National and Kapodistrian
University of Athens

School of Science
Faculty of Biology

Structure-function relationships in transmembrane transporters

PhD thesis

Athanasia Vasiliki Kourkoulou
Biologist

Athens 2021

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

Εθνικόν και Καποδιστριακόν

Πανεπιστήμιον Αθηνών

Σχολή Θετικών Επιστημών

Τμήμα Βιολογίας

Τομέας Βοτανικής

Σχέσεις δομής-λειτουργίας διαμεμβρανικών μεταφορέων

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Αθανασία Βασιλική Κούρκουλου

Βιολόγος

Αθήνα, 2021

«Η έγκριση της Διδακτορικής Διατριβής από το Τμήμα Βιολογίας της Σχολής Θετικών Επιστημών του ΕΚΠΑ δεν υποδηλώνει αποδοχή των απόψεων του συγγραφέα (ν. 5343/1932, άρθρο 202)» και «Το κείμενο της Διδακτορικής Διατριβής δεν αποτελεί προϊόν λογοκλοπής.»

Τριμελής Συμβουλευτική Επιτροπή

Γ. Διαλλινάς, Καθηγητής Πανεπιστημίου Αθηνών (Επιβλέπων)

B. Byrne, Καθηγήτρια Imperial College London

Ε. Φριλίγγος, Καθηγητής Πανεπιστήμιο Ιωαννίνων

Επταμελής Εξεταστική Επιτροπή

Γ. Διαλλινάς, Καθηγητής Πανεπιστημίου Αθηνών (Επιβλέπων)

B. Byrne, Καθηγήτρια Imperial College London

Ε. Φριλίγγος, Καθηγητής, Πανεπιστημίου Ιωαννίνων

Ε. Μικρός, Καθηγητής Πανεπιστημίου Αθηνών

Δ.Ι. Στραβοπόδης, Αναπληρωτής Καθηγητής Πανεπιστημίου Αθηνών

Κ. Χαραλαμπίδης, Αναπληρωτής Καθηγητής Πανεπιστημίου Αθηνών

Χ. Γουρνάς, Ερευνητής Γ, ΕΚΕΦΕ Δημόκριτος

Η υλοποίηση της διδακτορικής διατριβής χρηματοδοτήθηκε από ερευνητικό πρόγραμμα και υποτροφία του Ιδρύματος Σταύρος Νιάρχος (ΙΣΝ).

HELLENIC REPUBLIC

National and Kapodistrian

University of Athens

School of Science

Faculty of Biology

Department of Botany

Structure-function relationships in transmembrane transporters

Ph.D. Thesis

Athanasia Vasiliki Kourkoulou

Biologist

Athens 2021

Advisory Committee

G. Diallinas, Professor University of Athens (Supervisor)

B. Byrne, Professor Imperial College London

E. Frillingos, Professor University of Ioannina

Examining Committee

G. Diallinas, Professor University of Athens (Supervisor)

B. Byrne, Professor Imperial College London

E. Frillingos, Professor University of Ioannina

E. Mikros, Professor University of Athens

D.J. Stravopodis, Associate Professor University of Athens

K. Haralampidis, Associate Professor University of Athens

C. Gournas Associate Researcher NCSR Demokritos

This research has been funded by a research grant including a scholarship from Stavros Niarchos Foundation (SNF).

Περίληψη

Οι μεταφορείς είναι διαμεμβρανικές πρωτεΐνες που πραγματοποιούν την επιλεκτική μεταφορά ουσιών μέσω των μεμβρανών. Τα μέλη της οικογένειας μεταφορέων NAT (Nucleobase Ascorbate Transporter family) είναι συμμεταφορείς H^+ ή Na^+ ειδικοί για την πρόσληψη είτε πουρινών και πυριμιδινών είτε L-ασκορβικού οξέος [1]. Παρά το γεγονός ότι αρκετά μέλη έχουν μελετηθεί εκτενώς σε γενετικό, βιοχημικό και κυτταρικό επίπεδο και οι δομές μελών από το βακτήριο *Escherichia coli* and τον μύκητα *Aspergillus nidulans* έχουν δημοσιευθεί παρέχοντας μια πληθώρα δεδομένων για τον μηχανισμό λειτουργίας, τα υπάρχοντα δεδομένα δεν είναι ικανά να εξηγήσουν πλήρως το πως καθορίζεται η εκλεκτικότητα των υποστρωμάτων [2,3]. Καλά χαρακτηρισμένα μέλη από τα βακτήρια, τους μύκητες και τα φυτά μεταφέρουν ειδικά πουρίνες η/και πυριμιδίνες ενώ τα θηλαστικά και άλλα σπονδυλωτά διαθέτουν μέλη που είναι ειδικά για L-ασκορβικό οξύ (SVCT1/2) αλλά και μέλη ειδικά για νουκλεοτιδικές βάσεις (π.χ. rSNBT1) [4,5]. Τα σπονδυλωτά διαθέτουν ένα επιπλέον παράλογο άγνωστης λειτουργιας (SVCT3) [5].

Οι δομές από δυο μέλη της NAT οικογένειας είναι γνωστές [2,3,6]. Αυτές είναι η δομή του μεταφορέα ουρακίλης της *E.coli* UraA και του μεταφορέα ουρικού οξέος-ξανθίνης του *A. nidulans* UapA. Και οι δυο πρωτεΐνες αποτελούνται από 14 διαμεμβρανικά τμήματα που χαρακτηρίζονται από δύο ανεστραμμένες επαναλήψεις (7+7) που αντιστοιχούν σε δύο επικράτειες, την επικράτεια πυρήνα (core domain) και την επικράτεια διμερισμού (dimerization domain). Και οι δύο πρωτεΐνες σχηματίζουν διμερή, ο σχηματισμός των οποίων είναι απαραίτητος για την λειτουργία των μεταφορέων. Ο UapA είναι ένας συμμεταφορέας ουρικού οξέος-ξανθίνης/H^+ του μύκητα *A. nidulans* και θεωρείται το πρότυπο, ευκαρυωτικό μέλος αυτής της οικογένειας επειδή είναι ένας από τους πιο εκτενώς χαρακτηρισμένους ευκαρυωτικούς μεταφορείς σε ότι αφορά τις σχέσεις-δομής λειτουργίας, την εκλεκτικότητα υποστρώματος, την ρύθμιση της έκφρασης και την υποκυτταρική διακίνηση [7].

Όλοι οι NAT μεταφορείς περιέχουν ένα συντηρημένο μοτίβο στο $10^ο$ διαμεμβρανικό τμήμα που ονομάστηκε ιστορικά σαν αλληλουχία-αναγνώρισης NAT (NAT signature motif) το οποίο περιλαμβάνει κατάλοιπα που είναι απαραίτητα για την δέσμευση υποστρώματος και την εκλεκτικότητα ή για την κατάλυση της μεταφοράς [1,8]. Προηγούμενες μελέτες στον UapA έδειξαν ότι οι περισσότερες μεταλλαγές που επηρεάζουν την εκλεκτικότητα του, που έχουν προκύψει από τυχαίες μεταλλαξιγενέσεις, βρίσκονται εκτός της θέσης δέσμευσης υποστρώματος και της αλληλουχίας-μοτίβου NAT [9–11]. Από αυτές οι πιο γνωστές

μεταλλαγές αφορούν τα κατάλοιπα Arg^{481}, Thr^{526} και Phe^{528} που εντοπίζονται κατά μήκος της πορείας κύλισης της επικράτειας πυρήνα πάνω στο διμερές.

