species in binary comparisons. The UapD clade seems to be ancestral, as it has homologues in fungi as divergent as Aspergilli and *Tuber melanosporum* (black truffle), a basal species of the Pezizomycotina. It is also present in *Kuraishia capsulata*, a non-basal member of the Saccharomycotina (Morales et al., 2013), where an orthologue of the characterized Xut1 of *C. albicans* is also present (Goudela et al., 2005) (Fig. 1 and Supplementary Fig. S1). In *Aspergillus* species, UapD genes are syntenic and the six full-length genes have 3–4 introns in well-conserved positions (results not shown). A partial UapD-like sequence is found in *Aspergillus terreus*, which seems to represent a putative pseudogene. The full-length UapD proteins are predicted to have 14 TMSs, similar to canonical NATs, as shown by homology modeling (results not shown).

3.2. UapD does not conserve amino acid residues involved in nucleobase binding and transport

We compared the UapD primary sequences with those of UapC and UapA in order to detect similarities and differences in specific amino acid residues known to be involved in substrate binding and transport, or in other elements involved in specificity. Fig. 2 highlights our findings. Unlike all canonical NATs, which have been shown to be nucleobase transporters, UapD does not conserve the two major residues, a Glu in TMS8 and a Gln/Glu at TMS10, which are involved in strong H bond interactions with substrates. In canonical NATs, these two residues are irreplaceable for function in all cases tested (e.g. UapA, XanQ, UraA). In UapD, the analogous residues are replaced by Ser/Leu/Val and Gly/Ala, respectively (underlined residues depict most frequent replacements). In addition, UapD has some major differences when compared to UapA and UapC in a number of residues supposed to be within bonding distance from the putative substrate-binding site. For example, a well-conserved Phe/Tyr/Ser in TMS10, which has been shown to be critical for substrate specificity in UapA (Phe^{406}; Kosti et al., 2010), is replaced by Gln/His in UapD sequences. On the other hand, UapDs conserve some essential elements of the NAT family. They possess an intact Gln-His motif in TMS1 and an Asp in TMS8, which have been shown to be necessary for transport activity in UapA and XanQ. They also conserve residues shown to be elements of selective gates located in TMS13 (inward-facing gate) and TMS14 (outward-facing gate) of UapA. Importantly however, none of the conserved residues are directly involved in substrate binding. Overall, our analysis predicts that UapD cannot recognize purines, pyrimidines or ascorbic acid, as none of the necessary residues for this latter activity are conserved (Kosti et al., 2012).

The N-terminal region of UapD proteins is highly similar to canonical NATs. This region is essential for translocation and exit from the ER and also affects the turnover of UapA, and probably of other NATs (M. Evangelinos, S. Amillis and G. Diallinas, unpublished). In particular

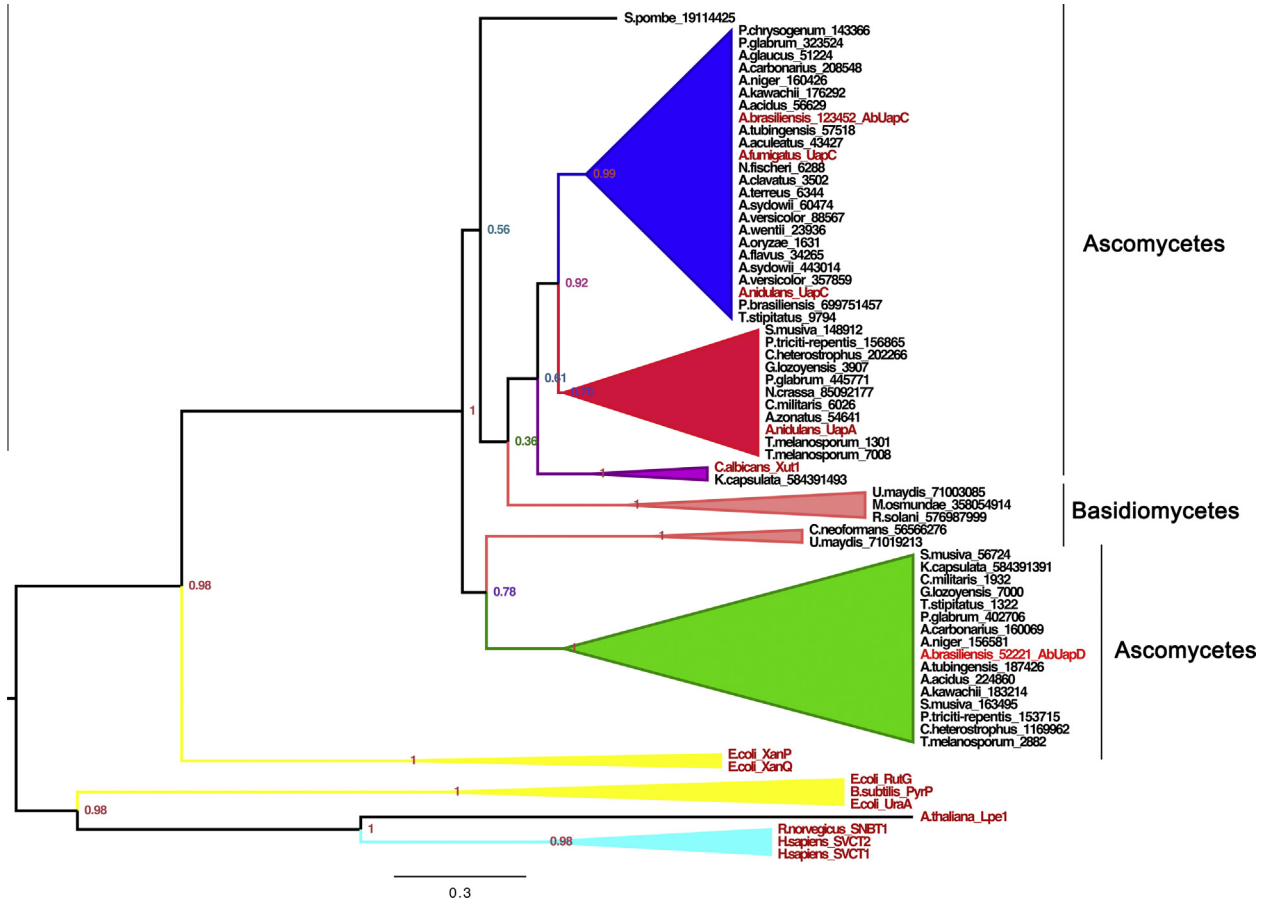


Fig. 1. Maximum Likelihood Phylogenetic rooted tree of selected fungal NAT proteins in cartoon form. A more comprehensive tree is shown in Supplementary Fig. S1. The functionally characterized proteins are in red lettering. The cartoons comprising major clades are color coded as: Pezimycotina: UapC-like, blue; UapA-like, red; UapD-like, green; Saccharomycotina: purple; Basidiomycetes: pink; Bacterial NATs: yellow; Mammalian NATs: cyan. The branches of the single members included from the Taphrinomycotina (*S. pombe*) and Plants (*A. thaliana*) respectively are in black. The only UapD paralogue within the Saccharomycotina (K.capsulata_584391391) is also included in the UapD (green) cartoon representation. The full names of species can be found in Supplementary Table S1. The protein IDs are as in JGI or NCBI, except for the previously characterized homologues which are indicated by their published names. Uniprot IDs for the characterized proteins shown are: UapA: Q07307; UapC: P48777; *A. fumigatus*_UapC: Q4WGR0; Xut1: Q5BN32; Lpe1: Q41760; UraA: P0AGM7; XanQ: P67444; XanP: P0AGM9; SVCT1: Q9UHI7; SVCT2: Q9UGH3; SNBT1: D2KX48. Alignment carried out with MAFFT, G-INS-i with Blosum62 scoring matrix, curation with BMGE 1.0 with default parameters, phylogenetic tree with PhyML with default parameters, redrawn with FigTree. The digits at the nodes indicate aLTR (Approximate Likelihood-Ratio Test values, Anisimova and Gascuel, 2006). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

an absolutely conserved Asp-Tyr-Asp-Tyr motif, also present in the UapD clade, in the middle of the N-terminal region has been shown to be necessary for UapA membrane trafficking and turnover. Thus, UapD seems to possess the proper signals for secretion toward the plasma membrane. On the other hand, the C-terminal regions of UapDs, which are approximately 10 amino acid residues long, are significantly shorter than those of canonical NATs and in this regard are more akin to bacterial homologues. The C-terminal region of UapA has been shown to be essential for ubiquitination, internalization from the plasma membrane and vacuolar turnover in response to endocytic signals, such as the presence of ammonium or excess substrate (Gournas et al., 2010; Karachaliou et al., 2013). In fact, removal of the entire C-terminal region of UapA stabilizes the protein in the plasma membrane (Gournas et al., 2010). Thus, these observations suggested that UapD might be a transporter that can be normally sorted and stay stably localized in the plasma membrane. As will be shown in this work, this is not the case (see below).

3.3. Expression of the two A. brasiliensis NAT/NCS2 homologues in A. nidulans

To investigate the function of UapD-like proteins, we cloned and functionally analyzed the UapD from *A. brasiliensis* (AbUapD) by heterologous expression in *A. nidulans*. We also cloned and studied the putative UapC transporter from the same species (AbUapC) as a control.

We constructed plasmid-borne GFP-tagged *AbuapD* and *AbuapC* fusions, driven by the well-studied *A. nidulans uapA* promoter (Gorfinkiel et al., 1993). Given our previous experience with *uapC* from *A. nidulans* (Krypotou and Diallinas, 2014) we expected the AbUapC-GFP fusion to be fully active. In fact, this fusion complements a *ΔuapA ΔuapC ΔazgA* deletion mutant strain, allowing growth on xanthine and uric acid and resulting in sensitivity to oxypurinol. In addition, AbUapC-GFP shows the expected subcellular localization in the plasma membrane and septa, and very little association with vacuoles (Fig. 3A). As we could not know *a priori* if and how a GFP fusion affects the folding of AbUapD, three constructions all driven by the *uapA* promoter were tested, namely $uapA_p$-*AbuapD*, $uapA_p$-*gfp-AbuapD*, $uapA_p$-*AbuapD-gfp*. In no case was complementation of the *ΔuapA ΔuapC ΔazgA* phenotype observed. Neither the amino terminal nor the carboxy terminal GFP fusions localize at the membrane, but instead in intracellular structures identified as the perinuclear endoplasmic reticulum (ER) and vacuoles (Fig. 3B). These results are consistent with the absence of complementation of the growth phenotype of the *ΔuapA ΔuapC ΔazgA* strain by AbUapD expression, as correct

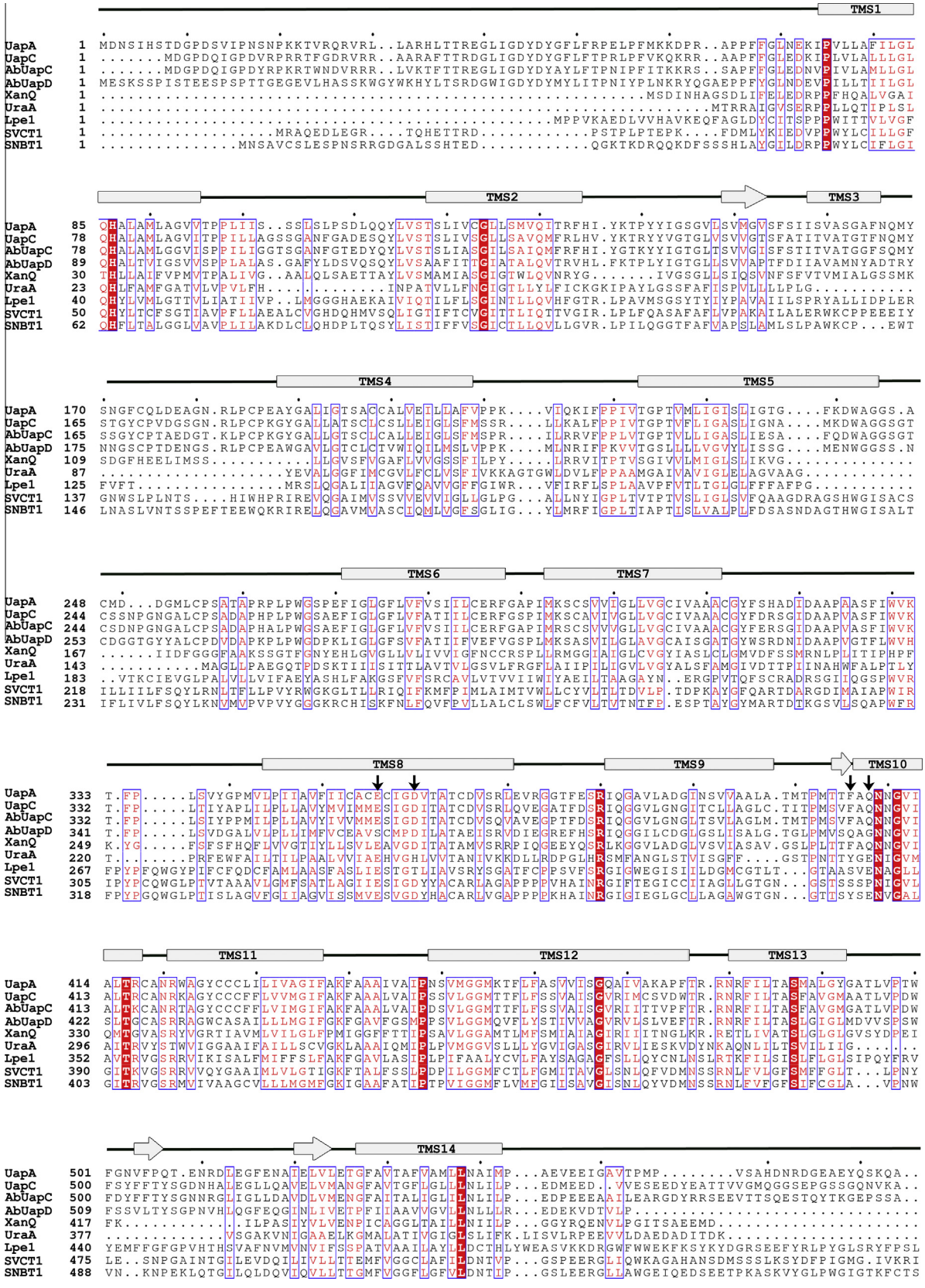


Fig. 2. Sequence alignment of selected NAT homologues. The sequences shown are: AbUapC, AbUapD (*A. brasiliensis*); UapA, UapC (*A. nidulans*); Lpe1 (*Arabidopsis thaliana*); UraA, XanQ (*E. coli*); SVCT1 (*Homo sapiens*); SNBT1 (*Rattus norvegicus*); for protein IDs see Fig. 1. Predicted transmembrane segments (TMS) for UapA are denoted with rectangles and TMSs forming short b-sheets are shown with horizontal arrows (Kosti et al., 2012). Residues of the NAT signature motif varying in UapD-like proteins are shown with vertical arrows. Invariant and highly conserved amino acids are shaded in red and blue-lined boxes, respectively. Absolutely conserved amino acids are in red shaded boxes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

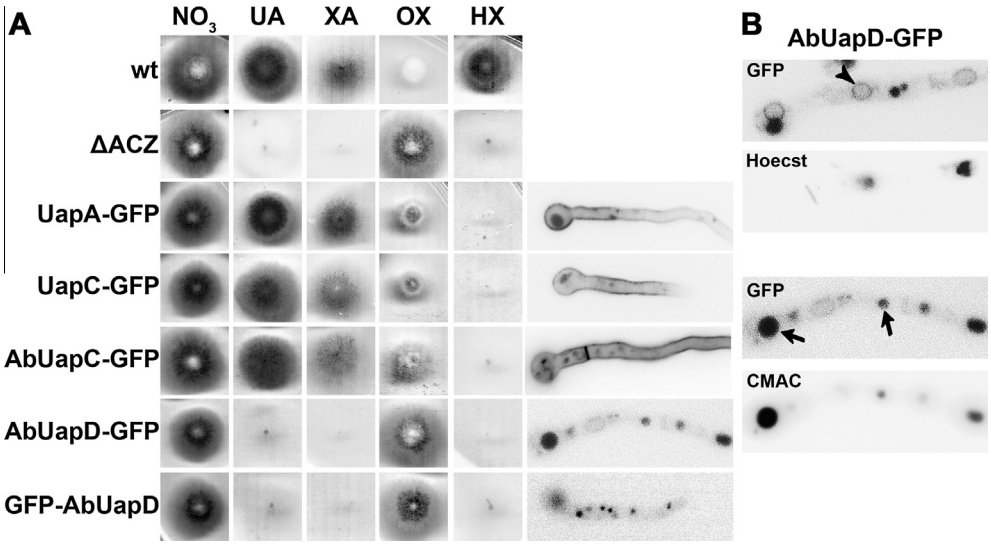


Fig. 3. (A) Left panel: Growth test of selected transformants expressing AbUapC-GFP and AbUapD-GFP. A wild type strain (wt), an isogenic negative control (ΔACZ), and isogenic strains expressing UapA-GFP or UapC-GFP are also shown. Right panel: Microscopic analysis for GFP-detection in strains shown in the growth test. (B) Microscopic analysis of the strain expressing AbUapD-GFP. In the upper panel perinuclear ER rings (arrowheads) were identified by nuclear staining with Hoechst 33342 dye. In the lower panel vacuoles (arrows) were identified by staining with CMAC dye. Conidiospores were inoculated in MM and grown for 16 h at 25 °C and pH 6.8 (NO_3: sodium nitrate, UA: uric acid, XA: xanthine, HX: hypoxanthine, OX: oxypurinol). In all cases (A and B), single-copy transformants are shown.

localization to the cell membrane is a necessary (but not sufficient) condition for transporter activity.

Given that AbUapC proved to be a genuine xanthine–uric acid transporter when expressed in *A. nidulans*, it is expected that it would function similarly in *A. brasiliensis*. This is in line with growth tests showing that *A. brasiliensis* grows similarly to *A. nidulans* on uric acid, xanthine and hypoxanthine as sole nitrogen sources, and shows moderate sensitivity to oxypurinol (Fig. 4).

3.4. Kinetics and specificity profile of AbUapC

To determine the kinetics and specificity profile of AbUapC we carried out uptake assays of radiolabelled [^{3}H]-xanthine in the *A. nidulans* strain expressing AbUapC-GFP. AbUapC-mediated xanthine transport reaches a plateau in 2 min, similarly to most other transporters studied so far in *A. nidulans* (Fig. 5A). AbUapC-mediated [^{3}H]-xanthine uptake is reduced ~70–90% in the presence of excess (1000-fold) xanthine, uric acid, hypoxanthine, 1-, 3- or 8-methylxanthine and ~50–30% in the presence of guanine, adenine, 6-thioxanthine, 8-azaxanthine, uracil or cytosine (Fig. 5B). Substrate binding affinities and inhibition constants ($K_{m/i}$) were determined for xanthine, uric acid and hypoxanthine (Fig. 5C). AbUapC shows high affinity for xanthine (K_m 23 μM), medium affinity for uric acid (K_i 160 μM) and low affinity for hypoxanthine (K_i 700 μM). This profile confirms its classification as a UapC-like, rather than a UapA-like, transporter.

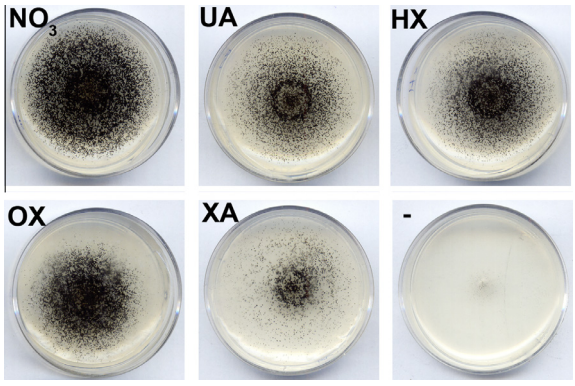


Fig. 4. Growth test of *A. brasiliensis* at 37 °C. (NO_3: sodium nitrate, UA: uric acid, XA: xanthine, HX: hypoxanthine, OX: oxypurinol, – : lack of any nitrogen source).