Επιπλέον, πρόσφατες μελέτες έδειξαν ότι συγκεκριμένες αληλλεπιδράσεις του UapA με μεμβρανικά λιπίδια στην επιφάνεια διμερισμού είναι απαραίτητες για τον σχηματισμό ή/και την σταθερότητα των λειτουργικών διμερών [12]. Πιο συγκεκριμένα, κατά την διαδικασία απομόνωσης του UapA συγκατακρημνίζονται και λιπίδια και η απομάκρυνση αυτών των λιπιδίων οδηγεί σε διάλυση του διμερούς συμπλόκου σε μονομερή. Προσθήκη φωσφο-ινοσιτιδίων (PIs) ή φωσφατιδυλοαιθανολαμίνης (PEs) οδήγησε στον επανασχηματισμό του διμερούς. Προσομοιώσεις μοριακής δυναμικής (MDs) προέβλεψαν την ύπαρξη μιας ειδικής θέσης δέσμευσης λιπιδίων στην επιφάνεια διμερισμού που αποτελείται από τρία κατάλοιπα αργινίνης Arg^{287}, Arg^{478} και Arg^{479}. Η αντικατάσταση αυτών των καταλοίπων οδήγησε σε πλήρη απώλεια λειτουργίας η οποία οφείλεται στην απώλεια σχηματισμού του λειτουργικού διμερούς σε μεγάλο ποσοστό της πρωτεΐνης όπως αποδείχτηκε από φασματομετρία μάζας (native MS) και δοκιμασίες εντοπισμού πρωτεϊνικών αλληλεπιδράσεων με το σύστημα BiFC.

Η παρούσα διατριβή είναι χωρισμένη σε τρία κεφάλαια. Στο πρώτο ερευνήθηκε η μοριακή βάση της εξειδίκευσης υποστρώματος στην NAT οικογένεια και μελετήθηκε η εξέλιξη των μεταφορέων ασκορβικού πραγματοποιώντας αρχικά μια εκτενή φυλογενετική ανάλυση και στη συνέχεια μεταλλαγές στο μοτίβο NAT του UapA. Την παραπάνω συστηματική μεταλλαξιγένεση ακολούθησε ορθολογικά σχεδιασμένος συνδυασμός υποκαταστάσεων ενώ απομονώθηκαν νέες επιπλέον υποκαταστάσεις μέσω τυχαίων μεταλλαξιγενέσεων. Τα αποτελέσματα συνολικά υποστηρίζουν ότι ο ρόλος κάποιων μερικώς συντηρημένων καταλοίπων του μοτίβου NAT στην εξειδίκευση του μεταφορέα UapA εξαρτάται από την ύπαρξη συγκεκριμένων αμινοξέων σε άλλες θέσεις. Επιπλέον παρουσιάζονται νέα δεδομένα για το πώς το κατάλοιπο Phe^{528}, που βρίσκεται εκτός της θέσης πρόσδεσης υποστρώματος, μπορεί να επηρεάζει την εκλεκτικότητα του UapA.

Το δεύτερο μέρος αυτής της διατριβής αφορά τον ρόλο των αλληλεπιδράσεων του UapA με λιπίδια στη λειτουργία, τη σταθερότητα και τη μεταφορά του στη μεμβράνη. Πιο συγκεκριμένα, εξετάστηκε περαιτέρω ο ρόλος των αλληλεπιδράσεων στην επιφάνεια διμερισμού και διερευνήθηκε ο πιθανός ρόλος άλλων αλληλεπιδράσεων, που έχουν προβλεφθεί από MDs, στην περιφέρεια της επικράτειας πυρήνα του UapA. Βρέθηκε πως διακριτές αλληλεπιδράσεις του UapA με μεμβρανικά λιπίδια είναι απαραίτητες για τον εξαρχής σχηματισμό διμερών στο ενδοπλασματικό δίκτυο, ή την έξοδο από αυτό και την

περαιτέρω στόχευση του στη μεμβράνη. Επιπλέον, μέσω τυχαίων μεταλλαξιγενέσεων απομονώθηκαν μεταλλαγές που επαναφέρουν τον σχηματισμό διμερών ή/και τη στόχευση στη μεμβράνη.

Τέλος, στο τρίτο μέρος, χρησιμοποιώντας αποτελέσματα της παρούσας διατριβής έγινε για πρώτη φορά λειτουργική ετερόλογη έκφραση μιας NAT ομόλογης πρωτεΐνης από τα θηλαστικά στον *A. nidulans*.

Abstract

Transporters are transmembrane proteins that mediate the selective translocation of solutes across biological membranes. Members of the ubiquitous Nucleobase Ascorbate Transporter (NAT) family are H^+ or Na^+ symporters specific for the cellular uptake of either purines and pyrimidines or L-ascorbic acid [1]. Despite the fact that several members have been extensively characterized at a genetic, biochemical or cellular level, and crystal structures of NAT members from *Escherichia coli* and *Aspergillus nidulans* have been determined pointing to a mechanism of transport, the current knowledge cannot explain how substrate selectivity is determined [2,3]. Functionally characterized NATs from bacteria, fungi and plants are specific for nucleobases, but rather surprisingly, mammals and other vertebrates possess NAT homologues that are specific for L-ascorbate transport (SVCT1/2) in addition to nucleobase specific members (i.e. rSNBT1) [4,5]. Vertebrates also include a third distinct paralogue of unknown function called SVCT3 [5].

High-resolution crystal structures from two NAT members have been obtained [2,3,6]. These are the UraA uracil transporter of *E. coli* and the UapA uric acid-xanthine transporter of *A. nidulans*. Both proteins are composed of 14 transmembrane segments characterized by a 7 helix inverted repeat (7+7) forming a *core* and a *dimerization* domain. Additionally, these proteins exist as dimers, the formation of which is essential for transport activity. UapA is considered the prototypic eukaryotic member of this family as it is one of the most extensively studied eukaryotic transporters in respect to structure-function relationships, substrate specificity, regulation of expression and subcellular trafficking.

All NATs include a highly conserved motif in TMS10 historically referred as the NAT signature motif which includes residues critical for substrate binding and specificity or transport catalysis [1,8]. Previous studies on UapA reported that most specificity substitutions in UapA, selected by direct genetic screens, map outside the major substrate binding site and the NAT signature motif [9,10,13]. The most prominent specificity substitutions concerned residues Arg^{481}, Thr^{526} or Phe^{528}, which are located along the proposed sliding trajectory of the core domain in the UapA dimer.

Moreover, it has been shown recently that specific interactions with plasma membrane phospholipids at the dimer interface of UapA are essential for the formation and/or stability of functional dimers [12]. More specifically, UapA co-purifies with lipid and delipidation results in dissociation into monomers. Addition of PIs or PEs resulted in the re-

formation of the UapA dimer. MDs predicted a specific lipid binding site at the dimer interface that is formed by three arginine residues Arg^{287}, Arg^{478} and Arg^{479}. Replacement of these arginines by alanine residues led to total loss of UapA function and both native MS and bifluorescence complementation (BiFC) assays indicated that a major fraction of UapA cannot dimerize.

This study is divided in three distinct chapters. In the first one the molecular aspects of NAT substrate specificity and the evolution of ascorbate transporters were investigated by making at first an extensive phylogenetic analysis and then a mutational analysis on the NAT signature motif of UapA. This mutational analysis was coupled with a rational combination of substitutions while new substitutions were also isolated by novel genetic screens. Overall, the results revealed cryptic context-dependent roles of partially conserved residues in the NAT signature motif in determining the specificity of the UapA transporter. Additionally we provided novel findings concerning how Phe528, a residue outside the substrate binding site, might function as a key amino acid in determining UapA specificity.

The second part of the study was focused on the role of lipid interactions in UapA function, stability and trafficking. In particular, the role of UapA-lipid interactions at the dimer interface was further examined and the possible role of other predicted, by MDs, interactions at the membrane facing regions of the core domain of the UapA dimer was identified. We found that distinct interactions of UapA with membrane lipids are essential for *ab initio* formation of functional dimers in the ER, or ER exit and further subcellular trafficking. Additionally, through genetic screens, we identified substitutions that restore defects in dimer formation and/or trafficking.

Finally, in the third part of the present study, using knowledge acquired from this work we achieved for the first time the functional expression of a mammalian NAT homologue in *A. nidulans*.

Acknowledgements

First of all I am grateful to my supervisor Professor George Diallinas, for his guidance and support and for his motivation to transfer his passion for science to me but also others. He made my Ph.D. experience productive and gave me the opportunity to grow as a scientist. Last but not least, I want to thank him for what he taught me about food and the wonderful experiences I have gained during travelling.