3.5. AbuapD is an intrinsically misfolded protein

The partial ER-retention and vacuolar degradation of AbUapD is typical of partially misfolded proteins. To further investigate the reasons for AbUapD apparent instability, we constructed two GFP-tagged, chimeric proteins, using sections of UapA. In the first chimaera, we replaced the short C-terminal region of AbUapD with that of UapA (named AbUapD-A). In the second case, we replaced a sequence of *A. nidulans* UapA (amino acids 311–443) with the corresponding sequence of AbUapD (amino acids 319–451) (named UapADA). The region replaced includes the major elements of the substrate-binding site of canonical NATs (TMS8–TMS10). Importantly, mutations made in several residues of UapA or bacterial NATs in this region do not affect the folding of these transporters. It is also noteworthy that a similar chimaera made by using UapA and UapC sequences was functional and normally localized in the plasma membrane (Diallinas et al., 1998; Diallinas, 2013). Fig. 6 shows that both chimaeras we made herein were mainly localized in the ER. This result strongly suggests that AbUapD contains elements leading to intrinsic misfolding and high turnover. Growth tests in Fig. 6 confirm that these chimaeras do not confer any apparent nucleobase transport activity.

3.6. Proto-pseudogenes or proto-neogenes? A hypothesis

It is extremely surprising to find a whole clade of UapA/UapC-like transporters, named UapD, with specifically conserved residues, which if the results obtained with AbUapD can be generalized, show characteristics of misfolded and highly turned-over proteins. We have recently reported that in *A. nidulans*, within the Fur family of transporters, one member called FurB, shows the same intracellular localization as AbUapD, typical of misfolded proteins, when its expression is driven by an heterologous promoter (Krypotou et al., manuscript under revision). Moreover,

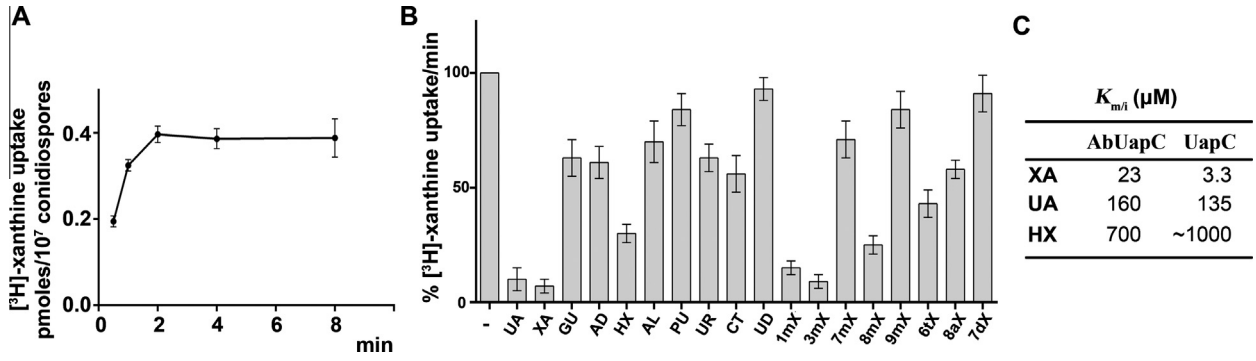


Fig. 5. Kinetic profile of AbUapC. (A) Time course of AbUapC mediated [^{3}H]-xanthine. (B) AbUapC mediated [^{3}H]-xanthine uptake in the presence of excess (1000-fold) of different substrates/inhibitors. (C) Affinity and inhibition kinetic constants $K_{m/i}$ (μM). Results are averages of at least three independent experiments in triplicate for each concentration point. Standard deviation was <20% (UA: uric acid, XA: xanthine, GU: guanine, AD: adenine, HX: hypoxanthine, AL: allantoin, PU: purine, UR: uracil, CT: cytosine, UD: uridine, 1mX: 1-methylxanthine, 3mX: 3-methylxanthine, 7mX: 7-methylxanthine, 8mX: 8-methylxanthine, 9mX: 9-methylxanthine, 6tX: 6-thioxanthine, 8aX: 8-azaxanthine, 7dX: 7-deazaxanthine). In all cases (A, B and C), single-copy transformants are shown.

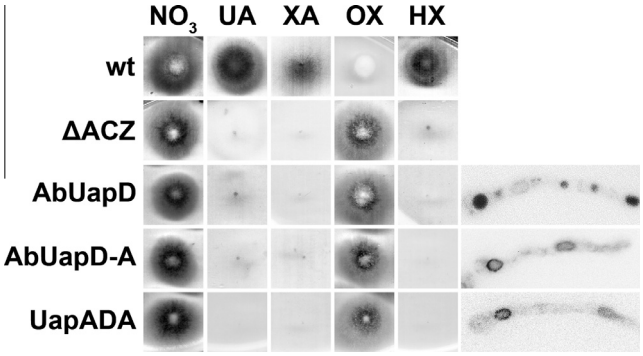


Fig. 6. Growth tests and microscopic analysis of the AbUapD chimeric proteins (NO_3: sodium nitrate, UA: uric acid, XA: xanthine, HX: hypoxanthine, OX: oxypurinol). Single-copy transformants are shown.

putative orthologues of FurB show also a pattern of patchy phylogenetic distribution. In addition, we have not detected transcription of *furB* from its native promoter (Hamari et al., 2009; E. Krypotou and G. Diallinas, unpublished results). We do not have transcription data for *uapD* in *A. brasiliensis*, but no transcription was detected for the orthologous *Aspergillus niger uapD* (Delmas et al., 2012; van Munster et al., 2014). We are aware of other cases where forced expression of transporters in *E. coli* (S. Frillingos, pers. com) or *Saccharomyces cerevisiae* (M. Casal, pers. com) do not result in a detectable polypeptides.

As gene duplication relaxes selection pressure and allows mutational divergence of paralogues, it is to be expected that some of these mutations could lead to partial misfolding and loss of activity. These 'junk' polypeptides would accumulate more mutations, which will, usually, lead to translation stops (pseudogenes) and eventual extinction. In fact we have detected a *bona fidae uapD* pseudogene (loss of start codon, presence of several stop codons and truncations) in *Pyrenophora teres*, while its orthologue in *Pyrenophora tritici*, shows an open reading frame (data not shown). AbUapD or FurB might constitute intermediate evolutionary steps of inactive polypeptides, which will most probably, but not necessarily, become extinct, and thus we might call the relevant genes as 'proto-pseudogenes'. However, as these genes in spite of encoding misfolded proteins show perfectly constituted open reading frames, they could evolve toward a new functionality, and thus could also be considered as 'proto-neogenes'. Despite the fact that it is generally accepted that novel function cannot arise in single mutational steps, the fate of molecules representing intermediate steps is usually not discussed. How long could these proto-pseudo (neo) genes persist without a function and no selective pressure imposed upon them? Could it be that their presence has a cryptic function unrelated to their canonical homologues (e.g. acting as misfolded polypeptides titrating out proteolytic mechanisms; Adrain and Freeman, 2012; Zettl et al., 2011), or is there a conservative mechanism at the DNA/chromatin level that shields them from evolutionary extinction mechanisms?

Acknowledgments

The *A. brasiliensis* strain was provided by Ronald De Vries (Fungal Biodiversity Centre Institute of the Royal Netherlands Academy of Arts and Sciences).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fgb.2015.01.009.

References

Adrain, C., Freeman, M., 2012. New lives for old: evolution of pseudoenzyme function illustrated by iRhoms. Nat. Rev. Mol. Cell Biol. 13, 489–498.
Amillis, S., Kosti, V., Pantazopoulou, A., Mikros, E., Diallinas, G., 2011. Mutational analysis and modeling reveal functionally critical residues in transmembrane segments 1 and 3 of the UapA transporter. J. Mol. Biol. 411, 567–580.
Anisimova, M., Gascuel, O., 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst. Biol. 55, 539–552.
Cerqueira, G.C., Arnaud, M.B., Inglis, D.O., Skrzypek, M.S., Binkley, G., Simison, M., Miyasato, S.R., Binkley, J., Orvis, J., Shah, P., Wymore, F., Sherlock, G., Wortman, J.R., 2013. The *Aspergillus* Genome Database: multispecies curation and incorporation of RNA-Seq data to improve structural gene annotations. Nucleic Acids Res. 42 (Database issue), D705–D710.
Criscuolo, A., Gribaldo, S., 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol. Biol. 10, 210.
Delmas, S., Pullan, S.T., Gaddipati, S., Kokolski, M., Malla, S., Blythe, M.J., Ibbett, R., Campbell, M., Liddell, S., Aboobaker, A., Tucker, G.A., Archer, D.B., 2012. Uncovering the genome-wide transcriptional responses of the filamentous fungus *Aspergillus niger* to lignocellulose using RNA sequencing. PLoS Genet. 8, e1002875.
Diallinas, G., 2008. Biochemistry. An almost-complete movie. Science 322, 1644–1645.
Diallinas, G., 2013. Allopurinol and xanthine use different translocation mechanisms and trajectories in the fungal UapA transporter. Biochimie 95, 1755–1764.
Diallinas, G., 2014. Understanding transporter specificity and the discrete appearance of channel-like gating domains in transporters. Front. Pharmacol. 5, 207.
Diallinas, G., Gournas, C., 2008. Structure–function relationships in the nucleobase-ascorbate transporter (NAT) family: lessons from model microbial genetic systems. Channels (Austin) 2, 363–372.
Diallinas, G., Valdez, J., Sophianopoulou, V., Rosa, A., Scazzocchio, C., 1998. Chimeric purine transporters of *Aspergillus nidulans* define a domain critical for function

and specificity conserved in bacterial, plant and metazoan homologues. EMBO J. 17, 3827–3837.
Frillingos, S., 2012. Insights to the evolution of Nucleobase-Ascorbate Transporters (NAT/NCS2 family) from the Cys-scanning analysis of xanthine permease XanQ. Int. J. Biochem. Mol. Biol. 3, 250–272.
Georgopoulou, E., Mermelekas, G., Karena, E., Frillingos, S., 2010. Purine substrate recognition by the nucleobase-ascorbate transporter signature motif in the YgfO xanthine permease: ASN-325 binds and ALA-323 senses substrate. J. Biol. Chem. 285, 19422–19433.
Gorfinkiel, L., Diallinas, G., Scazzocchio, C., 1993. Sequence and regulation of the uapA gene encoding a uric acid–xanthine permease in the fungus *Aspergillus nidulans*. J. Biol. Chem. 268, 23376–23381.
Goudela, S., Karatza, P., Koukaki, M., Frillingos, S., Diallinas, G., 2005. Comparative substrate recognition by bacterial and fungal purine transporters of the NAT/NCS2 family. Mol. Membr. Biol. 22, 263–275.
Goudela, S., Reichard, U., Amillis, S., Diallinas, G., 2008. Characterization and kinetics of the major purine transporters in *Aspergillus fumigatus*. Fungal Genet. Biol. 45, 459–472.
Gournas, C., Amillis, S., Vlanti, A., Diallinas, G., 2010. Transport-dependent endocytosis and turnover of a uric acid–xanthine permease. Mol. Microbiol. 75, 246–260.
Grigoriev, I.V., Nordberg, H., Shabalov, I., Aerts, A., Cantor, M., Goodstein, D., Kuo, A., Minovitsky, S., Nikitin, R., Ohm, R.A., Otillar, R., Poliakov, A., Ratnere, I., Riley, R., Smirnova, T., Rokhsar, D., Dubchak, I., 2012. The genome portal of the Department of Energy Joint Genome Institute. Nucleic Acids Res. 40 (Database issue), D26–D32.
Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.
Hamari, Z., Amillis, S., Drevet, C., Apostolaki, A., Vágvölgyi, C., Diallinas, G., Scazzocchio, C., 2009. Convergent evolution and orphan genes in the Fur4p-like family and characterization of a general nucleoside transporter in *Aspergillus nidulans*. Mol. Microbiol. 73, 43–57.
Karachaliou, M., Amillis, S., Evangelinos, M., Kokotos, A.C., Yalelis, V., Diallinas, G., 2013. The arrestin-like protein ArtA is essential for ubiquitination and endocytosis of the UapA transporter in response to both broad-range and specific signals. Mol. Microbiol. 88, 301–317.
Karena, E., Frillingos, S., 2011. The role of transmembrane segment TM3 in the xanthine permease XanQ of *Escherichia coli*. J. Biol. Chem. 286, 39595–39605.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability (outlines version 7). Mol. Biol. Evol. 30, 772–780.
Kazmier, K., Sharma, S., Quick, M., Islam, S.M., Roux, B., Weinstein, H., Javitch, J.A., McHaourab, H.S., 2014. Conformational dynamics of ligand-dependent alternating access in LeuT. Nat. Struct. Mol. Biol. 21, 472–479.
Kosti, V., Papageorgiou, I., Diallinas, G., 2010. Dynamic elements at both cytoplasmically and extracellularly facing sides of the UapA transporter selectively control the accessibility of substrates to their translocation pathway. J. Mol. Biol. 397, 1132–1143.
Kosti, V., Lambrinidis, G., Myrianthopoulos, V., Diallinas, G., Mikros, E., 2012. Identification of the substrate recognition and transport pathway in a eukaryotic member of the nucleobase-ascorbate transporter (NAT) family. PLoS One 7, e41939.
Koukaki, M., Giannoutsou, E., Karagouni, A., Diallinas, G., 2003. A novel improved method for *Aspergillus nidulans* transformation. J. Microbiol. Methods 55, 687–695.
Koukaki, M., Vlanti, A., Goudela, S., Pantazopoulou, A., Gioule, H., Tournaviti, S., Diallinas, G., 2005. The nucleobase-ascorbate transporter (NAT) signature motif in UapA defines the function of the purine translocation pathway. J. Mol. Biol. 350, 499–513.
Krypotou, E., Diallinas, G., 2014. Transport assays in filamentous fungi: kinetic characterization of the UapC purine transporter of *Aspergillus nidulans*. Fungal Genet. Biol. 63, 1–8.
Lu, F., Li, S., Jiang, Y., Jiang, J., Fan, H., Lu, G., Deng, D., Dang, S., Zhang, X., Wang, J., Yan, N., 2011. Structure and mechanism of the uracil transporter UraA. Nature 472, 243–246.
Meintanis, C., Karagouni, A.D., Diallinas, G., 2000. Amino acid residues N450 and Q449 are critical for the uptake capacity and specificity of UapA, a prototype of a nucleobase-ascorbate transporter family. Mol. Membr. Biol. 17, 47–57.
Mermelekas, G., Georgopoulou, E., Kallis, A., Botou, M., Vlantos, V., Frillingos, S., 2010. Cysteine-scanning analysis of helices TM8, TM9a, and TM9b and intervening loops in the YgfO xanthine permease: a carboxyl group is essential at ASP-276. J. Biol. Chem. 285, 35011–35020.
Morales, L., Noel, B., Porcel, B., Marcet-Houben, M., Hullo, M.F., Sacerdot, C., Tekaia, F., Leh-Louis, V., Despons, L., Khanna, V., Aury, J.M., Barbe, V., Couloux, A., Labadie, K., Pelletier, E., Souciet, J.L., Boekhout, T., Gabaldon, T., Wincker, P., Dujon, B., 2013. Complete DNA sequence of *Kuraishia capsulata* illustrates novel genomic features among budding yeasts (Saccharomycotina). Genome Biol. Evol. 5, 2524–2539.
Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I.V., Dubchak, I., 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res. 42, D26–D31.
Pantazopoulou, A., Diallinas, G., 2007. Fungal nucleobase transporters. FEMS Microbiol. Rev. 31, 657–675.
Pantazopoulou, A., Lemuh, N.D., Hatzinikolaou, D.G., Drevet, C., Cecchetto, G., Scazzocchio, C., Diallinas, G., 2007. Differential physiological and developmental expression of the UapA and AzgA purine transporters in *Aspergillus nidulans*. Fungal Genet. Biol. 44, 627–640.
Papageorgiou, I., Gournas, C., Vlanti, A., Amillis, S., Pantazopoulou, A., Diallinas, G., 2008. Specific interdomain synergy in the UapA transporter determines its unique specificity for uric acid among NAT carriers. J. Mol. Biol. 382, 1121–1135.
Penmatsa, A., Gouaux, E., 2014. How LeuT shapes our understanding of the mechanisms of sodium-coupled neurotransmitter transporters. J. Physiol. 592, 863–869.
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, second ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Simmons, K.J., Jackson, S.M., Brueckner, F., Patching, S.G., Beckstein, O., Ivanova, E., Geng, T., Weyand, S., Drew, D., Lanigan, J., Sharples, D.J., Sansom, M.S., Iwata, S., Fishwick, C.W., Johnson, A.P., Cameron, A.D., Henderson, P.J., 2014. Molecular mechanism of ligand recognition by membrane transport protein, Mhp1. EMBO J. 33, 1831–1844.
Tsao, T.F., Marzluf, G.A., 1976. Genetic and metabolic regulation of purine base transport in *Neurospora crassa*. Mol. Gen. Genet. 149, 347–355.
van Munster, J.M., Daly, P., Delmas, S., Pullan, S.T., Blythe, M.J., Malla, S., Kokolski, M., Noltorp, E.C., Wennberg, K., Fetherston, R., Beniston, R., Yu, X., Dupree, P., Archer, D.B., 2014. The role of carbon starvation in the induction of enzymes that degrade plant-derived carbohydrates in *Aspergillus niger*. Fungal Genet. Biol., pii:S1087-1845(14)00060-7
Vlanti, A., Amillis, S., Koukaki, M., Diallinas, G., 2006. A novel-type substrate-selectivity filter and ER-exit determinants in the UapA purine transporter. J. Mol. Biol. 357, 808–819.
Zettl, M., Adrain, C., Strisovsky, K., Lastun, V., Freeman, M., 2011. Rhomboid family pseudoproteases use the ER quality control machinery to regulate intercellular signaling. Cell 145, 79–91.

Molecular Microbiology (2015) 96(5), 927–950 ■
doi:10.1111/mmi.12982
First published online 7 April 2015

Origin, diversification and substrate specificity in the family of NCS1/FUR transporters

Emilia Krypotou,[1] Thomas Evangelidis,[2] Jacob Bobonis,[1] Alexandros A. Pittis,[3,4] Toni Gabaldón,[3,4,5] Claudio Scazzocchio,[6,7] Emmanuel Mikros[2] and George Diallinas[1]*

[1]*Faculty of Biology, University of Athens, Panepistimioupolis, Athens 15784, Greece.*
[2]*Faculty of Pharmacy, University of Athens, Panepistimioupolis, Athens 15771, Greece.*
[3]*Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), Dr. Aiguader, 88, Barcelona 08003, Spain.*
[4]*Department of Experimental and Health Sciences, Universitat Pompeu Fabra (UPF), Barcelona 08003, Spain.*
[5]*Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, Barcelona 08010, Spain.*
[6]*Department of Microbiology, Imperial College, London SW7 2AZ, UK.*
[7]*Institut de Génétique et Microbiologie, Université Paris-Sud, France.*

Summary

NCS1 proteins are H^+/Na^+ symporters specific for the uptake of purines, pyrimidines and related metabolites. In this article, we study the origin, diversification and substrate specificity of fungal NCS1 transporters. We show that the two fungal NCS1 sub-families, Fur and Fcy, and plant homologues originate through independent horizontal transfers from prokaryotes and that expansion by gene duplication led to the functional diversification of fungal NCS1. We characterised all Fur proteins of the model fungus *Aspergillus nidulans* and discovered novel functions and specificities. Homology modelling, substrate docking, molecular dynamics and systematic mutational analysis in three Fur transporters with distinct specificities identified residues critical for function and specificity, located within a major substrate binding site, in transmembrane segments TMS1, TMS3, TMS6 and TMS8. Most importantly, we predict and confirm that residues determining substrate specificity are located not only in the major substrate binding site, but also in a putative outward-facing selective gate. Our evolutionary and structure-function analysis contributes in the understanding of the molecular mechanisms underlying the functional diversification of eukaryotic NCS1 transporters, and in particular, forward the concept that selective channel-like gates might contribute to substrate specificity.

Accepted 22 February, 2015. *For correspondence. E-mail diallina@biol.uoa.gr; Tel. +30(210)7274649; Fax +30(210)7274702.

Introduction

The cellular uptake of purines and pyrimidines is mediated by specific plasma membrane transporters identified in prokaryotes and eukaryotes (De Koning and Diallinas, 2000; Pantazopoulou and Diallinas, 2007; Young *et al.*, 2013; Girke *et al.*, 2014). Nucleobase transporters also mediate the uptake of purine- and pyrimidine-based drugs widely used against different diseases, infections, cancer or as antiviral and antiprotozoan compounds (Kraupp and Marz, 1995; De Koning and Diallinas, 2000; Frillingos, 2012; Young *et al.*, 2013). Two families are highly specific for purines and pyrimidines. These are the NAT/NCS2 (Nucleobase Ascorbate Transporters or Nucleobase Cation Symporter family 2) and NCS1 (Nucleobase Cation Symporter family 1) families. The NAT family is ubiquitously present in all organisms, with few notable exceptions being some members of the Saccharomycotina, including *Saccharomyces cerevisiae* and most protozoa (Diallinas and Gournas, 2008; Gournas *et al.*, 2008; Frillingos, 2012).

Members of the second nucleobase transporter family, NCS1, are present in prokaryotes, fungi and plants, and seem to be absent from other domains of life (for reviews, see Pantazopoulou and Diallinas, 2007; Girke *et al.*, 2014). Fungal NCS1 transporters have been further classified, based on their primary amino acid sequences and specificity profiles into two sub-families, the Fcy-like and the Fur-like transporters (De Koning and Diallinas, 2000; Pantazopoulou and Diallinas, 2007). No rigorous phylogenetic analysis exists for bacterial NCS1 proteins, whereas the recently identified plant NCS1 proteins (Mourad *et al.*, 2012; Witz *et al.*, 2012; 2014; Schein *et al.*, 2013) seem to be closer to the Fur family, although a broader phylogenetic analysis is critically missing. Functionally characterised fungal Fcy-like transporters are high-affinity H^+ symporters

specific for cytosine-adenine-guanine-hypoxanthine or pyridoxine (Weber *et al*., 1990; Stolz and Vielreicher, 2003; Paluszynski *et al*., 2006; Vlanti and Diallinas, 2008; Krypotou *et al*., 2012; for a review, see Pantazopoulou and Diallinas, 2007). Fungal Fur-like proteins are also high-affinity H^+ symporters but have an entirely different and non-overlapping specificity profile from that of fungal Fcy transporters, being specific for allantoin, uracil, uridine, thiamine, nicotinamide riboside, but also for uric acid, xanthine (Jund *et al*., 1988; Yoo *et al*., 1992; Enjo *et al*., 1997; Singleton, 1997; de Montigny *et al*., 1998; Vickers *et al*., 2000; Amillis *et al*., 2007; Belenky *et al*., 2009; Hamari *et al*., 2009). Interestingly, the few functionally characterised plant NCS1 transporters, despite being more similar to fungal Fur sequences, show a specificity profile overlapping that of fungal Fcy and Fur proteins; they transport adenine, guanine, allantoin and uracil (Mourad *et al*., 2012; Witz *et al*., 2012; 2014; Schein *et al*., 2013). The two functionally known bacterial NCS1 transporters, CodB in *Escherichia coli* (Danielsen *et al*., 1992) and Mhp1 in *Microbacterium liquefaciens* (Weyand *et al*., 2008), are specific for cytosine and benzyl-hydantoin respectively. Crystal structures of Mhp1, caught in three different conformations, are available (Shimamura *et al*., 2010; Simmons *et al*., 2014), making the NCS1 family an excellent candidate to determine how substrate binding and transport specificity is determined.

It is apparent that substrate specificities within the NCS1 family cannot be predicted by amino acid sequence comparisons, simple phylogenetic analyses, or even from crystal structures. This conclusion is reinforced by a previously published phylogenetic analysis of Fur proteins in ascomycetes, which showed that the same transport activity could be catalysed by non-orthologous proteins (Hamari *et al*., 2009). More specifically, extent evidence shows that uracil and allantoin transport activities have appeared independently by convergent evolution at least three and two independent times, respectively, in *S. cerevisiae*, *S. pombe* and *A. nidulans* (Hamari *et al*., 2009).

The presence of large numbers of NCS1 paralogues of unknown function is evident in most species, especially in fungi. For example, *A. nidulans* has five Fcy-like and seven Fur-like proteins, the function of which is known only for three of them; FcyB is a cytosine-purine transporter (Vlanti and Diallinas, 2008; Krypotou *et al*., 2012), whereas FurA and FurD are allantoin and uracil transporters respectively (Amillis *et al*., 2007; Hamari *et al*., 2009). Null mutations of all other Fcy or Fur proteins do not lead to the apparent loss of any detectable transport activity (Hamari *et al*., 2009, E. Krypotou and G. Diallinas, unpublished). Importantly, although the three known *A. nidulans* transporters are absolutely conserved in all 23 species of *Aspergillus* and most ascomycetes (see later in *Results*), the orphan NCS1 genes are only partially conserved. The high number of orphan NCS1 paralogues contrasts with other transporter families analysed, such as the NAT family, which includes 1 or 2 paralogous proteins of known function in each species (Gournas *et al*., 2008). In other words, the NCS1 family seems to exhibit a striking structural and evolutionary plasticity leading to a plethora of putatively novel and unknown functions and substrate specificities.

In this article, we carry out wide-ranging rigorous phylogenetic analyses and functionally analyse all the Fur proteins of *A. nidulans*. Our results strongly suggest independent prokaryotic origins, through lateral gene transfer of the fungal and plant NCS1 homologues, expansion by gene duplication in dikarya, and absence in protozoa and metazoa. We characterise FurE as a promiscuous low-affinity transporter, FurC, FurF and FurG as minor uracil transporters, and FurB as a putative 'proto-pseudogene'. We subsequently identify, by modelling and mutational approaches, residues critical for substrate binding and transport in FurA, FurD or FurE, and present evidence for the existence of selective outward-facing gate in FurD. Our results are discussed in relation to the molecular events underlying the mechanism and evolution of substrate specificities in Fur transporters, and in particular the role of selective channel-like gates in transporters.

Results

Phylogenetics of NCS1 proteins

To gain an evolutionary insight into the NCS1 superfamily, we carried out an exhaustive phylogenomic analysis in the context of 1093 fully sequenced genomes from all life domains (see *Experimental procedures*). Profile-based searches revealed 2248 NCS1 family members in the inspected genomes, with a widespread presence in the major fungal clade Dikarya and plants, but a highly patchy distribution among other fungi and the remaining eukaryotic and prokaryotic groups (not shown). We next used sequence alignment and maximum likelihood (ML) methods to reconstruct the evolutionary relationships within this family. In the resulting phylogenetic tree (Fig. 1A) the two fungal transporter sub-families, Fur and Fcy, appeared well separated by the rest by long, highly supported branches, with most non-fungal sequences lying between these two groups, which is in agreement with previous observations (Hamari *et al*., 2009), pointing to independent origins in fungi. These two clades were highly homogeneous, with only three green algae sequences in the case of Fur and 2 sequences from *Phytophthora ramorum* in the case of Fcy, breaking their monophyly and showing patterns from recent horizontal gene transfer (HGT) from fungi to these groups. The plant NCS1 transporters, with the exception of the algal sequences men-

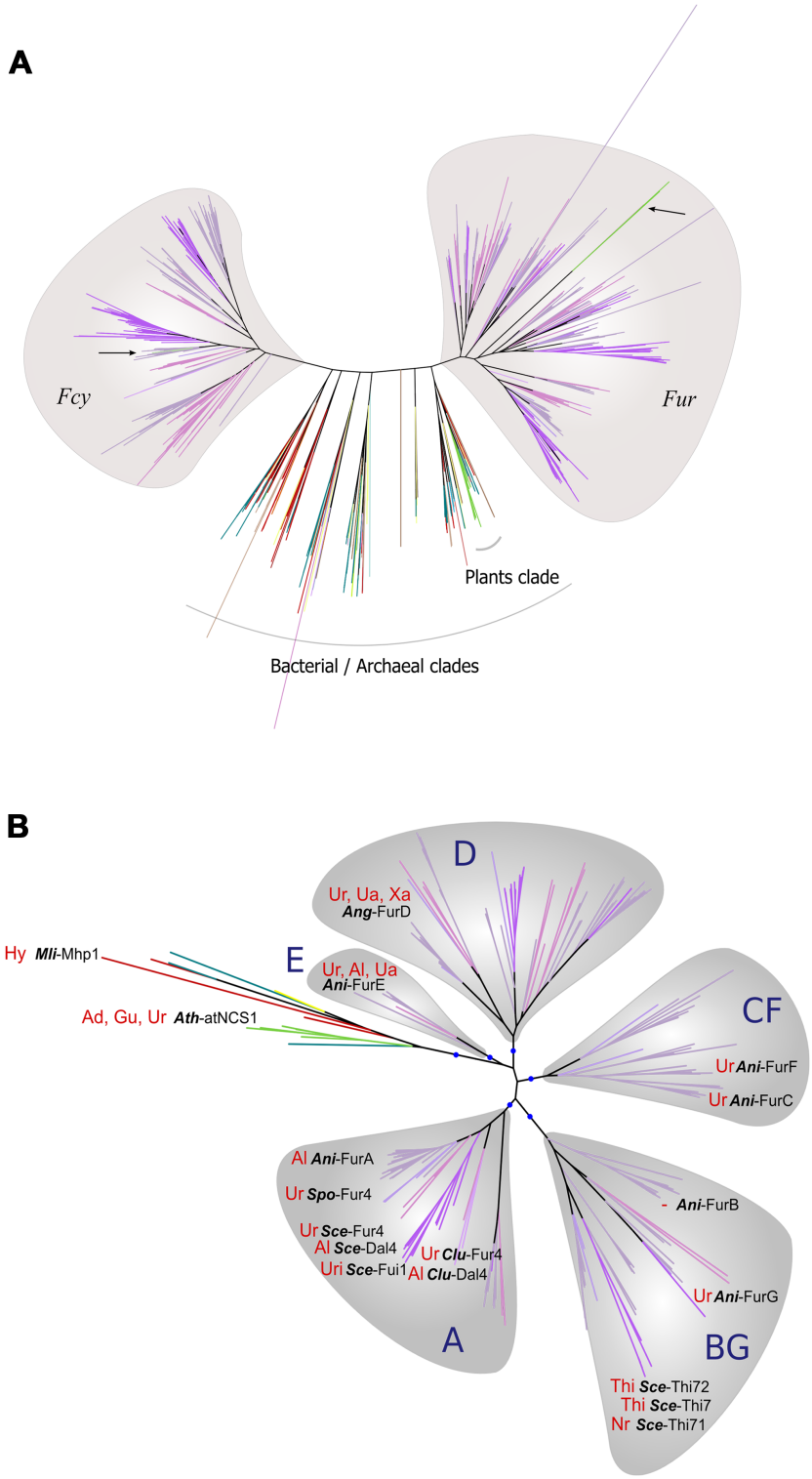


Fig. 1. Evolution of NCS1 transporter proteins.
A. Maximum Likelihood phylogenetic tree of all 2248 sequences from 1093 species from all domains, carrying the PF02133 Pfam domain (NCS1 superfamily). Fcy and Fur fungal proteins correspond to purple lines. Plant (Viridiplantae) NCS1 proteins are in green, Actinobacteria, Proteobacteria and Firmicutes are in red, blue and yellow respectively. The cases of recent horizontal gene transfer (HGT) from fungi to other groups are indicated with arrows.
B. Maximum Likelihood phylogenetic tree of the fungal Fur family of transporters (190 sequences). Colors are as in A. The sequences from plants, cluster within bacterial ones. The substrate specificities of experimentally characterised transporters are shown in red (Ur: uracil, Ua: uric acid, Al: allantoin, Uri: uridine, Xa: xanthine, Thi: thiamine, Nr: nicotinamide riboside). The protein names of those start with a three letters species code for clarity (*Ani*: *A. nidulans*, *Ang*: *A. niger*, *Sce*: *S. cerevisiae*, *Clu*: *C. lusitaniae*, *Spo*: *S. pombe*, *Ath*: *A. thaliana*, *Mli*: *M. liquefaciens*). The names of the five sub-families follow the protein names of the *A. nidulans* Fur proteins (FurA–G). The blue circles in the base of each sub-family clade indicate approximate Likelihood Ratio Test (aLRT) branch support of 100.

tioned above, are only distantly related to the fungal clade and appear associated with a prokaryotic clade, pointing to yet another independent horizontal acquisition in the hypothetical proto-plant cell. All the above observations were confirmed with a smaller sampling of species (339 sequences in 83 species), with the exception of the three putative transferred algal sequences, which moved out of the fungal Fur clade with long and poorly supported

branches (not shown). We cannot rule out the possibility that the two fungal groups result from vertical descent from ancient paralogous genes that diverged before the origin of eukaryotes. However, this would imply multiple parallel gene losses in every other eukaryotic lineage. Therefore, we conclude that the HGT scenario is more parsimonious.

We zoomed into the evolution of the fungal Fur subfamily in two steps. First, we reconstructed a phylogeny using 303 NCS1 sequences from 56 fungi sampled among all major fungal clades, four plants, the early opisthokont *Fonticula alba*, 20 bacterial genomes and the Mhp1 hydantoin transporter from *M. liquefaciens*. From this phylogeny, we extracted the 190 sequences from the monophyletic fungal Fur clade and its nearest neighbouring clade, plus the Mph1 sequence and reconstructed a new gene tree. Based on the branching patterns in the tree, and the known substrate specificities of characterised transporters, we defined five sub-families of fungal Furs (Fig. 1B). All of them were present throughout Dikarya and one of them was also present in the early-diverging aquatic fungus *Gonapodya prolifera* (Fig. 2). The presence of sequences from most dikaryal main groups in all the clades supports their emergence by gene duplication in the proto-Dikaryon or earlier, followed by subsequent independent losses in specific lineages. Gene duplication is considered one of the major sources for functional innovation (Ohno, 1970). Our phylogenetic analysis revealed that duplication has been rampant in the Fur family (Fig. S1). We used a phylogeny-based method to detect and date gene duplication events (Huerta-Cepas *et al*., 2007) and calculated the number of observed duplications for the family in each of the lineages in our taxonomy (Fig. 2). We observed 57 duplication events in a phylogeny of 56 species, with six, eight and nine ancient duplications mapped to the bases of fungi, Dikarya and Lleoteomyceta (Pezizomycotina) respectively. The largest duplication wave corresponds to 11 events specific to *Lipomyces starkeyi*, a basal species of the Saccharomycotina.

Overexpression of orphan fur *genes reveals cryptic transport activities and novel substrate specificities*

The FurA and FurD proteins have been functionally analysed in detail and shown to act as highly specific allantoin or uracil transporters respectively. FurD has very high affinity for uracil and 5-fluorouracil (5-FU) (0.5 μM) and medium affinity for uric acid and xanthine (~100 μM) (Amillis *et al*., 2007). No kinetic analysis was possible with FurA, as there is no radiolabelled allantoin commercially available. Genetic deletions of all other *fur* genes, namely *furB*, *furC*, *furE*, *furF* and *furG*, showed no distinct phenotype, when scored on various purines as nitrogen source or on purine or pyrimidine toxic analogues (Hamari *et al*., 2009). In addition, previous northern analysis (Hamari *et al*., 2009) and transcriptomics (Sibthorp *et al*., 2013) show that all *fur* genes, except *furD* and *furA*, are very lowly transcribed under most conditions studied.

To identify the possible transport activity of the five orphan Fur proteins, we overexpressed them using the *gpdA* strong promoter (Punt *et al*., 1990), in a genetic background lacking not only *furA* and *furD*, but also all transporters specific for nucleobases and nucleosides (*ΔfurD ΔfurA ΔfcyB ΔuapA ΔuapC ΔazgA ΔcntA*). Gene overexpression in this strain, called Δ7, permits the direct assessment of cryptic transport activities related to nucleobases or nucleosides (Krypotou and Diallinas, 2014).

Figure 3A shows a growth test of selected transformants overexpressing Fur transporters, compared with appropriate isogenic control strains. The Δ7 negative control strain cannot grow on purines as sole nitrogen sources, shows residual growth on allantoin due to a non-identified transporter and is resistant to 5-FU, 5-fluorocytosine (5-FC), 5-fluorouridine (5-FUd) or oxypurinol. Overexpression of FurD led to increased sensitivity to 5-FU and increased capacity for growth on uric acid or xanthine. Overexpression of FurA led to a compact morphology on allantoin, apparently due to over-accumulation of this metabolite, which is toxic (Darlington and Scazzocchio, 1967). FurD overexpression also led to highly increased sensitivity to 5-FC, 5-FUd and oxypurinol and that of FurA to 5-FU and moderately to 5-FC. This test revealed an enlarged substrate specificity profile of these transporters. FurE overexpression conferred increased capacity for growth on uric acid or allantoin, but not on xanthine, and led to moderately increased sensitivity to 5-FU, 5-FC, 5-FUd and oxypurinol. FurC, FurF and FurG overexpression led to increased sensitivity to 5-FU, albeit at different levels (FurC > FurG > FurF). Finally, FurB scored exactly as Δ7, revealing no transport activity related to its overexpression.

Several Fur transporters contribute to uracil or 5-FU accumulation

In order to further establish the physiological contribution of the different Fur transporters (except FurB, which has no apparent transport function, see also later) in the uptake of relevant substrates, and especially of uracil and 5-FU, which seem to be common substrates for all Fur transporters, we constructed combinations of *Δfur* null mutants among themselves and with the *pyrG89* auxotrophic marker, using standard genetic crossing. The *pyrG89* mutation blocks pyrimidine biosynthesis and leads to uracil requirement (Oakley *et al*., 1987) so that growth of *pyrG89* mutants depends on the efficiency of uracil accumulation from the nutrient. Thus, *pyrG89* serves as an ideal genetic background for estimating the contribution of different Fur transporters in uracil uptake. Highlights of growth

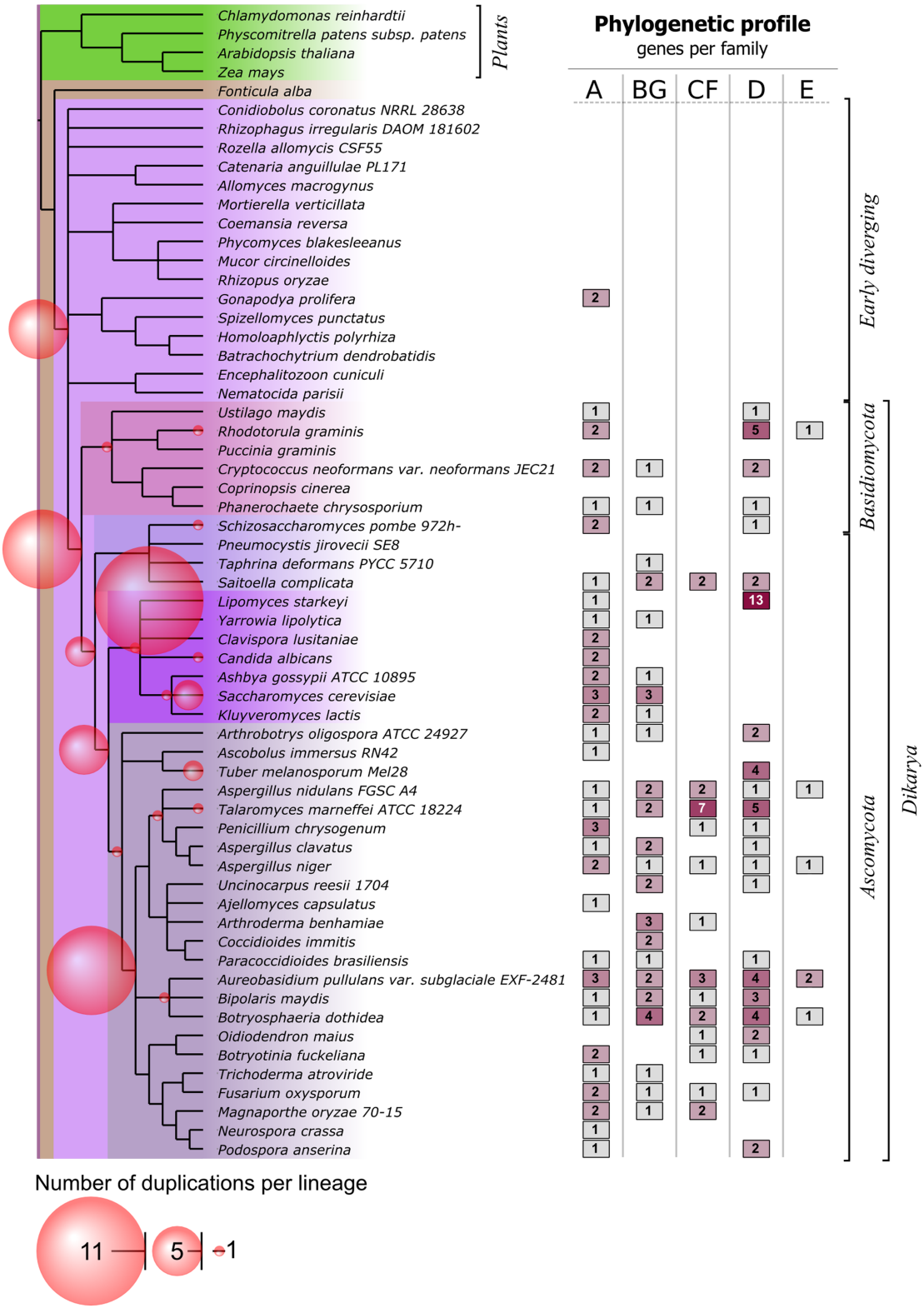


Fig. 2. Evolution of Fur transporter proteins. NCBI taxonomy tree structure of the 61 eukaryotic species included in the detailed analysis of the Fur family (56 fungi, four plants and *Fonticula alba*) and the distribution of the sub-family members among the fungal species. Sphere diameters are proportional to the number of duplications at a given lineage as inferred from the gene tree topology (see Huerta-Cepas *et al.*, 2007). Note that, due to subsequent gene loss, the number of current genes does not directly correspond to the number of duplications. In the different columns, the number of gene copies per family and per species is indicated.