Next I want to thank the rest of my Advisory committee, Professor Bernadette Byrne and Professor Stathis Frillingos for our collaboration and their support, suggestions and encouragement during all these years.

I am also grateful to everyone that I collaborated with and contributed to this work. These are Alexandros Pittis for our collaboration in my first paper, Professor Bernadette Byrne and her lab members for our collaboration in the lipid-interactions project, Professor Emmanuel Mikros and his team, George Lambrinidis and Iliana Zantza, for our collaboration over these years and finally my lab “kids” Pothos Grevias and Konstantina Foti for our collaboration and friendship but also their support, kindness and enthusiasm during the second and third year of my thesis.

I would also like to thank Dr. Sotiris Amillis, for his guidance, support and constant willingness to help me in every aspect of my lab life. I also want to thank him for our excellent collaboration and his patience and encouragement during a scientific project that is not described herein and it was one of my favorites.

My sincere thanks also goes to Professor E. Mikros, Associate Professor D.J. Stravopodis, Associate Professor K. Haralampidis and Associate Researcher C. Gournas, for serving as my Examining Commitee.

Additionally, I can’t thank enough Stavros Niarchos Foundation for giving me a scholarship and providing the financial support and thus giving me the opportunity to complete my Ph.D.

I am also thankful to all my fellow labmates for the team spirit and the atmosphere we created during working together in the lab. A special thanks goes to Dr. Georgia Papadaki, as she was my lab “teacher” during my undergraduate thesis and a labmate for all the three years of my Ph.D thesis. I thank also everyone from the Microbiology team (previous and current members) for being available to help me over the years and our small celebrations. A

special thanks goes to my Marilena Koukounia for her support and friendship inside and outside the lab during these four years that I know her.

Finally a great thanks goes to my family and friends for being there for me over these three years. First of all to my parents for literally doing everything for me and showing me what it is to be loved. And finally to all my friends and especially George, Artemis and Marilena for being a significant influence in my life, supporting me through everything, encouraging me during this experience and most importantly making me happy by being part of my life.

Contents

Περίληψη 9
Abstract 13
Acknowledgements 15
Contents 17
Abbreviations 19
Introduction 23
1.1 Transporters 23
1.2 Study of purine/pyrimidine transporters of the model fungus *A. nidulans* 35
1.3 Structure-function relationships in the NAT/NCS2 family 39
1.4 Transporter-lipid interactions 49
1.5 Goals of the study 53
Materials and methods 55
2.1 Strains and growth conditions 55
2.2 Genetic crosses and progeny analysis 60
2.3 Epifluorescence microscopy 61
2.4 DNA manipulations and molecular cloning 61
2.5 Transformation of *A. nidulans* 68
2.6 RNA extraction from *A. nidulans* 70
2.7 Northern blot analysis 70
2.8 Native MS of UapA 72
2.9 Phylogenetic analysis datasets and methods 73
2.10 Molecular simulations shown in chapter 3.2 74
2.11 Molecular simulations shown in chapter 3.3 and 3.4 76
Results 78
3.1 Evolution of substrate specificity in the Nucleobase Ascorbate Transporter (NAT) protein family 78
3.2 Context-dependent cryptic roles of specific residues in substrate selectivity of the UapA purine transporter 89
3.3. Specific residues in a purine transporter are critical for dimerization, ER exit, and function 107
3.4 Expression of mammalian NAT homologues in *A. nidulans* 119
Discussion 124
4.1 Evolution of L-ascorbate transporters 124

4.2 Context-dependent cryptic roles of residues in UapA substrate specificity 126

4.3 The role of Phe^{528} and other specificity residues- 'locks' or 'brakes' of transport dynamics 127

4.4 Lipid interactions are critical for UapA dimerization and ER exit..... 129

4.5 The successful functional expression of rSNBT1 in *A. nidulans* leads to new routes for NAT study..... 131

References..... 133

Appendix..... 153

CV and reprints of original publications..... 153

Abbreviations

5FC: 5-fluorocytosine

5FU: 5-fluorouracil

5FUd: 5-fluorouridine

AAA: ATP:ADP Antiporter family

AD: Adenine

ADO: Adenosine

AE: Anion Exchanger family

APC: Amino acid-Polyamine-organocation superfamily

ATP: Adenosine-Triphosphate

BiFC: Bimolecular Fluorescence Complementation assay

bp: Base pairs

BSA: Bovine Serum Albumin

°C: Celsius degrees

CE: Collision Energy

CM: Complete Media

COPII: Coat Protein Complex II

DDM: Dodecyl-β-D-maltoside

DEPC: Diethyl pyrocarbonate

DMSO: Dimethyl sulfoxide

DPPC: Dipalmitoylphosphatidylcholine

EDTA: Ethylenediaminetetraacetic Acid

ER: Endoplasmic Reticulum

EtBr: Ethidium Bromide

EtOH: Ethanol

fs: femtoseconds

GFP: Green Fluorescent Protein

GPH: Glycoside-Pentoside-Hexuronide Cation symporter family

GUO: Guanosine

h: Hour

H_2O_{dist}: Distilled water

HP: Helical hairpin

HPLC: High Performance Liquid Chromatography

HX: Hypoxanthine

INS: Inosine

K_m: Michaelis-Menten constant

LB: Luria-Bertani medium

MD: Molecular Dynamics

MFS: Major Facilitator Superfamily

min: Minute

ML: Maximum Likelihood

MM: Minimal Media

MS: Mass Spectrometry

NAT: Nucleobase Ascorbate Transporter family

NCBI: National Center for Biotechnology Information

NCS1: Nucleobase Cation Symporter-1 family

NCS2: Nucleobase Cation Symporter-2 family

ND: Not Determined

ns: nanoseconds

NSS: Neurotransmitter Sodium Symporter family

OD: Optical Density

ORF: Open Reading Frame

paba: p-aminobenzoic acid

panto: D-pantothenic acid

PC: Phosphatidylcholine

PCR: Polymerase Chain Reaction

PDB: Protein Data Bank

PE: Phosphatidylethanolamine

PI: Phosphatidylinositol

PIC: Protease Inhibitors Cocktail

PM: Plasma Membrane

PRO: Proline

PTS: Phosphotransferase System

pyro: Pyridoxine

ribo: Riboflavin

RMSD: Root Mean Square Deviation

rpm: Rounds per minute

RT: Room Temperature

SD: Standard Deviation

SDS: Sodium Dodecyl Sulfate

SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

sec: Second

SGP: Spore Germination Protein subfamily

SLC: Solute Carrier family

SulP: Sulfate Permease family

T: Thymine

TCA: Trichloroacetic Acid

TEV: Tobacco Etch Virus

T_m: Melting temperature

TMS: Transmembrane segments

U: Uracil

UA: Uric acid

UV: Ultraviolet

wt: Wild Type

X: Xanthine

YFP: Yellow Fluorescent Protein

Introduction

1.1 Transporters

1.1.1 Membrane transport systems

Nutrients, ions and generally most micro- or macro-polar molecules are dependent on membrane proteins in order to enter or exit the cell. Simple diffusion of molecules through the membrane is possible only for non-polar molecules such as steroids, O_2 and CO_2 and small uncharged polar molecules such as ethanol and glycerol (Figure 1.1). Most of the other molecules move through membranes via two types of transmembrane transport proteins, *transporters* and *channels/pores*. The major difference concerns their mechanism of function (see later) [7,14]. Transport proteins fulfill essential functions in living cells as they mediate the exchange of a large variety of molecules but also play an important role in other physiological and pathophysiological processes such as signal transduction and mechanosensing. Their importance is reflected by their association with many diseases including autism, epilepsy, cystic fibrosis and neurodegeneration ([15–17] and references therein). Transporters are also involved directly in drug sensitivity or resistance as gateways or efflux proteins. It is worth to mention that the targets of the two of the most widely prescribed drugs in the world, fluoxetine (Prozac) and omeprazole (Prilosec) are membrane transport proteins.