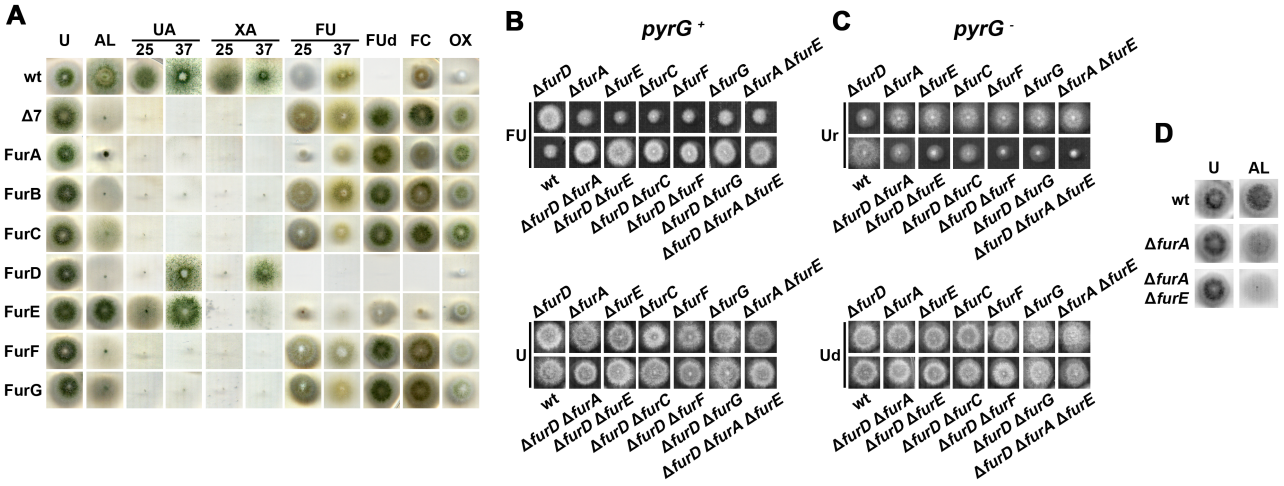


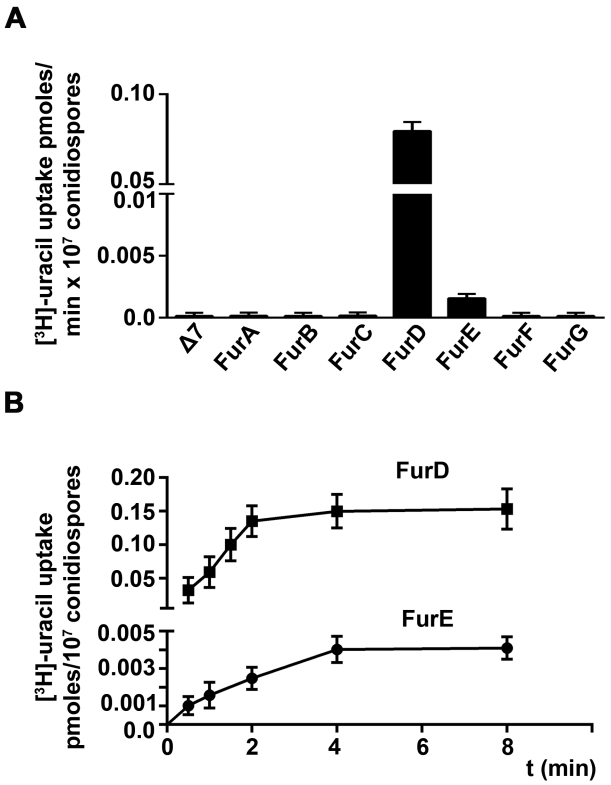
Fig. 3. Functional analysis of the Fur transporters.
A. Growth tests of transformants overexpressing Fur transporters. Nitrogen sources or toxic analogues used are: urea (U), allantoin (AL), uric acid (UA), xanthine (XA), 5-fluorouracil (FU), 5-fluorouridine (FUd), 5-fluorocytosine (FC) and oxypurinol (OX). Growth tests shown were at 25°C, or at 37°C when indicated. Toxic analogues were used in media containing urea as a nitrogen source. A wild-type strain was used as a positive control and Δ7 as a negative control (see text).
B–C. Growth tests of simple, double or triple null *Δfur* mutants. Panel B shows *pyrG*$^{+}$ strains tested in 5-fluorouracil (FU) in presence of urea (upper). Growth on urea is shown as a control (lower panel). Panel C shows *pyrG89* (*pyrG*$^{-}$) strains grown in media supplemented with 2.5 mM uracil (Ur), or 10 mM uridine (Ud) as a control. In both cases urea was used as sole nitrogen source. Growth tests shown are at 25°C.
D. Growth tests of single *ΔfurA* and double *ΔfurA ΔfurE* null mutants in the presence of 2 mM allantoin (Al) or urea (U) as sole nitrogen sources. Growth tests shown are at 25°C. Wt strains are also shown in B, C and D as controls.

phenotypes, relevant to uracil or 5-FU uptake, of selected single, double or triple mutants are shown in Fig. 3B and C. Results in the upper panel of Fig. 3B show that FurD has the highest contribution in 5-FU uptake (i.e. *ΔfurD* has highest resistance to 5-FU toxicity), whereas all other Fur transporters (A, C, E, F and G) do not seem to contribute to 5-FU accumulation, as the single mutants are more or less equally sensitive to 5-FU with a wild-type control. Furthermore, the double mutants *ΔfurEΔfurD*, *ΔfurAΔfurD*, *ΔfurCΔfurD*, *ΔfurFΔfurD* or *ΔfurGΔfurD* are also similarly resistant to 5-FU compared with the *ΔfurD* single mutant. In other words, despite what has been revealed by overexpression, when Fur transporters, other than FurD, are expressed from their native promoters, they do not contribute to 5-FU sensitivity/resistance. This also highlights the importance of overexpression in identifying the biochemical function of these transporters. Results in Fig. 3C proved more revealing in respect to the individual contribution of different Fur transporters in uracil uptake. Through this test, it became apparent that at least FurF and FurE contribute to uracil uptake, because the double mutants *ΔfurEΔfurD* and *ΔfurFΔfurD* grew slightly less than the *ΔfurD* single mutant in uracil-containing media. In addition, the triple mutant *ΔfurAΔfurEΔfurD* grew less than *ΔfurEΔfurD* and all other single and double mutants tested, revealing the minor contribution of also FurA in uracil uptake.

Overall, our results strongly suggest that: (i) FurD is the major transporter for uracil and 5-FU accumulation, but all Fur proteins, except FurB, have the capacity of mediating uracil or 5-FU transport, when overexpressed, in the order FurD ≫ FurE, FurA > FurF, FurC, FurG. (ii) FurE, FurA, FurF, FurC and FurG contribute very little, if not at all, to uracil or 5-FU transport, when expressed from their native promoters. (iii) When biosynthesis of pyrimidines is genetically blocked (*pyrG89*), FurE, FurA and FurF have low, but detectable, contribution to uracil transport. (iv) FurA is the main allantoin transporter, whereas FurE acts as a secondary allantoin transporter. This is as also supported by the observation that the double mutant *ΔfurA ΔfurE* grows less than *ΔfurA* single mutant on allantoin-containing media (Fig. 3D). (v) FurE has a broader specificity profile from other Fur paralogues.

FurE is a novel promiscuous nucleobase transporter

We measured Fur-mediated uptake of radiolabelled uracil in all strains overexpressing Fur transporters. We could not detect any measurable uracil uptake mediated by FurB, FurC, FurF, FurG or FurA, the latter being known to be specific for allantoin. In contrast, we could detect low capacity uracil transport mediated by FurE (Fig. 4A). The lack of uracil transport by FurB is expected given the total lack of any detectable relevant phenotype in growth tests. In contrast, the fact that we could not detect any uracil transport activity mediated by FurC, FurF, FurG or FurA, somehow contradicts the contribution of these transporters in 5-FU sensitivity. This indicates that these transport-



C

FurE specificity profile

Inhibitor (1mM)	[^{3}H]-uracil uptake (%)	$K_{m/i}$ (μM)
-	100	-
UR	45±5	~1000
AL	40±8	~1000
UA	27±3	20
XA	100±5	>1000
HY	74±10	>1000
TH	84±10	>1000
CCCP	0	-

Fig. 4. Kinetic analysis of FurE transporter.
A. Uracil transport assays. Initial uptake rates of [^{3}H]-radiolabelled uracil (0.1 μM) for 1 min were measured in transformants overexpressing Fur transporters. Standard deviation is depicted with error bars.
B. Time course of FurE-mediated [^{3}H]-uracil uptake (0.1 μM). Time course of FurD is shown as a control. Standard deviation is depicted with error bars.
C. Specificity profile of FurE. Middle column: FurE-mediated [^{3}H]-uracil uptake in the presence of excess (1 mM) unlabelled uracil (UR), allantoin (AL), uric acid (UA), xanthine (XA), hydantoin (HY) or thymine (TH), or in the presence of H^+-uncoupler carbonylcyanide chlorophenylhydrazone (30 μM, CCCP). Standard deviation values are shown. Right column: Approximate $K_{m/i}$ (μM) values of putative inhibitors, determined through IC_{50} values by the Cheng and Prusoff equation. As we cannot determine the exact K_m value of uracil for FurE, the values given are approximate (see text). We consider, however, that as the concentration of labelled substrate used (uracil) is much lower than the notional K_m value of uracil, the Cheng and Prusoff equation is appropriate in spite of the intrinsic limitations of the method (for more details, see *Experimental procedures*). The graphs from which the $K_{m/i}$ values are calculated are shown in Fig. S5). ~1000 or > 1000 means that at K_i values (inhibition close to 50%) are close to 1 mM or > 1 mM. In such cases, the low solubility of nucleobases (< 1 mM) does not allow for a more accurate estimation of binding constants. Results are averages of at least three independent experiments in triplicate for each concentration point. Standard deviation was < 20%.

ers might have very low substrate affinities and transport capacities for uracil, which is technically impossible to measure.

FurE-mediated uracil transport was time dependent, reaching a plateau in 4 min, similar to most other transporters studied so far in *A. nidulans* (Fig. 4B). FurE activity was inhibited by excess allantoin, uric acid and uracil (Fig. 4C). The profile and degree of FurE inhibition suggested that FurE might be a low-affinity transporter, with binding affinities following the order uric acid > allantoin > uracil. Although we could not measure the exact K_m value of FurE for uracil, which is necessary for rigorous estimation of K_i values of other putative substrates using the Cheng and Prusoff equation (1973), we could still establish approximate K_is based on apparent IC_{50} values. FurE seems to have very low-affinity binding capacity for uracil, allantoin, thymine and hydantoin (≥ 1 mM), but surprisingly relatively high affinity for uric acid (20 μM) (see Fig. 4C). This values are obviously approximate, and only give an order of magnitude but are valid for comparative purposes with mutant versions of the FurE protein. Overall, judging from both growth tests and uptake assays, FurE shows (i) relatively high affinity and transport capacity for uric acid, (ii) low affinity but high transport capacity for uracil and allantoin and (iii) no significant binding affinity and transport capacity for xanthine, adenine, guanine, hypoxanthine, thymine or hydantoin. Growth tests on relevant toxic analogues further suggested that FurE transports with high capacity cytosine (highly sensitive to 5-FC) and has moderate transport capacity for uridine (5-FUd) and oxypurinol (see Fig. 3A). Finally, FurE was also severely inhibited by a standard proton uncoupler, showing that it acts as a proton symporter, analogously to all *A. nidulans* transporters characterised so far.

Fur transporters, with the exception of FurB, are localised in the plasma membrane

None of the Fur proteins of *A. nidulans* has been previously analysed for subcellular localisation. We constructed C-terminally GFP-tagged chimeric Fur proteins and tested their subcellular localisation. The GFP versions of Fur proteins were expressed from the *gpdA* promoter and all proved functional, as judged from growth tests of the relevant transformants (not shown), except

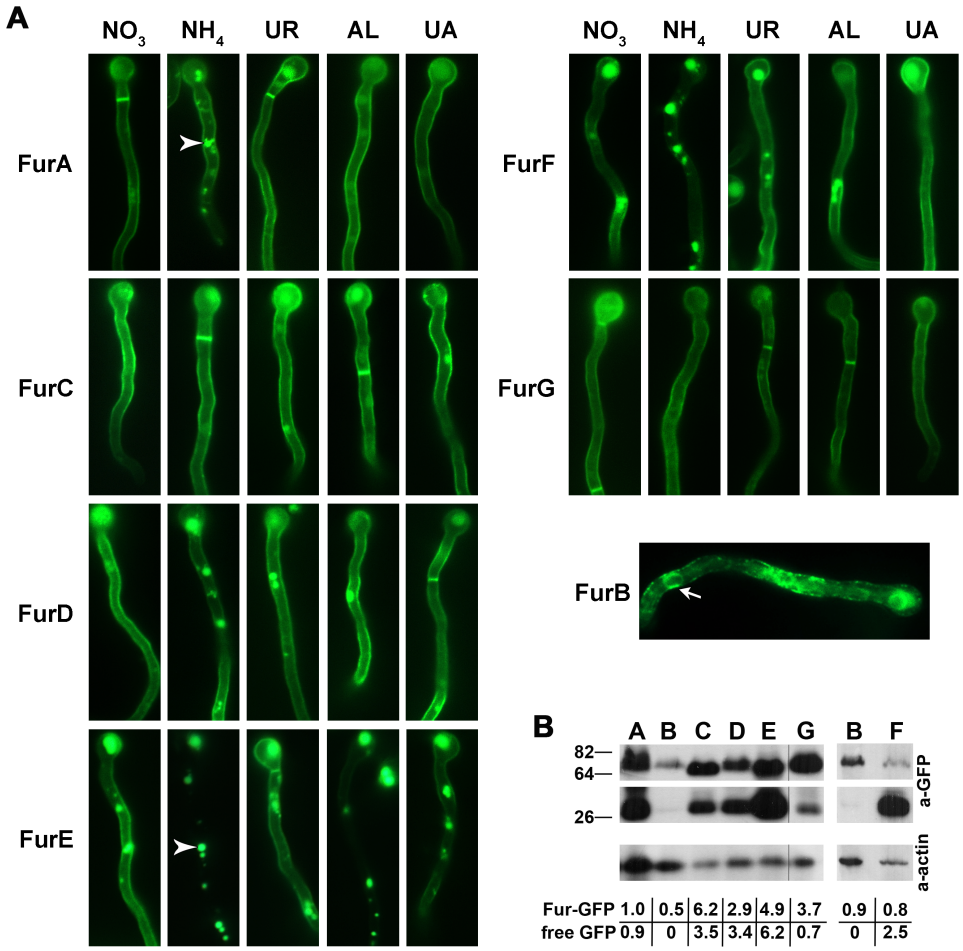


Fig. 5. *In vivo* subcellular localisation and protein expression levels of Fur transporters.
A. Epifluorescence microscopy of GFP-tagged Fur proteins expressed from the *gpdA* promoter. Notice that in all cases, except FurB, GFP fluorescence labels primarily the periphery of cells, as expected for plasma membrane (PM) transporters. FurB-GFP is localised in cytoplasmic structures and foci beneath the PM. *Arrows* and *arrowheads* depict vacuoles and perinuclear endoplasmic reticulum (ER) membrane rings respectively. (UR: uracil, AL: allantoin, UA: uric acid). Cultures were grown in MM for 14–16 h at 25°C.
B. Western blot analysis of total protein extracts of strains expressing GFP-tagged Fur proteins using an anti-GFP antibody. At the left panel, all lanes correspond to the same blot. FurF is shown in a separate blot with FurB as an internal control. Antibody against actin was used as an internal marker for equal loading. Molecular weight markers in kDa are shown. Relative amounts of Fur-GFP and free-GFP, normalised to actin protein levels, are shown below the blots (see *Experimental procedures*). Cultures for total protein extraction were grown in MM supplemented with urea at 25°C for 16 h.

from FurB-GFP, which had no apparent function to be tested. Figure 5A shows that FurA-GFP, FurC-GFP, FurD-GFP, FurE-GFP, FurF-GFP and FurG-GFP are all plasma membrane proteins. In all cases, Fur-GFP fluorescence associates nearly entirely with the periphery of mycelia and septa and secondarily with vacuoles. This is the normal picture obtained with most *A. nidulans* transporters functionally tagged with GFP. In contrast, FurB-GFP labels mostly the cytoplasm, characteristic perinuclear ER rings and prominent cortical foci which seem associated with the inner side of the plasma membrane. In addition, the level of fluorescence is lower than all, except FurF, Fur proteins.

Figure 5A also reveals that the Fur proteins show differential sensitivity towards endocytic signals. FurC and FurG do not show either ammonium- or substrate-elicited internalisation (Gournas *et al.*, 2010; Karachaliou *et al.*, 2013). On the other hand, FurE is highly sensitive to substrate-elicited and especially to ammonium-elicited endocytosis. Substrate-elicited endocytosis is triggered by uric acid and allantoin and less efficiently by uracil. FurA, FurD and FurF are sensitive to ammonium endocytosis, the latter two being also moderately sensitive to substrate-elicited endocytosis. The response of FurA, FurD, FurE and FurF transporters to substrate-elicited endocytic turnover is in agreement with their established substrate specificity profile.

Western blot analysis, at the same conditions as those used for epifluorescence microscopy analysis, using anti-GFP antibody (for details, see Gournas *et al.*, 2010;

Karachaliou *et al*., 2013), confirmed that all GFP-tagged Fur proteins are rather stable proteins of the expected molecular weight (Fig. 5B). FurA, FurC, FurD and FurE showed, more or less, the same degree of turnover, as evidenced by the amount of free GFP detected (free GFP levels originating from GFP-tagged transporters is a measure of vacuolar degradation). FurG was significantly more stable compared with other Furs, as very little free GFP was detected.

The relative amounts of free GFP detected in westerns agrees with the detected vacuolar fluorescence. For FurF, lower amounts of the intact fusion protein are very probably due to increased vacuolar degradation, as assessed by a dramatic increase in free GFP and by the relevant microscopic analysis. In contrast, in FurB, the low steady state levels of the GFP-tagged protein are not associated with an increase in free GFP. This indicates that FurB is synthesised at low rates or rapidly undergoes proteolysis.

Structural models of FurA, FurD and FurE

We constructed structural models of FurA, FurD and FurE based on the crystal structure of the Mhp1 benzylhydantoin permease from *M. liquefaciens* (PDB entry: 2JLO). The Fur proteins and Mhp1 share a significant sequence similarity (22.1–24.2% identity in 421 amino acid overlap, out of the 501 amino acids of the shorter Mhp1), which is adequate for sustaining solid homology models of FurA, FurD and FurE. The similarity of Furs with Mhp1 is significantly higher than the similarity of Fcy-like proteins to Mhp1 (~18%; Krypotou *et al*., 2012). The final sequence alignments used to build the Fur models are shown in Fig. S2.

The 3D structural models of the Fur proteins are shown in Fig. 6 and correspond to substrate-occluded conformers. The models also show the approximate substrate binding sites, highlighted in zoom-in panels. The topology of all transporters consists of two distinct domains, a compact core consisting of segments TMS1-10 and a C-terminal domain comprising TMS11-12. The distribution of the ionised residues on the protein surface is as expected, as most of these are positioned either at the cytoplasmic and periplasmic sides or along the protein pore in the protein interior. Positive charges are mostly concentrated in the cytoplasm-facing loops. The core of the transporters is subdivided in two inverted repeats, made of TMS1-5 and TMS6-10, arranged in twofold pseudo-symmetry axis, running parallel with the membrane plane. The two repeat units are completely intertwined giving two topologically distinct subdomains constituted by TMSs 1, 2, 6 and 7 (bundle motif) and TMSs 3, 4, 8 and 9 (hash motif), respectively, linked with helices TMS5 and TMS10. As it will be shown below, the substrate-binding site is located in the space between the two subdomains of the core (see below Fig. 7).