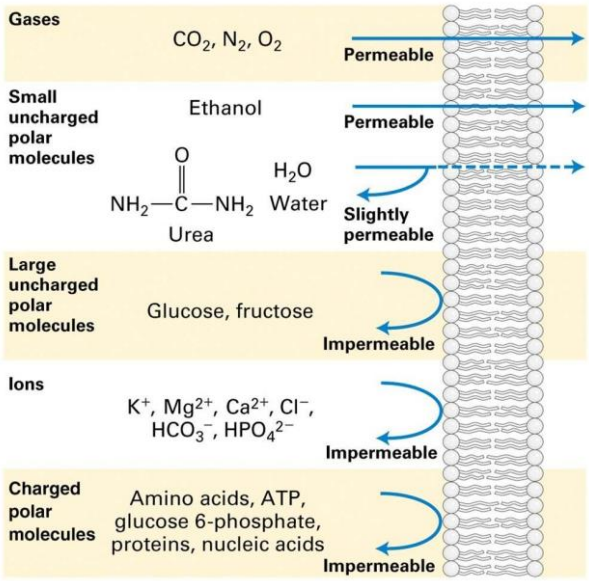


Figure 1.1. Membrane permeability graph. This figure is adapted from https://www.studyblue.com/notes/note/n/chapter-4/deck/5796523.

Transporters, also known as permeases or carriers, are polytopic transmembrane proteins that comprise a substrate binding site and their transport mechanism includes major conformational changes that result to either a structure being open to the one side of the membrane or to the other (Figure 1.2). In other words, the binding site is not accessible from both sides of the membrane simultaneously [7,14]. On the contrary, channels/pores do not have binding sites and can be open to both environments at the same time (Figure 1.2). Channels/pores are made by several transmembrane segments (TMSs) like transporters but often two or more subunits associate to create the functional unit. The term channels refers to proteins that are either in an open or a closed state whereas pores are continuously open and active. Opening and closing of channels is regulated by *gating domains or gates* in response to signals like ligand binding or membrane voltage changes [7,14,18]. As channels do not possess a binding site, selectivity is controlled by *selectivity filters* that restrict the size of the hole. Their charge is also adapted to the substrate ion. As a result of their mechanism channels/pores are much faster than transporters, translocating up to 10^8 molecules/s while transporters have a relatively low turnover rate with approximately 200-50000 molecules/s [19].

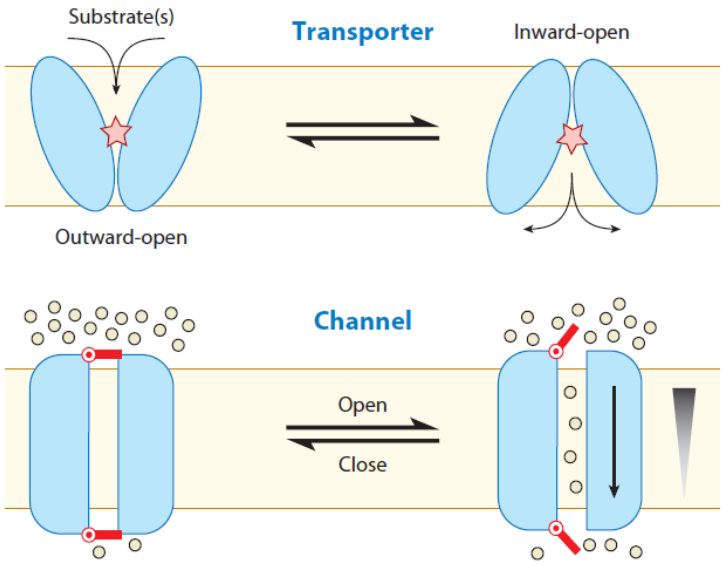


Figure 1.2. **Transporters vs Channels**. Schematic representation of the mechanism of function of a transporter (top) and a channel (bottom). Transport mechanism of a transporter includes major conformational changes so that the binding site is not accessible from both sides of the membrane simultaneously. Channels, on the other hand, have an open and a close state and when open substrate has access from both sides of the membrane and molecules move from the area of higher to lower concentration. This figure is adapted from [20].

Based on the consumption of energy, a transport system can be classified as *passive* or *active*. Passive transport does not have a requirement for energy spending as it concerns transport of molecules down their electrochemical-concentration gradient. Active transport instead requires energy in order to move molecules from an area of lower to higher concentration. Energy can be provided either directly (*primary active transport*) for example from ATP-hydrolysis or from the electrochemical potential difference (*secondary active transport*). Passive transport can be performed by both transporters (facilitated diffusion) and channels (passive diffusion) whereas active transport is only executed by transporters. In the first case transporters are called *facilitators*. Primary active transport is performed by transporters that most of them are also ATPases proteins hydrolyzing ATP to provide the energy needed for transport. There are four types of primary active transporters: P-type ATPases also referred as *pumps* e.g. Na^+-K^+ ATPase, F-ATPases e.g. mitochondrial ATP synthase, V-ATPases and ABC transporters e.g. CFTR. On the other hand, secondary active transport is carried out by transporters that couple the movement of a driving ion down its electrochemical gradient to the movement of another molecule or ion against its concentration gradient. It is considered as a type of active transport as it takes advantage of the electrochemical gradients made by primary active transport. Secondary active transporters can be *symporters*, such as the glucose symporter SGLT1, if transport of the substrates is in the same direction or *antiporters* if the substrates move in opposite directions. Figure 1.3 summarizes all the transport systems discussed above.

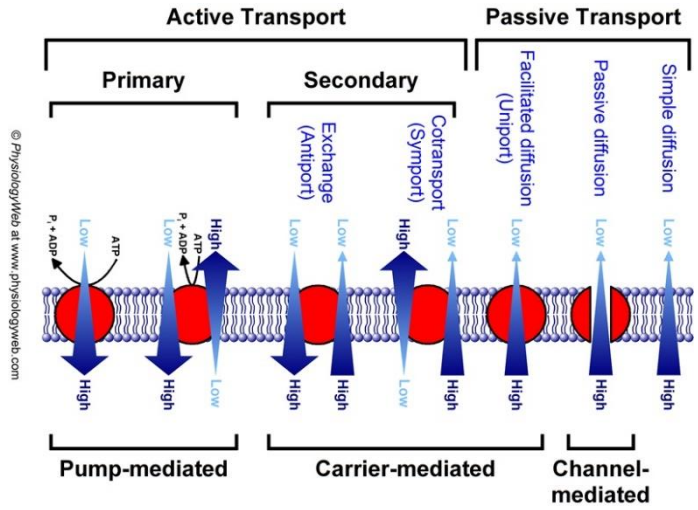


Figure 1.3. Mechanisms of membrane transport. Blue arrows indicate the concentration gradient of the transported molecule. The term carrier historically refers to the transport mechanism of transporters. Refers to a protein that uploads substrates from one side of the membrane, "swims" across the membrane, and translocate them to the other side [20]. This figure is adapted from http://www.physiologyweb.com.

1.1.2 Transporter classification

As mentioned before, membrane transport is accomplished by a wide variety of transmembrane proteins. Transporters represent a diverse group of proteins functioning as facilitators, pumps, group translocators, electron carriers, symporters and antiporters. Their study has grown dramatically over the past few years mainly due to progresses in the experimental methods and their biological and medical importance. Secondary active transporters are related with many human diseases and drug resistance/sensitivity as they transport a large variety of substrates involved in many physiological processes.

The Transporter Classification Database or TCDB has classified all kind of transport proteins based on a functional/phylogenetic hierarchical system [21]. All proteins are classified in classes, subclasses, families, subfamilies and transport systems. Over the years several families have been organized into sequence divergent superfamilies. The two largest superfamilies of transporters are the ATP-binding cassette (ABC) superfamily that consists of primary active transporters and the Major Facilitator Superfamily (MFS) which is mostly constituted by secondary active transporters. Of particular interest is also the Amino acid-Polyamine-organoCation (APC) superfamily that is the second largest superfamily of secondary active transporters.

1.1.2.1 Evolution of protein families

A protein family is a group of proteins that share a common ancestor, reflected by their sequence similarity which is the clearest indicator of common ancestry, related functions and structural similarities. Protein families are often organized into hierarchies, with proteins subdivided into smaller, more closely related groups. The term superfamily is often used for large groups of distantly related proteins whose association is not reflected in sequence similarity, but only in shared structural features. Proteins that are closely related within a protein family belong to a protein subfamily, class etc.