Substrate docking and Molecular Dynamics (MD) reveal residues critical for substrate binding and/or specificity in FurA, FurD or FurE

Using the Fur models, we addressed aspects of substrate recognition using induced fit docking (IFD) calculations and/or MD. In each case we investigated the possible interactions of the relevant physiological substrates. For FurA, this is allantoin, for FurD uracil and uric acid, and for FurE uracil, uric acid and allantoin. Figure 7 shows the modelled interactions of the Fur proteins with their recognised physiological substrates. The most probable substrate orientations are shown. Details on how these models are generated are provided in *Experimental procedures*.

FurD, FurA and FurE interactions with their respective substrates were identified by induced fit docking calculations (Fig. 7A). FurD interacts with uracil through two apparently strong H-bond interactions with Gln134 and Asn341. Given that the K_m of uracil is high (0.5 μM), more interactions with amino acids might be involved. This is confirmed using a MD approach described below (see also Fig. 7B). FurD also interacts with uric acid (K_m ~100 μM) via H bonds with Ala56, Asn249, Asn341 and π-π stacking with Trp130. Gln134, may also contribute to binding. FurA interacts with allantoin (no K_m available) mainly through H-bonds via the side-chains of Asn58, Gln136, Asn349 and Asp53 (neutral aspartate), but also with the backbone of Asn252. In this modelled state, allantoin has a net charge of +1 and Asp53 is protonated, suggesting that this may be an intermediate state in the substrate/proton symport of FurA. FurE interacts with uric acid (K_i ~ 20 μM), allantoin (~1 mM) or uracil (~1 mM). Interactions with uric acid involve primarily a π-π stacking interaction with Trp130 side-chain, an H-bond with its backbone, but also an H-bond with Asn341 side-chain. The negatively charged uric acid can also potentially form an H-bond with Gln134 and a salt bridge with Lys252 side-chain. In most of the docking poses, allantoin is stabilised in the substrate cavity by H-bonds with Ser54, Ser56 and Asn341 side-chains. There is also a large number of docking poses where allantoin forms H-bond interactions with the side-chains of Gln134 and Gln379 (not shown). Finally, uracil forms H-bonds with a wide range of residues, the most frequent of which were Gln134 and Asn341.

FurD-substrate interactions were also derived from the MD trajectories (Fig. 7B). The first 150 ns were discarded, as they corresponded to protein relaxation time (red arrows in Fig. S3), and only the last 250 ns were used in the analysis. According to these data, FurD interacts directly with uracil with constant H-bonds mediated by Gln134 and Asn341 side-chains and Leu253 backbone (Table S1). It

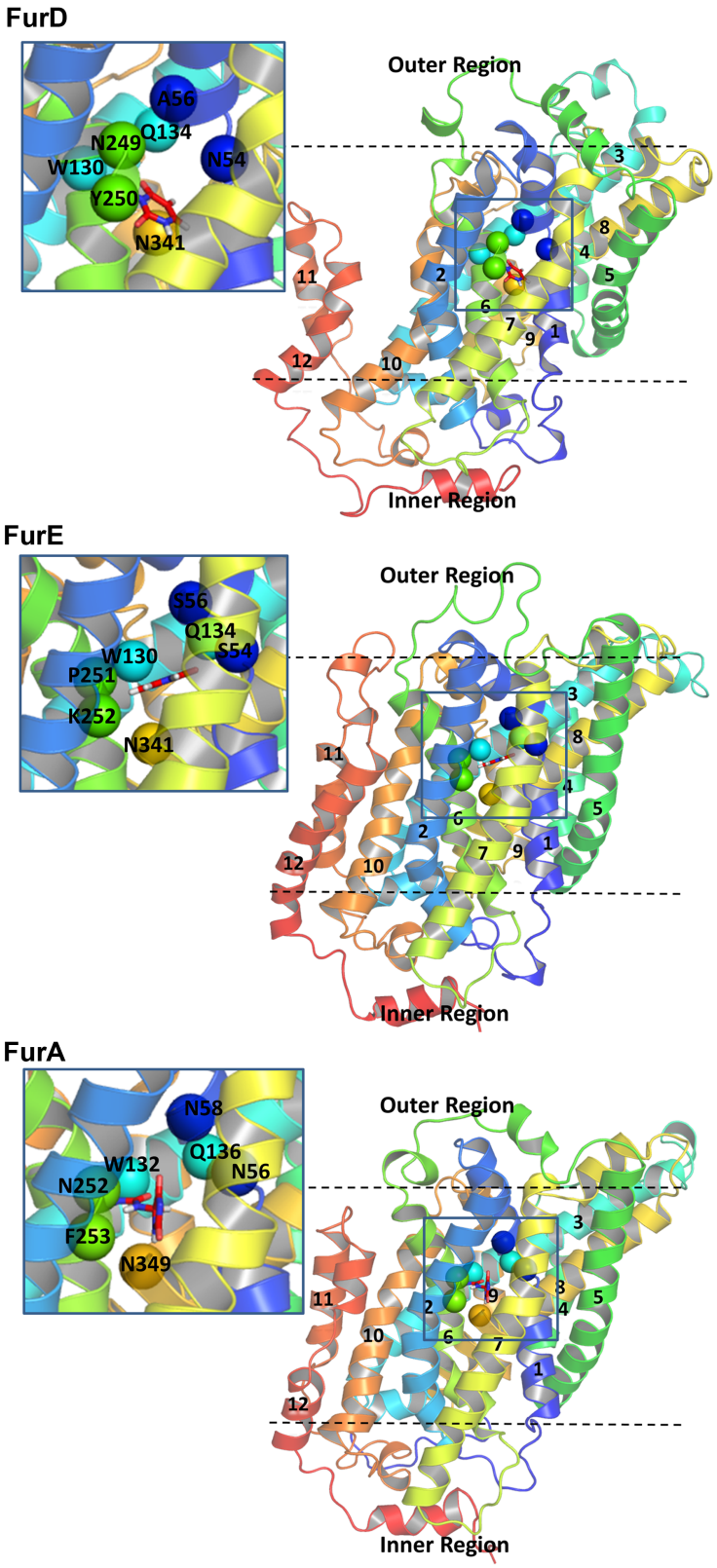


Fig. 6. 3D structural models of Fur proteins. Upper panel: topology of FurD occluded with uracil and orientation with respect to the membrane. The structure shown is taken from the MD trajectory upon relaxation of the protein in membrane-aqueous environment. Middle panel: topology of FurE occluded with uracil and orientation with respect to the membrane, as predicted by OPM server. Lower panel: topology of FurA occluded with allantoin and orientation with respect to the membrane, as predicted by OPM server. Models were built on the basis of alignments with Mhp1. The planes of the membrane are depicted with dashed lines. The substrate-binding site including residues of TMS1, TMS3 TMS6 and TMS8 is located in the space between the hash and the bundle domains. Major residues of the substrate binding site are designated with spheres. In all cases, the putative substrate binding cavity is highlighted in zoom-in panels.

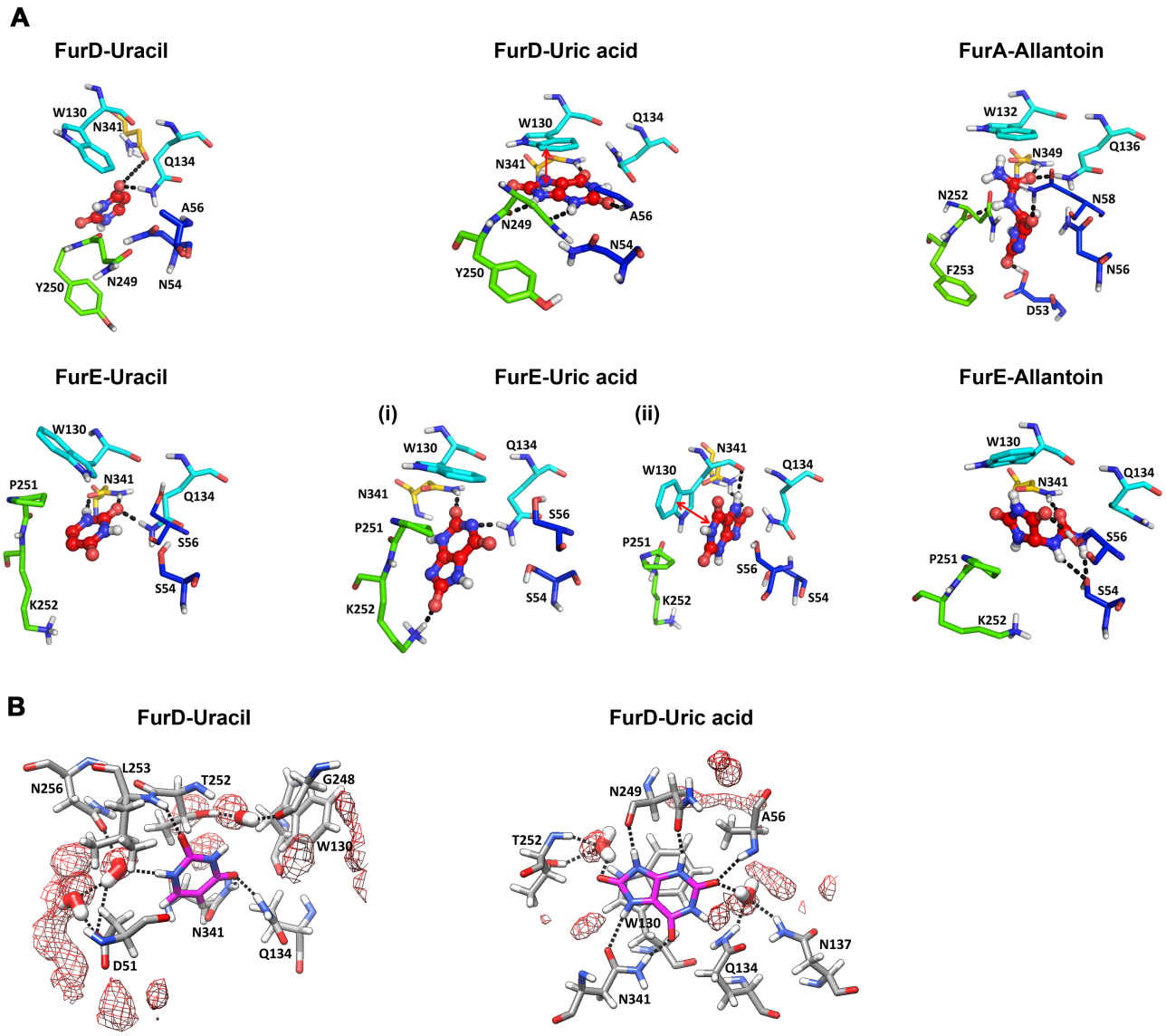


Fig. 7. Modelled substrate interactions with the binding sites of Fur proteins.
A. Predominant Docking Poses from IFD calculations. Upper panels: FurD with uracil (left) and uric acid (middle), FurA with R-allantoin (right). Lower panels: FurE with uracil (left), uric acid (i) and (ii) (middle) and R-allantoin (right). Similar interactions were observed for S-allantoin (results not shown). Hydrogen bonds are depicted in black dashes, and π-π-stacking is highlighted with two-headed red arrows.
B. Representative snapshots from MD trajectory of FurD with uracil (left) and uric acid (right). Water oxygen density is depicted in mesh form and the water molecules that form stable water bridges are shown in black dashes.

also interacts indirectly through infrequent water bridges mediated by Asp51 and Asn256 side-chains (Table S2). In addition, the side-chain of Thr252 is stabilised in a position to interact with uracil through a constant water bridge connecting its side-chain with Gly248 backbone. Another, less frequent water bridge, between Asp51 and Asn256, may provide extra stability to the protein-substrate binding. FurD also interacts directly with uric acid through a π-π stacking interaction with Trp130 and five frequent H-bonds, mediated by Asn249 side-chain and backbone, Asn341 side-chain, and the Ala56 backbone (Table S3). Uric acid binding is further stabilised by frequent water bridges connecting it with Thr252 side-chain and backbone, Gln134 side-chain, as well as less frequently with Asn137 side-chain (Table S4). It should be highlighted that although we used a physiological NaCl concentration, unlike Mhp1, no Na^+ or Cl^- ion was localised in proximity to the substrate binding site during the two 400 ns long simulations with ligands, as indicated by the respective density maps (data not shown). Overall, the MD approach

confirmed and extended the results obtained with docking studies concerning FurD-substrate interactions.

These results establish a small number of amino acid residues as primary elements of a single major substrate binding site built by TMS1, TMS3, TMS6 and TMS8 in all three Fur transporters analysed. Two residues, Gln in TMS3 and Asn in TMS8, are involved in strong H-bonds in all cases. The Trp residue in TMS3 seems to have a specific role in uric acid binding. The interactions with TMS1 and TMS6 are more variable, pointing to a role in determining substrate specificity (see also later).

Mutational analysis identifies residues critical for substrate binding and specificity in FurD and FurE

Results derived from homology modelling, docking and MD approaches suggested that specific residues in TMS1, TMS3, TMS6 and TMS8 of Furs are directly involved in substrate binding. In order to test experimentally the functional role of these residues and understand further the molecular basis of substrate specificity and differences among Furs, we constructed a series of relevant mutations in FurD and FurE and analysed functionally the corresponding mutants. We did not make mutations in FurA as this transporter cannot be analysed by transport assays.

The mutations made in FurD were N54A, T57A, W130A, Q134E, Q134A, N249A, Y250A, Y250F, Y250W, Y250K and N341A. The rationale for these mutations was to replace amino acids located at bonding distance from substrates according to the docking analysis, but also neighbouring conserved residues, with Ala residues or residues differentially conserved among NCS1 members. For example, Q134E was made because plant NCS1 members have a Glu residue at this position. Similarly, substitutions Y250F, Y250W and Y250K replaced the FurD residue (Tyr) with those found in FurE or FurA or plant NCS1. In FurE, we mutated solely residues, at bonding distance from substrates that are significantly different from the corresponding residues in FurD and FurA (Fig. S4). The substitutions made (S54N/S56N, S54N/S56A, P251N, K252F and P251N/K252F) introduce residues found in FurD, FurA or other functionally characterised NCS1 transporters. All mutant versions of FurD or FurE expressed from the *gpdA* promoter were introduced in the Δ7 mutant strain and selected transformants arising from single-copy plasmid integration events were analysed on different purines as sole nitrogen sources and on relative toxic analogues (for details, see *Experimental procedures*).

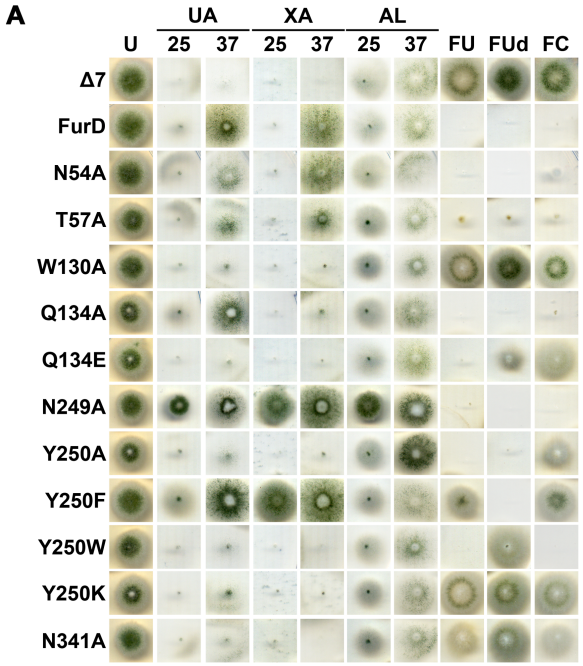
The analysis of FurD mutants is shown in Fig. 8A. As expected, all strains showed normal growth on nitrogen sources other than purines (e.g. urea). The functionality of FurD was reflected as conferring growth on uric acid or xanthine at 37°C and full sensitivity on 5-FU, 5-FC or 5-FUd. Based on this, Ala substitutions in Asn54, Thr57, Trp130 and Asn341, scored as either complete or partial loss-of-function mutations. Mutations in Gln134 seem to differentially affect FurD specificity. In particular, Q134A has led to loss of xanthine uptake, whereas Q134E led to loss of uptake of both uric acid and xanthine, and to some resistance to 5-FC and 5-FUd, but not to 5-FU. Mutation N249A conferred increased capacity for growth on uric acid, xanthine and allantoin, as well as high sensitivity on 5-FU, 5-FC and 5-FUd. The neighbouring mutation Y250A also conferred a capacity for allantoin uptake, despite loss of uric acid and xanthine uptake. Mutant Y250A also conserved 5-FU, 5-FUd and some 5-FC sensitivity, apparently due to uptake of these drugs. Other substitutions in Tyr250 scored mostly as general loss of function mutations. Overall, residues Asn54, Thr57, Trp130 and Asn341 were shown, within the limit of growth tests, to be necessary for FurD transport activity, whereas Gln134, Asn249 and Tyr250 proved critical for specificity.

For FurE (Fig. 8B), the functionality of the wild-type transporter was reflected as conferring growth on allantoin and uric acid (mostly at 37°C), and high sensitivity to 5-FU, compared with the Δ7 negative control strain. The double substitutions S54N/S56N and S54N/S56A, mimicking FurA and FurD, respectively, scored as loss-of-function mutations. P251N seemed to be a specificity mutation, as it scored like a partial loss-of-function mutation for allantoin and mostly uric acid transport, but also led to increased 5-FU, 5-FUd and 5-FC sensitivity. K252F is a gain-of-function mutation, increasing the uptake of most substrates, including xanthine. The double substitution P251N/K252F is a loss-of-function mutation for the transport of uric acid, allows allantoin transport, but leads to increased sensitivity to 5-FC. The profile of FurE mutants strongly suggested that Pro251 and Lys252 are important specificity determinants.

Kinetic analysis of mutants: residues determining substrate binding affinities and specificity

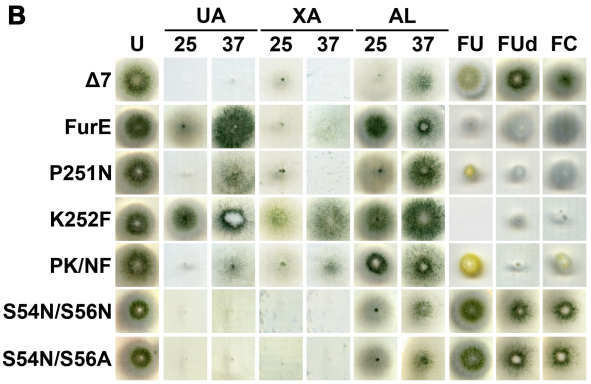
We performed comparative radiolabelled uracil uptake measurements in FurD and FurE mutants (Fig. 8C and D). In FurD, several mutants (N54A, T57A, W130A, Q134E, Y250A, Y250K and N341A) showed significantly reduced or undetectable rates of uracil uptake (0–25% of the wild-type). The rest (Q134A, N249A, Y250W and Y250F) conserved > 75% of the wild-type FurD transport rate. There was good agreement of FurD transport activities measured and FurD-mediated growth phenotypes at 37°C (no uptake measurements were performed at 25°C).

To further investigate the effect of specific mutations on substrate binding and transport, we measured K_m or K_i values of uracil, uric acid, xanthine, allantoin, hydantoin



C

	V (%)	K_m (μM) UR	K_i (μM) UA	XA	AL	HY	TH
FurD	100	0.5	99	100	>1000	40	3
N54A	5 ± 2	40	500	500	>1000	800	-
T57A	15 ± 5	0.5	18	110	>1000	-	-
W130A	0	n.m	n.m	n.m	n.m	n.m	n.m
Q134A	90 ± 5	2	66	140	>1000	100	2
Q134E	8 ± 3	8	216	300	>1000	500	40
N249A	75 ± 5	2.4	510	190	>1000	250	5
Y250A	25 ± 2	20	>1000	>1000	>1000	>1000	130
Y250F	90 ± 3	4.5	194	300	>1000	67	6
Y250W	80 ± 6	3	10	200	>1000	-	1
Y250K	0	n.m	n.m	n.m	n.m	n.m	n.m
N341A	0	n.m	n.m	n.m	n.m	n.m	n.m



D

	V (%)	K_m (μM) UR	K_i (μM) UA	XA	AL	HY	TH
FurE	100 ± 2	~1000	20	>1000	~1000	>1000	>1000
P251N	2000 ± 15	27	80	>1000	18	97	140
K252F	2000 ± 25	8	80	~1000	56	130	117
PK/NF	1280 ± 5	3	250	800	330	40	5
S54N/S56N	0	n.m	n.m	n.m	n.m	n.m	n.m
S54N/S56A	0	n.m	n.m	n.m	n.m	n.m	n.m

Fig. 8. Functional analysis of FurD and FurE mutants. Growth tests of FurD (A) and FurE (B) mutants. Details are as in Fig. 3. Initial uptake rates (V) and $K_{m/i}$ values (μM) were determined for FurD (C) and FurE (D) using [^{3}H]-uracil uptake competition, as described in *Experimental procedures*. The graphs from which the $K_{m/i}$ values are calculated are shown in Fig. S5. Results are averages of at least three independent experiments in triplicate for each concentration point. Standard deviation was < 20%. n.m., no measurable.

and thymine for the wild-type and mutant versions of FurD (Fig. 8C). The main conclusions coming from these results can be summarised as follows: (i) Ala substitutions of Asn54, Trp130, Asn249, Tyr250 and Asn341 led to either no transport activity (W130A and N341A) or to significant reduction in the binding of all substrates, albeit at different degrees (N54A, N249A and Y250A). (ii) Ala substitutions of Thr57 and Gln134 have more subtle and specific effects, related in each case, with specific substrates. In particular, T57A did not affect at all the affinity for uracil and xanthine, but increased fivefold the affinity for uric acid binding, whereas Q134A has practically no effect on xanthine, uric acid or thymine binding, but reduced uracil and hydantoin binding three to fourfold. (iii) Specific substitutions of Gln134 and Tyr250 showed varying effects on the binding of different substrates. Q134E showed a general but variable reduction in the binding of all substrates. Y250W increased 10-fold and threefold uric acid and thymine binding, respectively, reduced sixfold uracil binding and had no effect on xanthine binding. All mutants analysed did not show any measurable binding affinity for allantoin, adenine, guanine, hypoxanthine and cytosine (results not shown).

A similar kinetic analysis was performed for FurE mutants (see Fig. 8D panel). Both single mutations (P251N and K252F) and the double mutation (P251N/K252F) showed dramatically increased transport capacities. In addition, they had an impressive positive effect on

the affinity of FurE for uracil (~37- to 330-fold), but also for allantoin, thymine and hydantoin (~8–250 fold). The same mutations had a moderate negative effect on uric acid binding and little effect on xanthine binding. Given that, Pro251 and Lys252 substitutions were designed to mimic FurD or FurA, our results, provided strong evidence confirming the critical role of these residues in determining the particular transport kinetics and specificity in Fur transporters.

A putative outward-facing gate is critical for FurD specificity

Although the modelled structures of the FurD and FurA binding sites are rather similar (conserved residues : Asn54/56 in TMS1, Trp130/132 and Gln134/136 in TMS3, Asn249/252 and Tyr250/Phe253 in TMS6, Asn341/349 in TMS8), their physiological substrate specificities are quite different. This observation points to the existence of elements, outside the major substrate binding site, acting as additional determinants of substrate specificity. The best candidates for a synergistic role in substrate selection and transport are domains that might act as selective gates. This idea is strongly supported by work on the UapA transporter, where mutations in putative outward- and inward-facing gates modified substrate specificity (Kosti *et al*., 2010; 2012; Papageorgiou *et al*., 2008). We searched for analogous gating domains in the Fur transporters.

According to structural data, the transport mechanism in Mhp1 involves substrate binding to an outward-facing conformation of the transporter, between the bundle (TMS1, -2, -6, and -7) and hash or scaffold domain (TMS3, -4, -8, -9) (Shimamura *et al*., 2010), followed by bending of the N-terminal part of TMS10 (TMS10a) over the substrate and occlusion of the binding site (Weyand *et al*., 2008). In other words, TMS10a acts as a dynamic element of an outward-facing gate. Primary sequence alignments show that in TMS10a fungal Fur transporters differ in specific amino acid residues that might affect the bending flexibility of this α-helical segment (Fig. 9A). In addition, plant Fur transporters and transporters of the Fcy-like sub-family have an extra amino acid inserted in the same region. Most importantly, a specific mutation in Mhp1 (L363A) converts the analogue Naphthyl-methyl-hydantoin (NMH) from a competitive inhibitor to a substrate, suggesting that the modification of the outward-facing gate can affect the transport specificity of Mhp1 (Simmons *et al*., 2014).

Based on these observations, we constructed the following FurD mutations: L386A, L386S, M389A, M389L, L386-N387/L386-E387*-N388, where Met389 (Fig. 9A and B) corresponds to Leu363 in Mhp1, and Glu387 is an inserted residue, corresponding to the extra amino acid found in Fcy-like transporters. Strains expressing FurD alleles were made as previously described. Growth tests (Fig. 9C) show that insertion of Glu387 and substitutions L386A, L386S and M389L led to apparent loss or partial loss of FurD-mediated transport. Most interestingly, M389A led to an active version of FurD, which besides its natural substrates, could also transport allantoin, hypoxanthine and the toxic analogue 8-azaguanine. Subsequent transport and competition assays confirmed that FurD-M389A can recognise not only uracil, uric acid, xanthine with high/moderate affinities (8–50 μM), but also allantoin and hypoxanthine with low affinities (0.7–1 mM) (Fig. 7D). Thus, introducing a small aliphatic residue in a critical position in TMS10a enlarged the specificity of FurD.

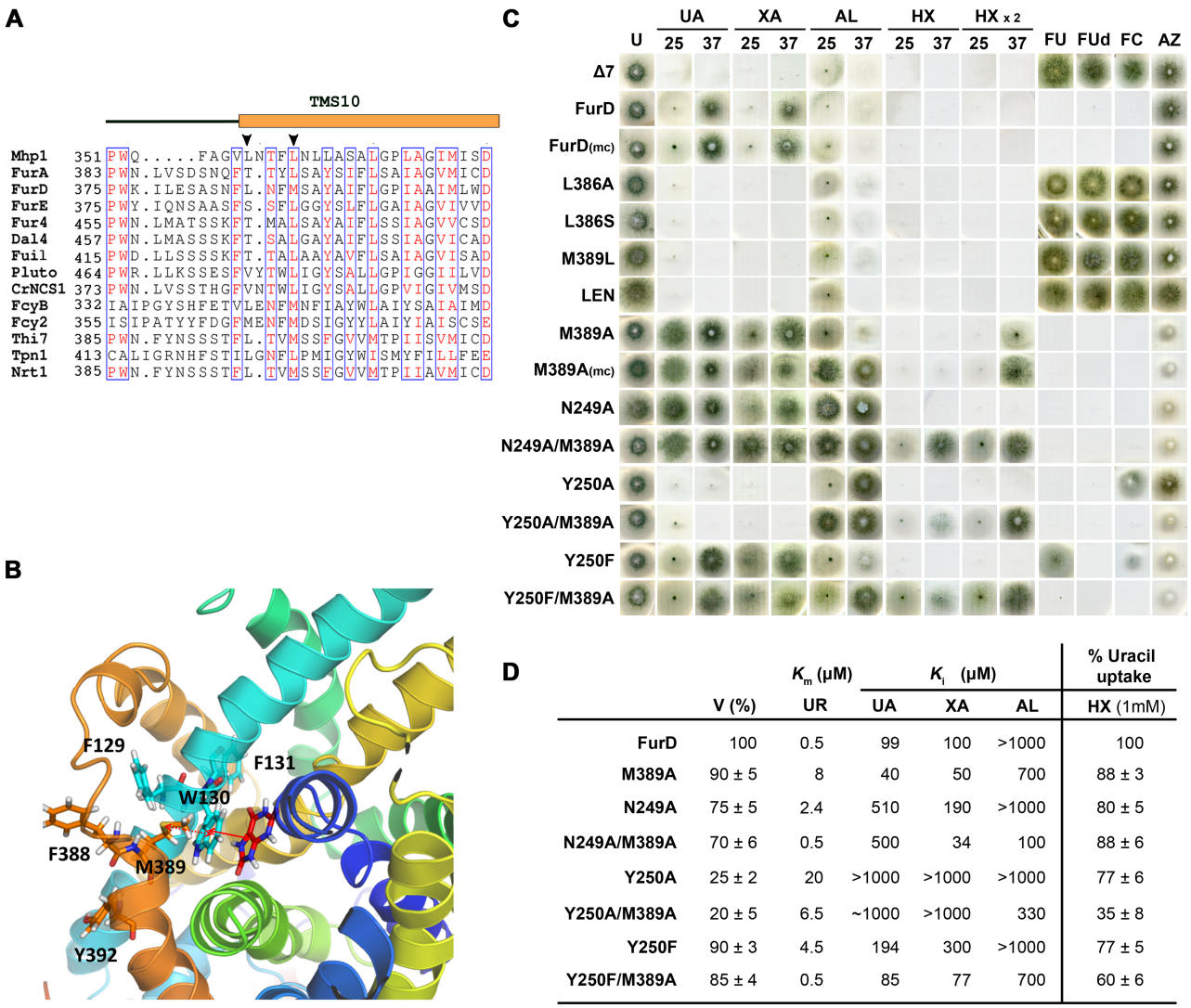
To further understand the role of Met389, we also combined mutation M389A with mutations N249A, Y250A or Y250F, located in the major substrate-binding site, as this was defined in this work. Results in Fig. 9C and D show that all combinations of mutations led to kinetically different versions of FurD, both in respect to substrate affinity and specificity. This constitutes further strong evidence that the function and specificity of FurD is determined by synergistic interactions of the binding site with TMS10a (see also *Discussion*).

To investigate the structural impact of M389A mutation that enlarges the specificity of FurD, we monitored the sulfur-aromatic interactions between Met389 and the proximal aromatic amino acids Phe129, Trp130, Phe131 of TMS3, as well as Phe388 and Tyr392 of TMS10, in the three MD trajectories (Fig. S6). We started our analysis after the 150th ns, where we believe the protein has been completely relaxed (see Fig. S3). We considered a sulfur-aromatic interaction to exist when the distance between the center of mass of Met is below 7 Å (Valley *et al*., 2012). Our results suggest that Met389 in the FurD-uracil complex interacts strongly with Phe129 and Phe388 and sporadically with Tyr392 aromatic side-chains. In the FurD-uric acid complex, Met389 interacts constantly with Phe129 and Trp130 and transiently with Phe388 and Tyr392 (see Fig. 9B). Finally, in the FurD-apo Asp51 trajectory, Met389 interacts constantly with Phe129 and Tyr392 (TMS10) and rarely with Phe388 (TMS10). Overall, all our results are in line with the idea that Met389 is an element of a putative outward-facing gate.

Discussion

Evolutionary and functional plasticity in NCS1 transporters

There are several basic conclusions arising from the phylogenetic analyses performed in this work. First, NCS1 sequences originated in prokaryotes, but their patchy distribution also suggests significant losses. Second, the two



	V (%)	K_m (μM) UR	K_i (μM) UA	K_i (μM) XA	K_i (μM) AL	% Uracil uptake HX (1mM)
FurD	100	0.5	99	100	>1000	100
M389A	90 ± 5	8	40	50	700	88 ± 3
N249A	75 ± 5	2.4	510	190	>1000	80 ± 5
N249A/M389A	70 ± 6	0.5	500	34	100	88 ± 6
Y250A	25 ± 2	20	>1000	>1000	>1000	77 ± 6
Y250A/M389A	20 ± 5	6.5	~1000	>1000	330	35 ± 8
Y250F	90 ± 3	4.5	194	300	>1000	77 ± 5
Y250F/M389A	85 ± 4	0.5	85	77	700	60 ± 6

Fig. 9. Functional characterisation of FurD-M389A.
A. Sequence alignment of TMS10 of selected NCS1 transporters (see also Fig. S4). The sequences shown are: Mhp1 (D6R8X8; *M. liquefaciens*), FurA (Q5BFM0; *A. nidulans*), FurD (A6N844; *A. nidulans*), FurE (Q5ATG4; *A. nidulans*), Fur4 (P05316; *S. cerevisiae*), Dal4 (Q04895; *S. cerevisiae*), Fui1 (Q04895; *S. cerevisiae*), PLUTO (Q9LZD0; *A. thaliana*), CrNCS1 (A8J166; *Chlamydomonas reinhardtii*), FcyB (C8V329; *A. nidulans*), Fcy2 (P17064; *S. cerevisiae*), Thi7 (Q05998; *S. cerevisiae*), Tpn1 (P53099; *S. cerevisiae*) and Nrt1 (Q08485; *S. cerevisiae*).
B. Topology of Met389 in TMS10a. Neighbouring aromatic amino acids are also shown. π-π-stacking interaction of substrate (uric acid)-Trp130 is illustrated with two-headed red arrow and the dotted two-headed arrow represents the sulfur-aromatic interaction Met389-Trp130.
C. Growth tests of TMS10 FurD mutants. Details are as in Fig. 3. (mc, multicopy plasmid integration).
D. Initial uptake rate (V) and $K_{m/i}$ values (μM) were determined for FurD-M389A and N249A,Y250A,Y250F/M389A double mutants by [^{3}H]-uracil uptake, as described in *Experimental procedures*. The other mutants in TMS10 had no measurable uptake (not shown). The graphs from which the $K_{m/i}$ values are calculated are shown in Fig. S5. Competition of [^{3}H]-uracil uptake by excess hypoxanthine is also shown (right column). Results are averages of at least three independent experiments in triplicate for each concentration point. Standard deviation was < 20%.

distinct fungal families, Fur and Fcy, originated independently from prokaryotes, very probably through HGT. Interestingly, early-diverging fungi do not have NCS1 sequences, with only two exceptions within the 56 fungi in our dataset (*G. prolifera* and *Batrachochytrium dendrobatidis*) and only three more within the full set with 271 fungal species included (not shown). Third, the plant NCS1 proteins, with the exception of some algal sequences, are only distantly related to the fungal clade and appear to represent an independent horizontal acquisition in the hypothetical proto-plant cell. Fourth, protozoa and metazoa do not contain NCS1 sequences, with very minor exceptions that may in fact represent artifact annotation due to symbiotic or contaminating bacteria. Fifth, the fungal NCS1 sequences

of dikarya, in both the Fur and Fcy families, have expanded in number due to repeated duplication. This seems to have allowed the evolution of novel functions and specificities through divergence or/and convergence of molecular elements necessary for substrate recognition and transport (Ohno, 1970). The previously proposed functional plasticity of the Fur family through convergent evolution (Hamari *et al*., 2009) is further reinforced here. It seems that uracil, allantoin or uridine transporters arise independently by convergent evolution in fungi (Fig. S7).

Interestingly, our analysis also suggests that promiscuity might have preceded specificity in NCS1 evolution. The FurE sequence is the closest, among all seven Furs analysed in this work, to a common ancestral Fur molecule (Fig. S8). In line with this, FurE orthologues are lost in most fungi. FurE is a rather promiscuous NCS1 transporter, recognising uracil, uric acid, allantoin, 5-fluorocytosine and 5-fluorouridine when overexpressed. Interestingly, plant NCS1 transporters, which have a prokaryotic origin independent from that of the fungi, are also quite promiscuous. Our molecular and genetic analysis has shown that mutations in selective gates can lead to a general increase in promiscuity (e.g. FurD-M389A; see Fig. 9), whereas mutations in the substrate binding site often lead to more specific modifications of substrate recognition and transport (e.g. FurD-Y250A, FurE-P251N, FurE- P251N/K252F; see Fig. 8). Thus, one can imagine an ancestral transporter with a promiscuous binding site and 'open' gates, which duplicates and evolves progressively to novel, more specific, transporters through either acquisition of selectivity in the opening of gates or/and modifications within the *bona fidae* substrate binding site.

The case of FurB: a persistent 'proto-pseudogene'?

Independent evidence showed that the *furB* gene is not transcribed under all conditions tested (Hamari *et al*., 2009; and results not shown, see below), whereas its forced transcription via a strong promoter led to very low steady-state levels of a polypeptide, which is localised in cytoplasmic structures of unknown identity, rather than in the plasma membrane. This picture has not changed when we followed FurB expression under various conditions concerning temperature, pH, N or C sources, or conditions inducing secondary metabolism (Cánovas *et al*., 2014; Yaegashi *et al*., 2014) (results not shown).

In no case we manage to detect *furB* transcription from its native promoter, or FurB-GFP localisation in the plasma membrane or other recognisable organellar membrane, except the endoplasmic reticulum (ER). The sequence of the putative FurB polypeptide does not show any particular feature distinguishing from the other six Furs with established transport activities. Phylogenetics showed that FurB has orthologues in two other *Aspergillus* species (*A. clavatus* and *A. niger*) and some other dikarya (Fig. S9), supporting that it must have been present in the ancestral Dikaryon and then lost in various lineages creating a patchy distribution. Could FurB be an unstable junk polypeptide or does it have a cryptic function? Cryptic functions were found for the so-called pseudo-enzymes in the rhomboid family of proteases (iRhoms). (Zettl *et al*., 2011; Adrain and Freeman, 2012; Reiterer *et al*., 2014). Large-scale sequencing of genomes has revealed that most enzyme families include inactive homologues. However, these pseudo-enzymes, unlike FurB, are expressed and often well conserved, implying a selective pressure to retain them during evolution, and therefore that they have significant function. We have recently obtained two cases, similar to FurB, concerning intrinsically unstable transporter paralogues showing limited conservation during evolution (Krypotou *et al*., 2015, Sá-Pessoa *et al*., 2015). Whether such cases correspond to orphan transporter genes with cryptic functions or represent persistent cases of 'proto-pseudogenes' is, we believe, an interesting issue.

Specificity determinants are partially determined by the Fur substrate-binding site

All evidence obtained supports the idea that substrate specificities within the NCS1 family cannot be predicted by simple sequence alignments, or through comparison of amino acid residues in the major substrate binding site defined in this work. Similar to what has been found in the UapA uric acid-xanthine transporter (Papageorgiou *et al*., 2008; Kosti *et al*., 2010; Diallinas, 2014), specificity seems to be determined by complex synergistic effects of the major substrate binding site and gating domains, which act at allowing or occluding access to or from the major substrate binding site. This explains how NCS1 proteins of very similar putative substrate binding sites have totally different specificities. A prominent example concerns the Dal4 and Fur4 transporters of *S. cerevisiae*, which are highly specific for either allantoin or uracil, but are very similar proteins and most importantly seem to have identical residues for substrate recognition (see Fig. S4). Their distinct specificity might well be explained through differences in gating domains.

The first point arising from the modelling and mutational analyses was that a single major binding site, made of residues mapping in TMS1, TMS3, TMS6 and TMS8, is present in all three Fur transporters analysed. In this site, the absolutely conserved Trp residue in TMS3 and Asn residue in TMS8 seem irreplaceable for function, as their substitution leads to total loss of transport activity. In agreement with this result, these amino acids are the only residues of the substrate binding site that are absolutely conserved in all NCS1 proteins. Their substitution has also

a negative effect on the transport activity of other NCS1 transporters, such as Fcy2 (Ferreira *et al*., 1997), FcyB (Krypotou *et al*., 2012), Mhp1 (Simmons *et al*., 2014) or PLUTO (Witz *et al*., 2014), further confirming that these residues are major elements for substrate binding in NCS1. Interestingly, no substitution of these residues led to a change in specificity, showing that they participate in rather flexible interactions as H donors or acceptors (Asn) or in π-π stacking (Trp). A third amino acid shown here to be important in the binding and transport of substrates is a Gln residue in TMS3. This residue, absolutely conserved in Fur proteins, is replaced by an Asn residue in Fcy-like transporters, or Glu in plant NCS1 and can vary in bacterial NCS1 members (Asn/Gln/Glu). Its replacement with other amino acids leads to a moderate negative effect on Fur activity, but a more dramatic loss of activity in Fcy transporters or a plant NCS1 (PLUTO). Specific mutations in this Gln residue can however differentially affect the binding of various substrates, showing that it contributes to specificity. Finally, a fourth residue shown to be important for FurD or FurE functioning was Asn54 or Ser54, respectively, within TMS1. Replacement of these polar residues with Ala led to dramatic loss in substrate binding affinity and transport. Interestingly, this amino acid position is variably conserved in NCS1 proteins, in Fur transporters is either an Asn (FurD and FurA) or a Ser (FurE), a Val in Fcy-like proteins, a Gly or Val in plant NCS1s, and variable in bacterial homologues. In the plant PLUTO transporter, Asn54/Ser54 in FurD/FurA corresponds to Gly147 (Witz *et al*., 2014). However, the presence of a conserved Asn in uracil (FurD and Fur4), allantoin (FurA, Dal4) or uridine (Fui1) fungal transporters, argues against a simple explanation on how this residue affects specificity. The analogous residue in the crystal structure of Mhp1, Gln42, was shown to be in close distance, but not to interact, with the substrate. Mutations replacing Gln42 with aliphatic residues led to significant loss of transport activity. Overall, it seems that this residue participates in flexible and variable interactions critical for substrate binding and/or transport in different NCS1 transporters.