Protein families are basically created by two major events. Firstly, when parent species are separated into two genetically isolated descendent species, the descendent genes/proteins can independently accumulate mutations and thus variations. The resulting proteins that share the same function in different species are called orthologs. Secondly, a gene duplication event can create a second copy of a gene that is free to acquire independent

random mutations as the original copy continues to perform its function [22,23]. In this case the two homologous genes/proteins are called paralogs and usually have similar but not identical functions. Thus duplication creates redundancy and redundancy may provide the material for novelty. In the course of evolution, several gene/protein families, especially in eukaryotes, undergo extreme expansions and contractions sometimes together with whole genome duplications.

Maintenance of the gene duplicates has been explained by different evolutionary scenarios, such as dosage effects, neo-functionalization or sub-functionalization [24]. The first scenario refers to the fixation of duplicate genes driven by positive selection for increased gene dosage [25]. Neo-functionalization is the process by which, after gene duplication, one copy retains the original function of the gene, while the second copy, free from selective pressure, acquires mutations and explores novel functions. If a new function is beneficial, natural selection positively "selects" for it and becomes fixed in the population. Finally sub-functionalization occurs when the two paralogs retain a subset of the ancestral functions. This model requires the existence of a multi-functional pre-duplication protein and introduces the possibility that positive selection pressure might drive evolution after gene duplication event. Thus novelty might also arise via side-activities or sub-functions of proteins, which are optimized in the course of evolution and under specific selective pressures.

1.1.2.2 Major Facilitator Superfamily (MFS)

The major facilitator superfamily (MFS) is a ubiquitous superfamily that includes both facilitators (uniport) and secondary active transporters with the latter being driven by the co-transport (symport or antiport) of either protons (H^+) or sodium ions (Na^+). Up to date, TCDB classifies the primary MFS group ("canonical" MFS family) into 89 families, but 15 more families, such as the Glycoside-Pentoside-Hexuronide (GPH):Cation symporter family and the ATP:ADP Antiporter (AAA) family, are also clustered together as distant members [26]. MFS transporters were initially identified as sugar transporters but today the MFS superfamily is known to consist of members moving also a variety of other small solutes such as amino acids, organic and inorganic anions, siderophores, drugs, neurotransmitters, but not macromolecules [26]. Bacterial members transport mostly nutrients, such as the well-studied lactose permease LacY, the glycerol-3-phosphate transporter GlpT and the xylose/H^+ symporter XylE, although some members function as efflux pumps for toxic compounds. Similarly, human MFS members are responsible for functions such as nutrient absorption and

transport and renal and hepatic clearance but also for transport of other molecules including signaling molecules ([27] and references therein). The best studied eukaryotic members are the glucose transporters GLUT1, GLUT2, GLUT3, and GLUT4.

The crystal structure of a number of MFS transporters, most from bacteria, have been solved. The first known structures were those of the LacY and GlpT of *E. coli* that were published in 2003 [28,29]. Recently, the structures of the first three eukaryotic members have been resolved. These include the fungal phosphate transporter PiPT [30], the plant nitrate transporter NRT1.1 [31] and the human glucose transporter GLUT1 [32]. Most proteins consist of 400-600 amino acid residues. All known MFS structures share a common protein-fold known as the MFS fold (Figure 1.4a). This is characterized by twelve transmembrane segments (TMSs) organized into two symmetrically related domains of six TMSs (6+6) that are connected by a cytoplasmic loop [33]. Each one of these two domains, referred as the N- and C-terminal domains, is made of two structurally inverted repeats composed of three segments (TMS1-3 and 4-6 in the N-domain and TMS7-9 AND 10-12 in the C-domain). In other words the MFS fold is made up of four structural repeats comprised of three TMSs. Some members may contain extra TMSs that seem to contribute to stabilization of the overall structure. Additionally, few MFS transporters are composed of 24 TMSs as a result of duplication or fusion to other functional domains resulting in proteins with dual functions [34]. The MFS transporters seem to function as monomers although there is evidence of oligomerization in some plant members [35]. For example, NRT1.1 structure revealed that this transporter can homodimerize [31]. Furthermore, the oligomerization state of this transporter is regulated by the phosphorylation of a Thr residue that seems to decouple the dimer in order to adjust affinity [31,36].

1.1.2.3 Amino acid-Polyamine-organoCation (APC) superfamily

The APC superfamily is the second largest superfamily of secondary carriers being also present in all domains of life. It consists of numerous families of transporters that move amino acids, peptides, purines, pyrimidines, metal ions, vitamins etc. and function as solute/cation symporters or solute/solute antiporters [37]. Based on TCDB, besides the APC family other 17 families are categorized in this superfamily including the Neurotransmitter Sodium Symporter (NSS) family, the Nucleobase Cation Symporter-1 (NCS1) Family, the Nucleobase/Ascorbate Transporter (NAT) or Nucleobase Cation Symporter-2 (NCS2) Family, the Sulfate Permease (SulP) Family and the Anion Exchanger (AE) Family. Members of the Spore Germination

Protein (SGP) subfamily (within the APC family) are amino acid receptors rather than transporters truncated at their C-termini.

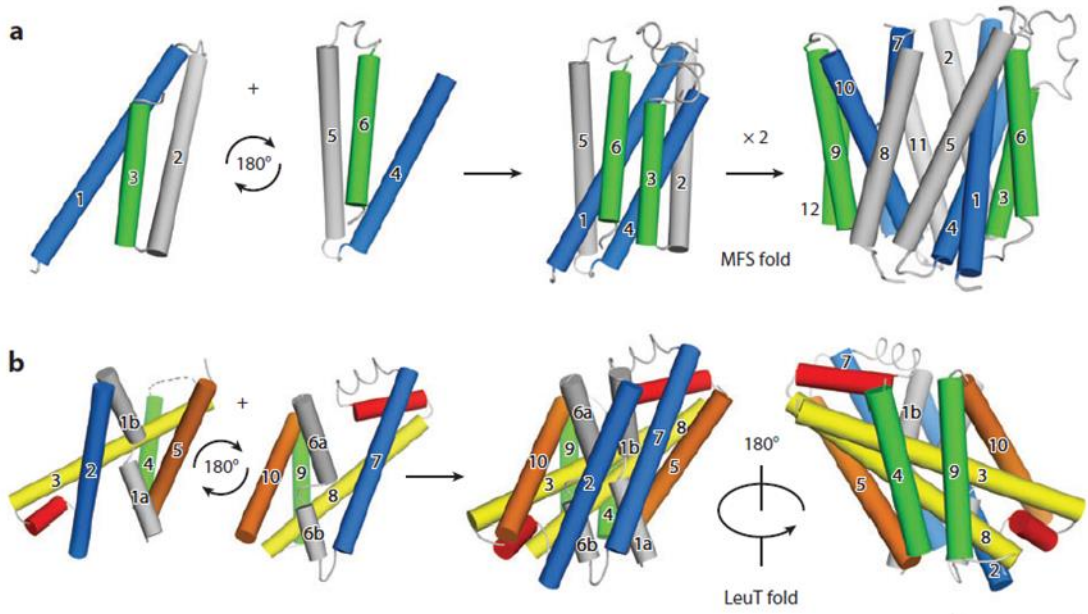


Figure 1.4. Common folds of secondary active transporters. (A) Representation of the MFS fold. (B) Representation of the LeuT fold. The figure is adapted from [20].

A large number of high resolution crystal structures of APC proteins from different families is published. The structure of most members contains a 5+5 topology known as the LeuT fold (Figure 1.4b) [38]. LeuT is an amino acid/sodium symporter from *Aquifex aeolicus*, homolog of the sodium-dependent neurotransmitter transporter from the NSS family. The LeuT fold consists of two structurally inverted repeats of five TMSs (TMSs1–5 and TMSs 6–10) that form two structurally distinct bundles referred as the scaffold and core domain [39]. The first segments of each domain (TMS1 and 6) are discontinuous, made up by two α-helices connected by a highly conserved unwound sequence that is important for the transport mechanism of transporters with the LeuT fold [40]. Two families, SulP and NCS2, display a different structure, adopting a 7+7 topology. Vastermark et al. [41] showed that both topologies share structural features confirming their possible relationship, and proposed an evolutionary pathway for their creation. The 7+7 topology is also found in other transporters classified within the APC family, such as the plant boron transporters Bor1-3 and the human Band3 anion exchanger. Apart from the 5+5 or 7+7 fold many members have extra TMSs at the N or C-terminus or in the middle of the inverted repeats.