The more pronounced specificity changes due to mutations in the major substrate binding site in FurD and FurE were obtained by substituting residues in TMS6. For both proteins, specific mutations, replacing these residues with those found in other Fur-proteins shifted substrate specificities in a pronounced fashion. For example, we could make a FurE version recognising with high-affinity pyrimidines, but not purines (FurE-P251N/K252F), or a version recognising with relatively high-affinity several substrates (FurE-K252F, e.g. uracil, allantoin, uric acid, xanthine, thymine or hydantoin). Or, make a FurD version transporting with high-capacity, but low-affinity, allantoin (FurD-N249A). In line with the importance of TMS6 residues in determining the specificity of NCS1s, these residues are quite variably conserved among NCS1 members possessing distinct specificities, in close distance from substrates (see Fig. S4).

A selective gate in FurD is critical for specificity

As some of the specificity changes obtained by mutations in the substrate binding site of FurD were moderate or somehow non-expected, we predicted and showed that specificity is also determined through the functioning of putative gating domains, acting independently from, but synergistically with, the major substrate binding site. Mutant FurD-M389A could transport an enlarged set of substrates, including hypoxanthine and 8-azaguanine, which no known Fur protein can recognise. In the modelled structure, residue Met389 lies over the substrate binding site, at the N-terminal part of TMS10 (see Fig. 9B). In Mhp1, TMS10 is shown to be an element of an outward-facing gate. Crystallography has shown that TMS10 is relatively free in the outward-open ligand-free state, but upon substrate binding, it bends over it, occluding it in the binding site and, subsequently, the protein switches to the inward-facing form (Weyand *et al*., 2008; Shimamura *et al*., 2010). Moreover, MD simulations suggest that in the outward-facing state, this helix is in equilibrium between open and closed conformations (Shimamura *et al*., 2010; Adelman *et al*., 2011; Shi, 2013; Zhao and Noskov, 2013; Kazmier et al., 2014). Very recently, X-ray crystallography has shown that in presence of a bulkier molecule (NMH; 5-(2-naphthylmethyl)-L-hydantoin) in the substrate-binding site, TMS10 retains the conformation seen in the ligand free structure. However, that condition is inverted by a single mutation at position 363 (L363A), which relieves a steric clash between a leucine and the ligand, hence allowing the closure of TMS10, which is a prerequisite for transport (Simmons *et al*., 2014). Thus, it can be conjectured that the M389A mutation in FurD has an effect analogous to mutation L363A in Mhp1. An important difference though between the two homologues is that Met389 is surrounded by many aromatic residues, including the TMS10 amino acids Phe388 and Tyr392, as well as the TMS3 residues Phe129, Trp130 and Phe131. Sulfur-aromatic interactions are of mixed electrostatic and dispersion nature and have an associated energy comparable with salt bridges in proteins (Valley *et al*., 2012). Thus, the simplest scenario of the transport properties of FurD-M389A is that in the wild-type protein Met confers a degree of substrate selectivity through specific interactions with other parts of the protein. When replaced with a small residue like Ala, the Met389-Trp130 sulfur-aromatic interaction is lost allowing the bulky Trp130 to rotate easier, thus facilitating fitting of more diverse substrates in the binding pocket, as well as the transition to the inward-facing substrate occluded state.

Noticeably, there is an impact of the temperature on FurD- or FurE-mediated growth, which might be related to the gating process. In fact, there is general trend towards better growth on uric acid or xanthine at 37°C compared with 25°C, but this is not the case for growth on allantoin or pyrimidine analogues (see Figs 3, 8 and 9). The apparent functional cryosensitivity of FurD and FurE suggests that uric acid and xanthine must have a different binding mode compared with allantoin or uracil and that this difference becomes more visible at lower temperatures, where protein flexibility is lower. Importantly, in some mutants, the cryosensitivity on uric acid or xanthine is partially or totally suppressed (e.g. N249A, Y250F, M389A, N249A/M389A in FurD, or K252F in FurE). In addition, FurE-mediated accumulation of allantoin also becomes cryosensitive in P251N (see Fig. 8). Thus, suppression or modification of the effect of temperature on Fur activities always concerns mutations in the putative gate (TMS10) or in the part of the substrate binding domain (TMS6) that is just below the gate. In other words, mutations that modify the effect of temperature on Fur activities are located in dynamic domains that synergise to define the specificity of these transporters.

Experimental procedures

Sequence data/identification

For the large-scale profiling and phylogeny, proteins encoded in 1093 fully sequenced genomes available at PhylomeDB (Huerta-Cepas *et al*., 2014) as of May 2014 were downloaded. The final database comprises 8 962 138 unique protein sequences. For the second part of the analysis, where we focused on the evolution of Fur family using a reduced species sampling, we selected 81 species (Table S5), covering all fungal taxonomic ranges available (61 genomes), four plant genomes, 20 bacterial genomes and the Mhp1 sequence from *M. liquefaciens* as the only member of the superfamily with a solved structure. The NCBI taxids, species names, source and date of each genome getting integrated in PhylomeDB are provided in the Table S5. Proteins within our genomes databases containing at least one Nucleobase cation symporter-1 domain were detected by searching for the PF02133 'Transp_cyt_pur' profile as found in Pfam using HMMER 3.1b1 (Eddy, 2011). The above domain is specific for the NCS1 superfamily being part of all of its members sequence and only in it. The scan was performed using the Gathering Cut-Off threshold (–cut_ga), minimising the number of false positives. In the case of *A. nidulans*, the superfamily members were added manually, given that for one member of the Fcy family the open reading frame is split between two scaffolds and thus absent from the proteome.

Phylogenetic analysis

In all the cases of gene family trees, the execution of the different phylogenetic workflows was done using the bioinformatics tool ETE-NPR as an environment (J. Huerta-Cepas, P. Bork and T. Gabaldón, unpublished) and by using the single gene tree execution mode without the npr functionality. For the first larger tree of 2248 sequences, multiple sequence alignments were built using CLUSTALO v1.2 (Sievers *et al*., 2011). The resulting alignment was trimmed using trimAl v1.3 (Capella-Gutierrez *et al*., 2009) using a gap score cut-off of 0.01 due to the large alignment size. Phylogenetic trees were reconstructed using an ML approach as implemented in FastTree 2.1.7 (Price *et al*., 2010) using the WAG evolutionary model and the gamma parameter. In all the other gene family trees, we used a similar pipeline to that described in Huerta-Cepas *et al*. (2011). Multiple sequence alignments were built from each set of homologous sequences using M-COFFEE v8.80 (Wallace *et al*., 2006) to combine the results of three different alignment programs: MUSCLE v3.8.31 (Edgar, 2004), MAFFT v6.814b (Katoh, 2008) and DIALIGN-TX (Subramanian *et al*., 2008). Alignments were performed in forward and reverse direction, thus evaluating six alignments per query. The resulting alignment of each family was trimmed using trimAl v1.3 (Capella-Gutierrez *et al*., 2009) using a gap score cut-off of 0.1. For each sequence alignment, the best fit evolutionary model selection was done prior to phylogenetic inference using ProtTest 3 (Darriba *et al*., 2011). In each case three different evolutionary models were tested (JTT, WAG and LG). The model best fitting the data was determined by comparing the likelihood of all models according to the AIC criterion (Akaike, 1973). An ML tree was inferred with RAxML v8.0.2 (Stamatakis, 2014) using the best-fitting model. A discrete gamma-distribution model with four rate categories plus invariant positions was used. The gamma parameter and the fraction of invariant positions were estimated from the data. Branch support was computed using an aLRT (approximate likelihood ratio test) based on a chi-square distribution using PhyML (Guindon and Gascuel, 2003).

Phylogenetic tree analysis and visualisation was performed by using the methods provided in the ETE toolkit (Huerta-Cepas *et al*., 2010) apart from the visualisation of the unrooted trees that was done using the web-based tool iTOL (Letunic and Bork, 2011).

Homology modelling

Full-length sequences of FurA, and FurE from *A. nidulans*, were retrieved from AspGD (Cerqueira *et al*., 2013) (accession codes AN0660 and AN8416, respectively) and FurD of the same species from UniProt (accession code A6N844). For homology model construction, we used a previous published alignment between NSC1 members, including FurA, FurD, FurE and Mhp1 from *M. liquefaciens* (Krypotou *et al*., 2012; Fig. S2). The full-length sequences of FurA, FurD and FurE were submitted to TOPCONS meta-server (Bernsel *et al*., 2009) for consensus membrane protein topology prediction. Two thousand models of FurA residues 20–500, 2000 models of FurD residues 22–493 and 2000 models of FurE residues 20–507 were constructed in total, using the Rosetta threading protocol with the membrane specific scoring function terms (Yarov-Yarovoy *et al*., 2006; Barth *et al*., 2007; 2009). The secondary structure of the proteins, as predicted by SAM server (Karplus, 2009), as well as the

consensus membrane protein topology prediction made by TOPCONS meta-server, were also taken into account in the form of restraints during model construction. The 3-mer and 9-mer peptide fragments for modelling were created locally using the SAM-predicted secondary structure of each protein. Unfolded and broken structures were excluded by retaining only the 90% top-score models. A random model among them was superimposed on the crystal structure of Mhp1 using the combinatorial extension algorithm (Shindyalov and Bourne, 1998), and the putative substrate binding cavities of FurA, FurD and FurE were determined by selecting every residue that was poised within 8 Å from the benzyl-hydantoin ligand. Trajectories were assembled from the remained low-energy models and were clustered with GROMACS Tools v4.6.1 (Pronk *et al.*, 2013) using the backbone heavy atoms and Cβ carbon of the putative substrate binding pocket. The algorithm used for clustering was 'gromos' with cut-off 1.5 Å.

Induced fit docking

The proteins were prepared for docking calculations using the Schrödinger Suite 2014 Protein Preparation workflow (Schrödinger Inc., New York, NY, USA). Molecular docking was performed on the most dominant cluster representative receptor structures using the Schrödinger Suite 2014 IFD protocol, which is intended to circumvent the inflexible binding site and accounts for the side-chain and backbone movements upon ligand binding (Sherman *et al.*, 2006). The details of IFD are described in the Supporting information.

Molecular dynamics simulations

The orientation of FurD with respect to the membrane plane (x-y plane) was predicted by Orientations of Proteins in Membranes (OPM) server (Lomize *et al.*, 2006). Subsequently, FurD-apo with neutral Asp51, FurD with uric acid and FurD with uracil structures were prepared and submitted to CHARMM-GUI (Jo *et al.*, 2008), to be inserted into pre-equilibrated DPPC bilayers and solvated into rectangular boxes with final NaCl concentration of 150 mM. These three systems were processed by the LEaP module in AmberTools14 (Case *et al.*, 2014), where Amber ff14SB force field parameters were applied to the protein, lipid14 (Dickson *et al.*, 2014) to the DPPC lipids and TIP3P to the water molecules. For the ligands, we used the atomic charges we calculated using the R.E.D. Server development (Dupradeau *et al.*, 2010; Vanquelef *et al.*, 2011; Wang *et al.*, 2013) along with PARM10 parameters (Zgarbová *et al.*, 2011), as described in the Supporting information. Constant pressure and constant temperature (NPT) runs were performed on the three protein-membrane systems using the GPU implementation of the AMBER 12 code (Case *et al.*, 2014), following a short energy minimisation and equilibration in the NVT ensemble, as described in detail in the Supporting information. A 4th system containing FurD-apo Asp51/M389A was modelled using the 300th nanosecond snapshot of FurD-apo Asp51 trajectory, to avoid relaxing the protein again, and a new simulation was seeded. The total simulation time for FurD-uric acid and FurD-uracil complex was 400 ns, for FurD-apo 555 ns and for FurD-apo M389A 255 ns (starting from the relaxed structure). Finally the trajectories were analysed using AmberTools14 and GROMACS Tools 4.6.1. Water-Bridge frequencies were calculated through combination of g_hbond module of GROMACS Tools and a post-processing python script we developed. H-bonds were considered to be formed when the donor-acceptor distance was less than 3.5 Å and the donor-H-acceptor angle less than 30°. Simulations were performed in a membrane consisting of saturated lipids, like the well-studied DPPC, in order to emulate as much as possible the membrane constitution of the fungal lipid raft (ergosterol, sphingolipids and glycerophospholipids). Unfortunately, there were no sterol, sphingolipid or glycerophospholipid parameters in LIPID14 force field and hence we used homogenous DPPC bilayers.

Media, strains, genetic techniques, growth conditions and A. nidulans *transformation*

Standard complete and minimal media (MM) for *A. nidulans* were used. Media and supplemented auxotrophies were at the concentrations given in http://www.fgsc.net. Nitrogen sources and analogues were used at the final concentrations: urea 5 mM, purines 0.5 mM, xanthine (XA) 1 mM, 5-fluorouracil (5-FU) 50–100 μM, 5-fluorouridine (5-FUd) 10 μM, 5-fluorocytosine (5-FC) 50 μM, 8-azaguanine (8-AZ) 1 mM, oxypurinol (OX) 100 μM. *E. coli* was grown on Luria–Bertani medium. Media and chemical reagents were obtained from Sigma-Aldrich (Life Science Chemilab SA, Hellas) or AppliChem (Bioline Scientific SA, Hellas). Derivatives of mutant strains were made with standard genetic crossing using auxotrophic markers for heterokaryon establishment. Multiple *Δfur* strains were identified by phenotype and PCR. A list of strains used in this study is included in Table S6. *A. nidulans* transformation was performed as described previously (Koukaki *et al.*, 2003). A *ΔfurD::riboB ΔfurA::riboB ΔfcyB::argB ΔazgA ΔuapA ΔuapC::AfpyrG ΔcntA::riboB pabaA1 pantoB100* mutant strain, named Δ7, was the recipient strain in transformations with plasmids carrying *fur* genes or alleles. This strain has been described in Krypotou and Diallinas (2014). In brief, it lacks, due to genetic deletions, all genes encoding major transporters for purines, pyrimidines and nucleosides. Selection was based on complementation of pantothenic acid auxotrophy *pantoB100*. Transformants expressing intact *fur* or *fur–gfp* alleles, through either single-copy or multi-copy plasmid integration events, were identified by PCR and Southern analysis. As a negative control a transformant obtained with an empty vector *pantoB* was used and as a wild type strain a *pabaA1* strain. Growth tests were performed at 25°C or 37°C, at pH 6.8.

Plasmid constructions and FurD/E mutations

The principal vector used is a modified pGEM-T-easy vector carrying the *gpdA* promoter (~1000 bp), cloned at AatII and SpeI restriction sites of the multicloning site (MCS), the *trpC* 3′ termination region, cloned at NotI and PstI restriction sites of the MCS, and the *pantoB* gene, at the NdeI restriction site of the MCS. The *fur* orfs were obtained from wild-type genomic DNA by PCR using appropriate primers and cloned

at the SpeI-NotI sites of the vector. For *fur-gfp* constructions, the *fur* orfs lacking the translation stop codon were amplified with relevant primers and the *gfp* orf was amplified using primers with NotI extensions and cloned in frame at the end of the *fur* orfs. Mutations were constructed by site-directed mutagenesis according to the instructions accompanying the Quik-Change® Site-Directed Mutagenesis Kit (Stratagene) on the above vectors and were confirmed by sequencing. Oligonucleotides used for site-directed mutagenesis, cloning or diagnostic purposes are also listed in Table S7.

Standard nucleic acid manipulations

Genomic DNA extraction from *A. nidulans* was as described in http://www.fgsc.net. Plasmid preparation from *E. coli* strains and DNA bands purification from agarose gels were done with the Nucleospin Plasmid kit and the Nucleospin ExtractII kit according to the manufacturer's instructions (Macherey-Nagel, Lab Supplies Scientific SA, Hellas). Southern blot analysis was performed as described in Sambrook *et al*. (1989). [^{32}P]-dCTP-labelled molecules used as *fur or pantoB* specific probes were prepared using a random hexanucleotide primer kit following the supplier's instructions (Takara Bio, Lab Supplies Scientific SA, Hellas) and purified on MicroSpin™ S-200 HR columns, following the supplier's instructions (Roche Diagnostics, Hellas). Labelled [^{32}P]-dCTP (3000 Ci mmol^{-1}) was purchased from the Institute of Isotops, Miklós, Hungary. Restriction enzymes were from Takara Bio (Lab Supplies Scientific SA, Hellas). Conventional PCR reactions were done with KAPATaq DNA polymerase (Kapa Biosystems, Lab Supplies Scientific SA, Hellas). Cloning and amplification of products were done with Kapa HiFi (Kapa Biosystems, Lab Supplies Scientific SA, Hellas).

Total protein extraction and western blot analysis

Cultures for total protein extraction were grown in MM supplemented with urea at 25°C for 16 h. Total protein extraction was performed as previously described (Galanopoulou *et al*., 2014). Equal sample loading was estimated by Bradford assays and Coomasie staining. Total proteins (30 μg) were separated by SDS-PAGE (10% w/v polyacrylamide gel) and electroblotted (Mini PROTEAN™ Tetra Cell, BIORAD) onto PVDF membranes (Macherey-Nagel, Lab Supplies Scientific SA, Hellas) for immunodetection. The membrane was treated with 2% (w/v) non-fat dried milk, and immunodetection was performed with a primary mouse anti-GFP monoclonal antibody (Roche Diagnostics, Hellas), a mouse anti-actin monoclonal (C4) antibody (MP Biomedicals Europe) and a secondary goat anti-mouse IgG HRP-linked antibody (Cell Signaling Technology Inc, Bioline Scientific SA, Hellas). Blots were developed by the chemiluminescent method using the LumiSensor Chemiluminescent HRP Substrate kit (Genscript USA, Lab Supplies Scientific SA, Hellas) and SuperRX Fuji medical X-Ray films (FujiFILM Europe, Lab Supplies Scientific SA, Hellas). Densitometry analysis of Fur-GFP, free-GFP and actin-specific bands was performed using the *ImageJ* software (Schneider *et al*., 2012). The relative density of the Fur-GFP and free-GFP bands were quantified and normalised by dividing with the respective actin-band density.

Kinetic analysis

The kinetic analysis was performed as recently described in detail in Krypotou and Diallinas (2014). In more detail, [^{3}H]-uracil (40 Ci mmol^{-1}, Moravek Biochemicals, CA, USA) uptake was assayed in *A. nidulans* conidiospores germinating for 3.5–4 h at 37°C, at 130–150 r.p.m., in liquid MM supplemented with 1% w/v Glucose as a carbon source and urea or nitrate as a nitrogen source, pH 6.8. The conidiospores were collected by centrifugation and resuspended in fresh MM at a concentration of 10^7 conidiospores/100 μl. Initial velocities were measured by incubation with concentrations of 0.2–2.0 μM of [^{3}H]-uracil at 37°C. The time of incubation was defined through time-course experiments and the period where each transporter showed linear increased activity was chosen respectively. $K_{m/i}$ values were obtained using labelled uracil at 0.2–0.5 μM in the presence of various concentrations (0.5–2000 μM) of non-labelled substrates. Reactions were terminated with addition of equal volumes of ice-cold MM containing 1000-fold excess of non-radiolabelled substrate. K_i values were calculated by satisfying the criteria for use of the Cheng and Prusoff equation $K_i = IC_{50}/(1 + (L/K_m))$, in which L is the permeant concentration (Cheng and Prusoff, 1973). In our assays, K_i values equal IC50 values as the L is very low (at least 10-fold lower than the K_m value). Radioactive counts are converted to the amount of moles of substrate taken up by a defined number of germinating conidia over a defined period of time, based on the concentration and specific activity of the stock of radioactive substrate used. Background, transporter-independent, radioactive counts, obtained due to unspecific labelling of cells or in some cases by simple diffusion, are estimated using control assays with an isogenic transporter-deleted strain, or in the simultaneous presence of 100- to 1000-fold excess non-labelled substrate. Background counts are subtracted from the values obtained in strain expressing the relevant transporter. K_m, K_i and V_{max} determination is carried out using standard Lineweaver–Burk or Eadie–Hofstee equations formulas, or a relevant software (e.g. GraphPad Prism). The analysis is performed through a non-linear regression curve fit and sigmoidal dose response analysis. Quality factors for the analysis result are: R^2 which should be > 0.99 and the Hill co-efficient (Weiss, 1997), which should be approximately –1 for a transporter with one binding site. All transport assays were carried out in at least three independent experiments, with three replicates for each concentration or time point. Standard deviation was < 20%.