1.1.2.3.1 Nucleobase Cation Symporter-1 (NCS1) Family

The NCS1 family includes over 1000 currently sequenced proteins from archaea, bacteria, fungi and plants but seem to be absent from other domains of life. They are H^+/Na^+ symporters specific for the uptake of purines, pyrimidines and related metabolites. NCS1 proteins are usually 419-635 residues long and possess twelve putative α-helical TMSs interconnected with rather short loops and cytosolic N- and C-termini. TMSs 1–10 are arranged as the LeuT-fold, while the last two TMSs (11 and 12) appear crucial for the oligomerization state of some NCS1-like transporters and not the transport mechanism. Fungal NCS1 transporters are classified into two sub-families, the Fcy-like and the Fur-like transporters based on their primary amino acid sequences and specificity profiles [42,43]. These two fungal NCS1 sub-families along with the plant homologues seem to originate through independent horizontal transfers from prokaryotes and probably expanded in number due to repeated duplication [43].

1.1.2.3.2 Nucleobase/Ascorbate Transporter (NAT) or Nucleobase Cation Symporter-2 (NCS2) Family

Besides NCS1, NAT/NCS2 is another distinct family that includes members highly specific for purines and pyrimidines. The NAT/NCS2 family is thought to contain two distinct clusters, one with transporters that recognize 2-oxypurines (xanthine, uric acid) and are members of the NAT (Nucleobase Ascorbate Transporter)/NCS2 family, and another with transporters recognizing 2-non-oxypurines (hypoxanthine, guanine and adenine), which form a distinct cluster known as AzgA-like and are distant structural homologues with NATs [44]. In contrast to the NCS1, the NAT family is present in all domains of life with few notable exceptions. NATs are made of 414-650 residues and are highly specific for either nucleobases/H^+ or L-ascorbate/Na^+ [1]. All bacterial NATs are highly specific for either uracil or xanthine or uric acid, while the fungal and plant members are specific for either xanthine-uric acid, or for adenine-guanine-hypoxanthine-uracil (only in plants). Mammalian NATs, that correspond to the human solute carrier family SLC23, are specific for L-ascorbate (hSVCT1 and hSVCT2 in human) however there is a case reported of a NAT transporter from rat, called rSNBT1, which is specific for nucleobases [5]. Mammals also include a third homologue called SVCT3 however its function is unknown. Moreover, NAT is one of the two families of the APC superfamily adopting a 7+7 fold. High resolution structures of the bacterial uracil transporter UraA and the eukaryotic uric acid/xanthine transporter from *A. nidulans* UapA (Figure 1.5) are published [2,6]. These structures showed that the 14 transmembrane segments are organized into a

core (TMSs 1–4 and 8–11) and a gate/dimerization domain (TMSs 5–7 and 12–14). Additionally, these structures revealed a pair of antiparallel β-strands located between TMS 3 and TMS 10 that were shown to be important for structural organization and substrate recognition. UapA structure and subsequent publications on UraA further confirmed that NATs dimerize and that the dimer is probably the functional unit [2].

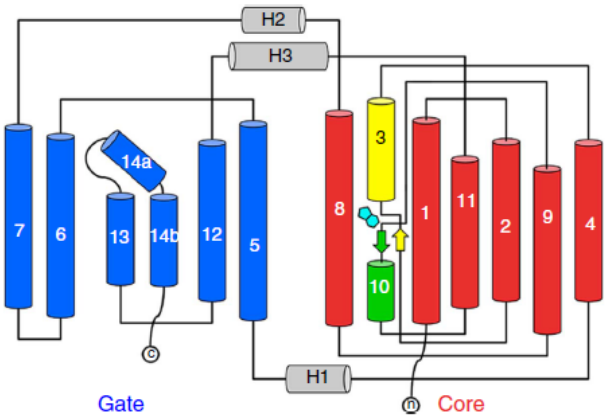


Figure 1.5. Topology diagram of UapA (7+7 fold). TMSs of the core domain are colored red and TMS of the dimerization/gate domain blue. The half helices and short β-strand regions of TMS 3 and 10 are colored yellow and green, respectively (both are part of the core domain). A schematic of xanthine is also included colored cyan. This figure is adapted from [2].

1.1.2.3.3 Anion Exchanger (AE) Family

Members of the AE family are present in bacteria, fungi, plants and animals. Characterized members perform mainly four different transport reactions that include bicarbonate: HCO_3^- /Cl^- exchange, Na^+/HCO_3^- cotransport, Na^+-HCO_3/H^+-Cl^- exchange and boric acid (boron) efflux ([38] and references therein). In certain cell types have also been detected Na^+ driven Cl^- /HCO_3^- exchange and K^+/HCO_3^- exchange activities. Bacterial members are still uncharacterized. Plants and yeast contain anion transporters. Humans code for 10 AEs that perform sodium independent Cl-HCO_3^- exchange (e.g. AEs 1-3), Na^+-coupled anion exchange (e.g. NDCBE), and Na^+/HCO3$^-$ cotransport (e.g. NBCn1, NBCe1 and NBCe2). Animal AE proteins consist of 900-1250 residues that form homodimeric complexes and interact via their N-terminal domains with cytoskeletal proteins. Their structure is made up of 14 TMSs with cytoplasmic relatively long N-terminal and short C-terminal regions. Although the C-terminal is small it contains binding motifs important for trafficking, protein-protein interactions etc. The crystal structure of the C-terminal anion exchanger domain (CTD) of AE1 (also known as band 3 or SLC4A1) has been published [45]. AE1 is a glycoprotein built from two domains, a cytosolic N-terminal domain that functions as an anchoring point for other proteins and an

integral membrane domain that catalyzes anion exchange. As mentioned before, the $AE1_{CTD}$ adopts a 7+7 fold and its structure is very similar to the structure of the NAT members, although they have very small sequence identity. Similarly to NATs, $AE1_{CTD}$ dimerizes, but dimerization seems not necessary for transport ([45] and references therein).

1.1.3 Mechanisms of transporter function

All transporters seem to alternate between two distinct conformations, an outward- and an inward-facing conformation (as suggested by the so-called *alternating access model*; [20,33,46] and references therein). Substrate accessibility is achieved by the opening of the major substrate binding site to either the extracellular or the intracellular space (outward- or inward facing respectively), via significant rearrangement of the transmembrane body of transporters. In addition, besides the gross transition from the outward- to the inward-facing conformation, local conformational changes occur that control the closing and opening of extracellular and intracellular gates. The existence of gates is an additional control element in transport catalysis and avoidance of substrate leakage in the wrong direction that also contributes to the specificity determination. Current evidence supports the idea that gating closure may be ahead of the alteration from the outward- to the inward-facing conformation [14]. Thus, 5 transporter states have been recognized; outward-open, outward-occluded (i.e. outer gate closed), fully-occluded (outer and inner gates closed), inward-occluded (inner gate closed), inward-open. Based on structures and biochemical data, three major transport mechanisms have been described, the gated rocker-switch, the rocking bundle, and the elevator mechanism.

1.1.3.1 Rocker-switch mechanism

The *rocker-switch* mechanism describes two structurally symmetric domains moving around a binding site that locates at the interface, approximately halfway across the membrane (Figure 1.6a) [47,48]. The two structurally similar bundles form a characteristic V-shaped architecture so that a barrier is formed between them, present either at the one side or the other. During transport this barrier moves from the one side and re-forms on the other (moving barrier mechanism) [33,49,50]. Thus, moving barrier transporters work with two separate “thick” gates. The binding site is at first accessible from the one side, a

conformational change closes the path (gate closure) and a new path opens on the other side (gate opening). In the simplest version of this mechanism the substrate binds symmetrically (e.g. SWEETs transporters) [51–53] but substrate binding can also be asymmetric to varying degrees (e.g. MFS transporters) [48,54]. Accordingly, in SWEET transporters the occluded conformation is mainly symmetric forming an O-like shape while in most occluded MFS structures, with the exceptions of EmrD and OxlT, cavities still have a clear orientation ([33] and references therein). In these cases the states are typically referred to as outward-occluded or inward-occluded conformations. Additionally, substrate binding induces the local movement of the gating helices located in one or both of the two domains that occlude the substrate from the outside or the inside achieving alternating access. These helices in MFS transporters are often discontinuous i.e. containing an unwound region [55,56]. Finally, it has been proposed for some MFS transporters that the breakage and re-formation of salt bridges controls transition between the two states [48,57,58]. Substrate binding seems to induce direct or indirect movements of helices involved in a salt bridge network resulting in salt-bridge breakage.

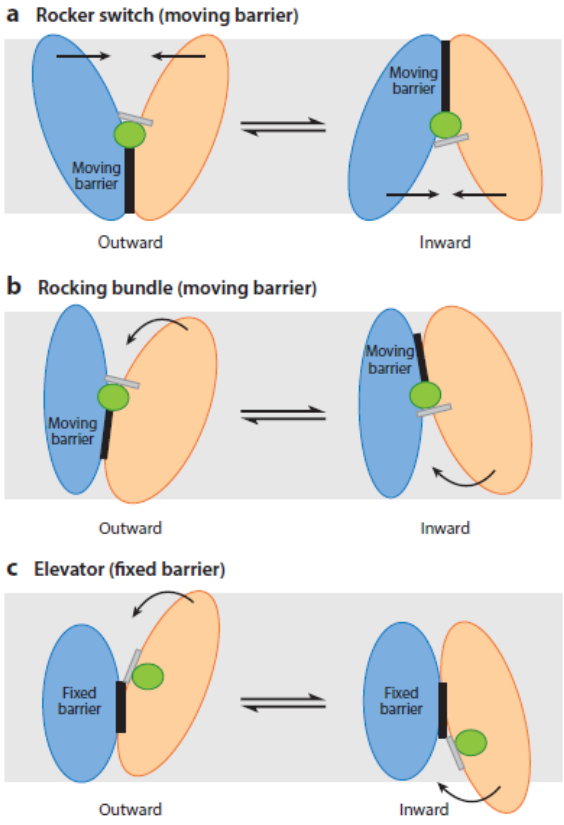


Figure 1.6. Mechanisms of alternating access. Schematic representation of the (A) rocker switch mechanism, (B) rocking bundle and (C) elevator-like mechanism. The substrate is presented with a green sphere and gates are depicted as a thick line over the substrate. In (B) and (C) the core domain is colored pink while the scaffold (in the rocking-bundle mechanism) or dimerization/gate (in the elevator mechanism) domain blue. This figure is adapted from [33].

1.1.3.2 Rocking-bundle mechanism

The prototypical example of a transporter moving with the *rocking-bundle* mechanism is LeuT [59]. In contrast to the rocking-switch, transporters with this mechanism contain two domains that are structurally dissimilar (referred as *scaffold* and *core* domains), with one domain predominantly rocking around the other [60]. In both cases the substrate binding site is centrally located, a moving barrier mechanism is used and substrate binding is associated with local rearrangements of gating elements, often formed by discontinuous helices (Figure 1.6b). However, the gates in transporters with the rocking-bundle mechanism are more fine and gating rearrangements vary substantially among them [33]. Additionally, ionic interactions seem to also be important for LeuT function as an important salt bridge that is conserved is formed during gating [61,62]. Finally, rocking bundle proteins include more complex transport pathways, probably as a result of substrate transport coupling with the movement of a driving ion, whereas most facilitators (passive transport) use the apparently "simpler" rocker-switch mechanism [33].

1.1.3.3 Elevator mechanism

The third major mechanism, the *elevator mechanism* is performed by transporters made up of two distinct domains, similar to transporters functioning with the rocking bundle mechanism [33]. The defining feature of this mechanism is that the substrate binding site locates exclusively in one of the two domains and more specifically in the core domain (or *elevator* domain), which slides through the lipid bilayer supported by the dimerization/gate domain (referred also as *scaffold* domain; Figure 1.6c). Thus, the substrate is carried across the membrane by only one domain, in contrast to the previously described mechanisms. Due to this vertical translocation the mechanism has been described also as the *moving carrier* mechanism [63]. The term *fixed barrier* was also used but is not totally representative as some elevator proteins may not have a fixed barrier (see also later). The dimerization/scaffold domain is rigid and immobile, but seems to provide most of the gating elements and mediates oligomerization, as transporters with this mechanism are usually dimers or trimmers [33]. The elevator mechanism can be classified into three types based on the gating mechanism: the fixed-barrier with one gate, the fixed barrier with two gates and the moving-barrier with two gates [63]. Classification is possible only for proteins that their structure is determined in all different conformational states, thus the mechanism of many elevator proteins has not been experimentally confirmed.

The human neutral amino acid exchanger ASCT2 that belongs to the SLC1 family uses the fixed barrier with a one-gate mechanism [64]. The transport domain of this transporter contains two structural elements called helical hairpins (HPs). The HP2 changes positions in the transition between the inward- and the outward-facing conformations and works both as an extracellular and an intracellular gate [64,65]. This mechanism is probably conserved in the SLC1 family [66,67] and likely in also other transporters such as the members of the Phosphotransferase System (PTS) [68–70]. The scaffold domain includes also two highly tilted helices that determine the minimal distance that the substrate-binding site moves, and have been named as the fixed barrier [63].

Transporters using the fixed-barrier with a two-gate elevator mechanism use different gates on the extra- and intra-cellular sides [63]. The nucleoside transporter CNT of the SLC8 family is the only elevator transporter with published structures in multiple intermediate conformations [71]. CNT uses this mechanism using a half-TM as an extracellular gate and an HP as an intracellular gate. Two other TMSs of the scaffold domain form the fixed barrier. The location of the binding site in transporters using this mechanism is at the interface between the two domains but the substrate makes more interactions with the core domain. It has been proposed that the location of the binding site is the reason for the existence of two gates as transporters using the first type have an occluded binding site within the core domain [33]. Most elevator-transporters conserve a binding site at the interface of the two domains suggesting that they may also use this mechanism. Examples include transporters of the AbgT family [72], anion exchanger 1 (AE1) [45], and UapA and UraA from the NAT family [2,6].

The last type, the moving-barrier with a two-gate mechanism has been described for the bile acid transporter ASBT [63,73,74]. This transporter combines sliding of the transport domain relative to the scaffold, thus an elevator movement, with a moving barrier, which is used by non-elevator mechanisms. Additionally, the transport domain of ASBT, but also the sodium-proton antiporters NapA and NhaA, does not contain HPs [75–78]. Finally, it should be noticed that ASBT is one of the elevator few transporters that is monomeric.

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

Several aspects of transporter biology, such as structure-function relationships, substrate specificity, regulation of expression and subcellular trafficking, have been studied in our lab

using the fungus *Aspergillus nidulans* as a model system. The transporters of interest are mostly nucleobase transporters of the NCS1 and the NAT/NSC2 families. Knowledge on how nucleobase transporters work has an obvious scientific interest as the metabolism of nucleobases is implicated in different diseases, but are also used as antimetabolites against several infections and cancer [79,80]. Understanding the structure-function relationships of these transporters will be also important for drug design [81].

1.2.1 The life cycle of *A. nidulans*

*A. nidulans (*also referred to as *Emericella nidulans)* is a homothallic filamentous fungus that is widespread and grows commonly on plants, soil, decaying organic matter, in/on animal systems but also in indoor environments such as damp walls. The term 'homothallic' refers to the ability of an organism, to reproduce sexually and thus having no mating types. The Aspergillus genus belongs to the phylum of Ascomycota and more specifically to the Pezizomycotina subphylum and the class of Eurotiomycetes [82,83]. The other two subphyla, Saccharomycotina and Taphrinomycotina, contain yeast-like taxa. As a member of the Ascomycota phylum, *A. nidulans* produces saclike structures named "ascus" that are formed as a result of its sexual cycle (see also later). Aspergillus species have both medical and commercial importance as some species can cause infections in humans or animals (e.g. *A. fumigatus*) while others are used for the production of natural products or enzymes (e.g. *A. niger* is the major source of citric acid).