Epifluorescence microscopy

Samples for epifluorescence microscopy were prepared as previously described (Gournas *et al*., 2010; Karachaliou *et al*., 2013). In brief, germlings incubated on coverslips in liquid MM supplemented with urea as nitrogen source for 12–14 h at 25°C, were observed on an Axioplan Zeiss (Carl Zeiss AG Corporate, Oberkochen, Germany) phase-contrast epi-fluorescent microscope or a Zeiss Axio Observer Z1 inverted epi-fluorescent microscope, and the resulting images were acquired with a Zeiss-MRC5 digital camera using the AxioVs40 V4.40.0 software. Image processing, contrast adjustment and color combining were made using the Adobe Photoshop CS4 Extended version 11.0.2 software

(Adobe Systems Incorporated San Jose, California, USA) or the ImageJ software. Images were converted to RGB and annotated using Photoshop CS4 before being saved to TIFF.

Acknowledgements

We thank George Lambrinidis for help in the initial modelling approaches. T.E. acknowledges Bodossaki Foundation. E.K. and this research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program 'Education and Lifelong Learning' of the National Strategic Reference Framework (NSRF). Research Funding Program: THALES – Investing in knowledge society through the European Social Fund. The structural part of the work was supported by the Cy-Tera Project (NEA YPODOMI/STRATI/0308/31), which is co-funded by the European Regional Development Fund and the Republic of Cyprus through the Research Promotion Foundation.

References

Adelman, J.L., Dale, A.L., Zwier, M.C., Bhatt, D., Chong, L.T., Zuckerman, D.M., and Grabe, M. (2011) Simulations of the alternating access mechanism of the sodium symporter Mhp1. *Biophys J* **101:** 2399–2407.

Adrain, C., and Freeman, M. (2012) New lives for old: evolution of pseudoenzyme function illustrated by iRhoms. *Nat Rev Mol Cell Biol* **13:** 489–498.

Akaike, H. (1973) Information theory and extension of the maximum likelihood principle. In: *Proceedings of the 2nd international symposium on information theory: 1973*. Budapest, Hungary. pp. 267–281.

Amillis, S., Hamari, Z., Roumelioti, K., Scazzocchio, C., and Diallinas, G. (2007) Regulation of expression and kinetic modelling of substrate interactions of a uracil transporter in *Aspergillus nidulans*. *Mol Membr Biol* **24:** 206–214.

Barth, P., Schonbrun, J., and Baker, D. (2007) Toward high-resolution prediction and design of transmembrane helical protein structures. *Proc Natl Acad Sci USA* **104:** 15682–15687.

Barth, P., Wallner, B., and Baker, D. (2009) Prediction of membrane protein structures with complex topologies using limited constraints. *Proc Natl Acad Sci USA* **106:** 1409–1414.

Belenky, P., Christensen, K.C., Gazzaniga, F., Pletnev, A.A., and Brenner, C. (2009) Nicotinamide riboside and nicotinic acid riboside salvage in fungi and mammals. Quantitative basis for Urh1 and purine nucleoside phosphorylase function in NAD+ metabolism. *J Biol Chem* **284:** 158–164.

Bernsel, A., Viklund, H., Hennerdal, A., and Elofsson, A. (2009) TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Res* **37:** W465–W468.

Cánovas, D., Marcos, A.T., Gacek, A., Ramos, M.S., Gutiérrez, G., Reyes-Domínguez, Y., and Strauss, J. (2014) The histone acetyltransferase GcnE (GCN5) plays a central role in the regulation of *Aspergillus* asexual development. *Genetics* **197:** 1175–1189.

Capella-Gutierrez, S., Silla-Martinez, J.M., and Gabaldon, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25:** 1972–1973.

Case, D.A., Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., *et al*. (2014) AMBER 14. San Francisco: University of California.

Cerqueira, G.C., Arnaud, M.B., Inglis, D.O., Skrzypek, M.S., Binkley, G., Simison, M., *et al*. (2013) The *Aspergillus* Genome Database: multispecies curation and incorporation of RNA-Seq data to improve structural gene annotations. *Nucleic Acids Res* **42:** D705–D710.

Cheng, Y., and Prusoff, W.H. (1973) Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol* **22:** 3099–30108.

Danielsen, S., Kilstrup, M., Barilla, K., Jochimsen, B., and Neuhard, J. (1992) Characterization of the *Escherichia coli* codBA operon encoding cytosine permease and cytosine deaminase. *Mol Microbiol* **6:** 1335–1344.

Darlington, A.J., and Scazzocchio, C. (1967) Use of analogues and the substrate-sensitivity of mutants in analysis of purine uptake and breakdown in *Aspergillus nidulans*. *J Bacteriol* **93:** 937–940.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27:** 1164–1165.

Diallinas, G. (2014) Understanding transporter specificity and the discrete appearance of channel-like gating domains in transporters. *Front Pharmacol* **5:** 207.

Diallinas, G., and Gournas, C. (2008) Structure-function relationships in the nucleobase-ascorbate transporter (NAT) family: lessons from model microbial genetic systems. *Channels (Austin)* **2:** 363–372.

Dickson, C.J., Madej, B.D., Skjevik, Å.A., Betz, R.M., Teigen, K., Gould, I.R., and Walker, R.C. (2014) Lipid14: The Amber Lipid Force Field. *J Chem Theory Comput* **10:** 865–879.

Dupradeau, F.-Y., Pigache, A., Zaffran, T., Savineau, C., Lelong, R., Grivel, N., *et al*. (2010) The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Phys Chem Chem Phys* **12:** 7821–7839.

Eddy, S.R. (2011) Accelerated Profile HMM Searches. *PLoS Comput Biol* **7:** e1002195. doi: 10.1371/journal.pcbi.1002195.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32:** 1792–1797.

Enjo, F., Nosaka, K., Ogata, M., Iwashima, A., and Nishimura, H. (1997) Isolation and characterization of a thiamin transport gene, THI10, from *Saccharomyces cerevisiae*. *J Biol Chem* **272:** 19165–19170.

Ferreira, T., Brèthes, D., Pinson, B., Napias, C., and Chevallier, J. (1997) Functional analysis of mutated purine-cytosine permease from *Saccharomyces cerevisiae*. A possible role of the hydrophilic segment 371–377 in the active carrier conformation. *J Biol Chem* **272:** 9697–9702.

Frillingos, S. (2012) Insights to the evolution of Nucleobase-Ascorbate Transporters (NAT/NCS2 family) from the Cys-scanning analysis of xanthine permease XanQ. *Int J Biochem Mol Biol* **3:** 250–272.

Galanopoulou, K., Scazzocchio, C., Galinou, M.E., Liu, W.,

Borbolis, F., Karachaliou, M., *et al*. (2014) Purine utilization proteins in the Eurotiales: cellular compartmentalization, phylogenetic conservation and divergence. *Fungal Genet Biol* **69:** 96–108.

Girke, C., Daumann, M., Niopek-Witz, S., and Möhlmann, T. (2014) Nucleobase and nucleoside transport and integration into plant metabolism. *Front Plant Sci* **5:** 443.

Gournas, C., Papageorgiou, I., and Diallinas, G. (2008) The nucleobase-ascorbate transporter (NAT) family: genomics, evolution, structure-function relationships and physiological role. *Mol Biosyst* **4:** 404–416.

Gournas, C., Amillis, S., Vlanti, A., and Diallinas, G. (2010) Transport-dependent endocytosis and turnover of a uric acid-xanthine permease. *Mol Microbiol* **75:** 246–260.

Guindon, S., and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52:** 696–704.

Hamari, Z., Amillis, S., Drevet, C., Apostolaki, A., Vágvölgyi, C., Diallinas, G., and Scazzocchio, C. (2009) Convergent evolution and orphan genes in the Fur4p-like family and characterization of a general nucleoside transporter in *Aspergillus nidulans*. *Mol Microbiol* **73:** 43–57.

Huerta-Cepas, J., Dopazo, H., Dopazo, J., and Gabaldón, T. (2007) The human phylome. *Genome Biol* **8:** R109.

Huerta-Cepas, J., Dopazo, J., and Gabaldón, T. (2010) ETE: a python Environment for Tree Exploration. *BMC Bioinformatics* **11:** 24.

Huerta-Cepas, J., Capella-Gutierrez, S., Pryszcz, L.P., Denisov, I., Kormes, D., Marcet-Houben, M., and Gabaldón, T. (2011) PhylomeDB v3.0: an expanding repository of genome-wide collections of trees, alignments and phylogeny-based orthology and paralogy predictions. *Nucleic Acids Res* **39:** D556–D560.

Huerta-Cepas, J., Capella-Gutiérrez, S., Pryszcz, L.P., Marcet-Houben, M., and Gabaldón, T. (2014) PhylomeDB v4: zooming into the plurality of evolutionary histories of a genome. *Nucleic Acids Res* **42:** D897–D902.

Jo, S., Kim, T., Iyer, V.G., and Im, W. (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *J Comput Chem* **29:** 1859–1865.

Jund, R., Weber, E., and Chevallier, M.R. (1988) Primary structure of the uracil transport protein of *Saccharomyces cerevisiae*. *Eur J Biochem* **171:** 417–424.

Karachaliou, M., Amillis, S., Evangelinos, M., Kokotos, A.C., Yalelis, V., and Diallinas, G. (2013) The arrestin-like protein ArtA is essential for ubiquitination and endocytosis of the UapA transporter in response to both broad-range and specific signals. *Mol Microbiol* **88:** 301–317.

Karplus, K. (2009) SAM-T08, HMM-based protein structure prediction. *Nucleic Acids Res* **37:** W492–W497.

Katoh, T. (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* **9:** 286–298.

Kazmier, K., Sharma, S., Islam, S.M., Roux, B., and Mchaourab, H.S. (2014) Conformational cycle and ion-coupling mechanism of the Na+/hydantoin transporter Mhp1. *Proc Natl Acad Sci USA* **111:** 14752–14757.

de Koning, H., and Diallinas, G. (2000) Nucleobase transporters. *Mol Membr Biol* **17:** 75–94.

Kosti, V., Papageorgiou, I., and Diallinas, G. (2010) Dynamic elements at both cytoplasmically and extracellularly facing sides of the UapA transporter selectively control the accessibility of substrates to their translocation pathway. *J Mol Biol* **397:** 1132–1143.

Kosti, V., Lambrinidis, G., Myrianthopoulos, V., Diallinas, G., and Mikros, E. (2012) Identification of the substrate recognition and transport pathway in a eukaryotic member of the nucleobase-ascorbate transporter (NAT) family. *PLoS ONE* **7:** e41939.

Koukaki, M., Giannoutsou, E., Karagouni, A., and Diallinas, G. (2003) A novel improved method for *Aspergillus nidulans* transformation. *J Microbiol Methods* **55:** 687–695.

Kraupp, M., and Marz, R. (1995) Membrane transport of nucleobases: interaction with inhibitors. *Gen Pharmacol* **26:** 1185–1190.

Krypotou, E., and Diallinas, G. (2014) Transport assays in filamentous fungi: kinetic characterization of the UapC purine transporter of *Aspergillus nidulans*. *Fungal Genet Biol* **63:** 1–8.

Krypotou, E., Kosti, V., Amillis, S., Myrianthopoulos, V., Mikros, E., and Diallinas, G. (2012) Modelling, substrate docking, and mutational analysis identify residues essential for the function and specificity of a eukaryotic purine-cytosine NCS1 transporter. *J Biol Chem* **287:** 36792–36803.

Krypotou, E., Scazzocchio, C., and Diallinas, G. (2015) Functional characterization of NAT/NCS2 proteins of *Aspergillus brasiliensis* reveals a genuine xanthine-uric acid transporter and an intrinsically misfolded polypeptide. *Fungal Genet Biol* **75:** 56–63. doi: 10.1016/j.fgb.2015.01.009

Letunic, I., and Bork, P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* **39:** W475–478.

Lomize, M.A., Lomize, A.L., Pogozheva, I.D., and Mosberg, H.I. (2006) OPM: orientations of proteins in membranes database. *Bioinformatics* **22:** 623–625.

de Montigny, J., Straub, M.L., Wagner, R., Bach, M.L., and Chevallier, M.R. (1998) The uracil permease of *Schizosaccharomyces pombe*: a representative of a family of 10 transmembrane helix transporter proteins of yeasts. *Yeast* **14:** 1051–1059.

Mourad, G.S., Tippmann-Crosby, J., Hunt, K.A., Gicheru, Y., Bade, K., Mansfield, T.A., and Schultes, N.P. (2012) Genetic and molecular characterization reveals a unique nucleobase cation symporter 1 in *Arabidopsis*. *FEBS Lett* **586:** 1370–1378.

Oakley, B.R., Rinehart, J.E., Mitchell, B.L., Oakley, C.E., Carmona, C., Gray, G.L., and May, G.S. (1987) Cloning, mapping and molecular analysis of the *pyrG* (orotidine-5′-phosphate decarboxylase) gene of *Aspergillus nidulans*. *Gene* **61:** 385–399.

Ohno, S.S. (1970) *Evolution by Gene Duplication*. New York: Springer.

Paluszynski, J.P., Klassen, R., Rohe, M., and Meinhardt, F. (2006) Various cytosine/adenine permease homologues are involved in the toxicity of 5-fluorocytosine in *Saccharomyces cerevisiae*. *Yeast* **23:** 707–715.

Pantazopoulou, A., and Diallinas, G. (2007) Fungal nucleobase transporters. *FEMS Microbiol Rev* **31:** 657–675.

Papageorgiou, I., Gournas, C., Vlanti, A., Amillis, S., Pantazopoulou, A., and Diallinas, G. (2008) Specific interdomain synergy in the UapA transporter determines its

unique specificity for uric acid among NAT carriers. *J Mol Biol* **382:** 1121–1135.
Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5:** e9490.
Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., *et al*. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29:** 845–854.
Punt, P.J., Dingemanse, M.A., Kuyvenhoven, A., Soede, R.D., Pouwels, P.H., and van den Hondel, C.A. (1990) Functional elements in the promoter region of the *Aspergillus nidulans gpdA* gene encoding glyceraldehyde-3-phosphate dehydrogenase. *Gene* **93:** 101–109.
Reiterer, V., Eyers, P.A., and Farhan, H. (2014) Day of the dead: pseudokinases and pseudophosphatases in physiology and disease. *Trends Cell Biol* **24:** 489–505.
Sá-Pessoa, J., Amillis, S., Casal, M., and Diallinas, G. (2015) Expression and specificity profile of the major acetate transporter AcpA in *Aspergillus nidulans*. *Fungal Genet Biol* doi: 10.1016/j.fgb.2015.02.010.
Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Second edn. NY: Cold Spring Harbor Laboratory, Cold Spring Harbor.
Schein, J.R., Hunt, K.A., Minton, J.A., Schultes, N.P., and Mourad, G.S. (2013) The nucleobase cation symporter 1 of *Chlamydomonas reinhardtii* and that of the evolutionarily distant *Arabidopsis thaliana* display parallel function and establish a plant-specific solute transport profile. *Plant Physiol Biochem* **70:** 52–60.
Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9:** 671–675.
Sherman, W., Day, T., Jacobson, M.P., Friesner, R.A., and Farid, R. (2006) Novel procedure for modelling ligand/receptor induced fit effects. *J Med Chem* **49:** 534–553.
Shi, Y. (2013) Common folds and transport mechanisms of secondary active transporters. *Annu Rev Biophys* **42:** 51–72.
Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N.G., Hadden, J.M., Sharples, D., *et al*. (2010) Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* **328:** 470–473.
Shindyalov, I.N., and Bourne, P.E. (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* **11:** 739–747.
Sibthorp, C., Wu, H., Cowley, G., Wong, P.W., Palaima, P., Morozov, I.Y., *et al*. (2013) Transcriptome analysis of the filamentous fungus *Aspergillus nidulans* directed to the global identification of promoters. *BMC Genomics* **14:** 847.
Sievers, F., Wilm, A., Dineen, D.G., Gibson, T.J., Karplus, K., Li, W., *et al*. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7:** 539.
Simmons, K.J., Jackson, S.M., Brueckner, F., Patching, S.G., Beckstein, O., Ivanova, E., *et al*. (2014) Molecular mechanism of ligand recognition by membrane transport protein, Mhp1. *EMBO J* **33:** 1831–1844.
Singleton, C.K. (1997) Identification and characterization of the thiamine transporter gene of *Saccharomyces cerevisiae*. *Gene* **199:** 111–121.
Stamatakis, A. (2014) RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312–1313.
Stolz, J., and Vielreicher, M. (2003) Tpn1p, the plasma membrane vitamin B6 transporter of *Saccharomyces cerevisiae*. *J Biol Chem* **278:** 18990–18996.
Subramanian, A., Kaufmann, M., and Morgenstern, B. (2008) DIALIGN-TX: greedy and progressive approaches for segment-based multiple sequence alignment. *Algorithms Mol Biol* **3:** 6.
Valley, C.C., Cembran, A., Perlmutter, J.D., Lewis, A.K., Labello, N.P., Gao, J., and Sachs, J.N. (2012) The methionine-aromatic motif plays a unique role in stabilizing protein structure. *J Biol Chem* **287:** 34979–34991.
Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J.C., *et al*. (2011) R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nuc Acids Res* **39:** W511–W517.
Vickers, M.F., Yao, S.Y., Baldwin, S.A., Young, J.D., and Cass, C.E. (2000) Nucleoside transporter proteins of *Saccharomyces cerevisiae*. Demonstration of a transporter (FUI1) with high uridine selectivity in plasma membranes and a transporter (FUN26) with broad nucleoside selectivity in intracellular membranes. *J Biol Chem* **275:** 25931–25938.
Vlanti, A., and Diallinas, G. (2008) The *Aspergillus nidulans* FcyB cytosine-purine scavenger is highly expressed during germination and in reproductive compartments and is downregulated by endocytosis. *Mol Microbiol* **68:** 959–977.
Wallace, I.M., O'Sullivan, O., Higgins, D.G., and Notredame, C. (2006) M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res* **34:** 1692–1699.
Wang, F., Becker, J.P., Cieplak, P., and Dupradeau, F.-Y. (2013) R.E.D. Python: Object oriented programming for Amber force fields, Université de Picardie – Jules Verne, Sanford Burnham Medical Research Institute.
Weber, E., Rodriguez, C., Chevallier, M.R., and Jund, R. (1990) The purine-cytosine permease gene of *Saccharomyces cerevisiae*: primary structure and deduced protein sequence of the FCY2 gene product. *Mol Microbiol* **4:** 585–596.
Weiss, J.N. (1997) The Hill equation revisited: uses and misuses. *FASEB J* **11:** 835–841.
Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., *et al*. (2008) Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322:** 709–713.
Witz, S., Jung, B., Fürst, S., and Möhlmann, T. (2012) De novo pyrimidine nucleotide synthesis mainly occurs outside of plastids, but a previously undiscovered nucleobase importer provides substrates for the essential salvage pathway in *Arabidopsis*. *Plant Cell* **24:** 1549–1559.
Witz, S., Panwar, P., Schober, M., Deppe, J., Pasha, F.A., Lemieux, M.J., and Möhlmann, T. (2014) Structure-function relationship of a plant NCS1 member – homology modelling and mutagenesis identified residues critical for substrate specificity of PLUTO, a nucleobase transporter from *Arabidopsis*. *PLoS ONE* **9:** e91343.
Yaegashi, J., Oakley, B.R., and Wang, C.C. (2014) Recent

advances in genome mining of secondary metabolite biosynthetic gene clusters and the development of heterologous expression systems in *Aspergillus nidulans*. *J Ind Microbiol Biotechnol* **41:** 433–442.
Yarov-Yarovoy, V., Schonbrun, J., and Baker, D. (2006) Multipass membrane protein structure prediction using Rosetta. *Proteins* **62:** 1010–1025.
Yoo, H.S., Cunningham, T.S., and Cooper, T.G. (1992) The allantoin and uracil permease gene sequences of *Saccharomyces cerevisiae* are nearly identical. *Yeast* **8:** 997–1006.
Young, J.D., Yao, S.Y., Baldwin, J.M., Cass, C.E., and Baldwin, S.A. (2013) The human concentrative and equilibrative nucleoside transporter families, SLC28 and SLC29. *Mol Aspects Med* **34:** 529–547.
Zettl, M., Adrain, C., Strisovsky, K., Lastun, V., and Freeman, M. (2011) Rhomboid family pseudoproteases use the ER quality control machinery to regulate intercellular signaling. *Cell* **145:** 79–91.
Zgarbová, M., Otyepka, M., Šponer, J., Mládek, A., Banáš, P., Cheatham, T.E., III, and Jurečka, P. (2011) Refinement of the Cornell *et al*. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J Chem Theory Comput* **7:** 2886–2902.
Zhao, C., and Noskov, S.Y. (2013) The molecular mechanism of ion-dependent gating in secondary transporters. *PLoS Comput Biol* **9:** e1003296.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.