A. nidulans can reproduce through three reproductive sub-cycles, the asexual, the sexual and the parasexual (Figure 1.7) [84]. During the asexual cycle, it grows by apical extension of multinucleate cells called hyphae that, under the appropriate conditions, form asexual reproductive structures called conidiophores (Figure 1.7 blue colored part). These structures are formed after differentiation of some hyphal cells into foot cells that create long stalks with a vesicle at their end. The vehicle gives rise through budding to a layer of 50-70 sterigmata called metulae. These in turn bud and form another layer of sterigmata called phialides that produce long chains of mitotically derived uninucleate spores called conidiospores (asexual spores; they confer *A. nidulans* green colour). Conidiospores are dispersed from the conidiophores and under favourable conditions germinate and produce haploid vegetative filamentous hyphae.

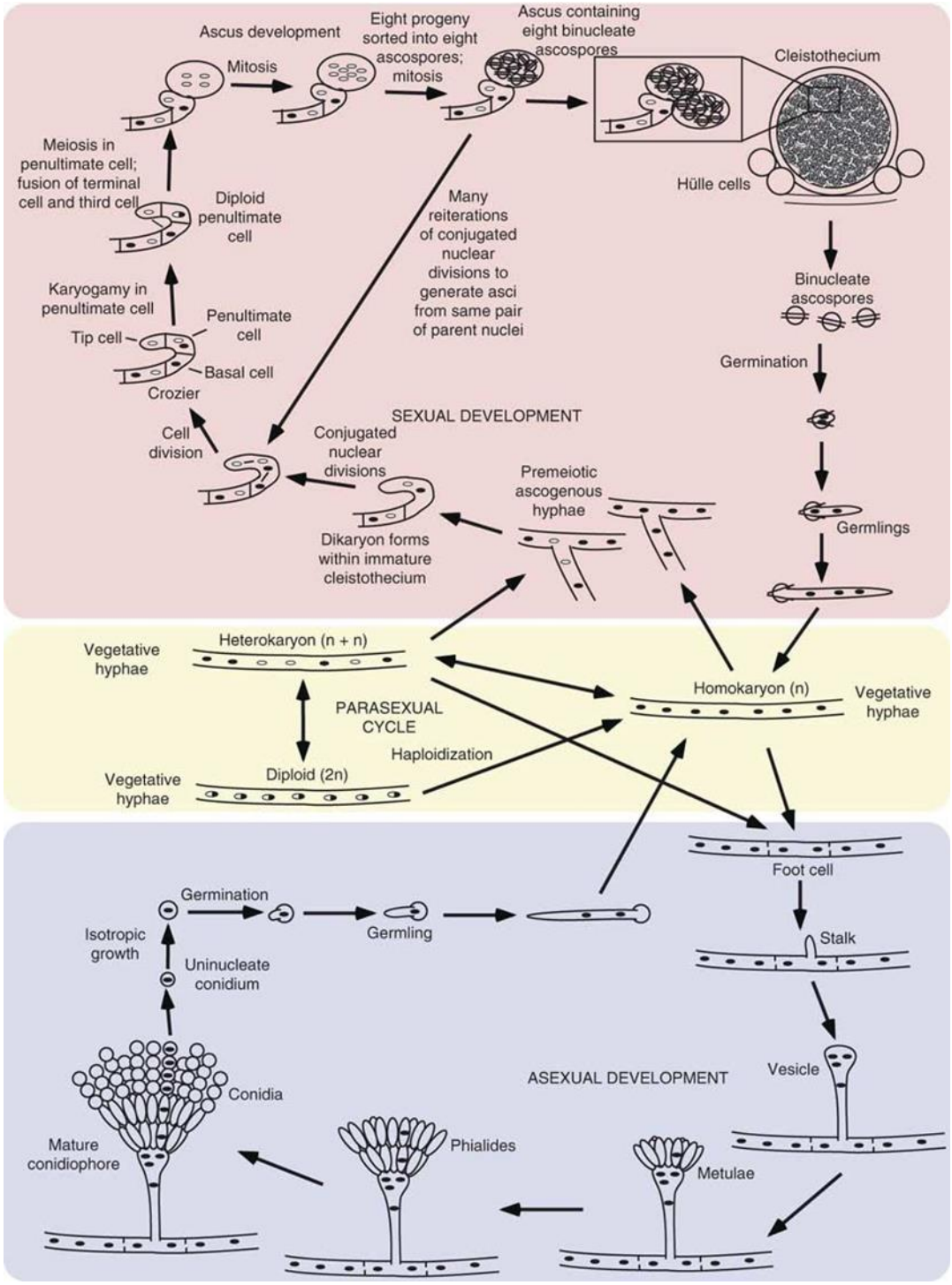


Figure 1.7. Diagrammatic representation of *A. nidulans* life cycle. Three different cycles exist: Sexual (colored pink), asexual cycle (colored blue) and parasexual cycle (colored yellow). The figure is from [84].

A hypha can enter the sexual cycle when conditions for vegetative growth are restraining and sexual spores, called ascospores, are generally recruited for long-term survival (Figure 1.7 pink colored part) [84]. During the sexual cycle two hyphae of the same (*selfing*) or two distinct individuals (heterothallic) fuse and form a heterokaryon that contains nuclei from both parental hyphae. The end of an ascogenous heterokaryon contains a uninucleate tip cell,

a binucleate penultimate cell and a uninucleate basal cell. The penultimate cell differentiates and increases in size to become an ascus and its two nuclei are fused creating a diploid cell. Then due to meiosis in the penultimate cell four nuclei are produced and the tip cell fuses with the basal cell to generate a binucleate cell. The four nuclei divide mitotically to generate eight nuclei that are sorted in eight ascospores and each nuclei undergoes again mitosis resulting in eight mature binucleate ascospores (in an ascus). The binucleate cell that was created by the tip and the basal cell repeats this process and creates a second ascus. All the asci are dispersed in sexual reproductive structures called cleistothecia that are surrounded by thick-walled nursing cells called Hülle cells [85]. These cells provide nutrients for the developing cleistothecia that mature and contain more than 10,000 binucleate ascospores, generated by a single ascogenous hypha in the ascus. Ascospores can germinate and enter the vegetative growth. Heterokarya can instead of entering the sexual cycle transform to a diploid or homokaryon (parasexual cycle; Figure 1.7 yellow colored part) [84]. If a heterokaryon is formed there is a 10^{-6} probability of nuclei fusion and creation of stable heterozygous diploid nuclei that generate diploid mycelia with the same architecture as the haploid ones. Diploids can revert to haploids involving mitotic non-disjunctions which results in random chromosome reassortments and production of aneuploid and haploid cells. Mitotic recombination is also possible leading in exchange of genes on chromosomes [86].

1.2.2 Dissection of structure function relationships of transporters in *A. nidulans*

At least 10-12% of the genome of filamentous fungi codes for transporters. Aspergilli seem to possess more than 120 different families of transporters with some of them appearing only within the genus, others being fungus-specific and others being present in all domains of life [7]. *A. nidulans* encodes twelve transporters of the NSC1 family, from both Fcy and Fur subfamilies, two members of the canonical NAT/NCS2 family, UapA and UapC, and a single AzgA-like purine transporter.

Model fungi, such as *A. nidulans,* provide unique classical and reverse genetic tools for the study of many aspects of transporter biology, including the study of structure-function relationships. In addition, a major advantage of using specifically *A. nidulans* for studying purine and pyridine transporters is that purines can be catabolized and used as sole nitrogen sources, while several purine or pyrimidine analogues, incorporated in the cell specifically by nucleobase transporters, lead to severe cytotoxicity. Thus, mutations in nucleobase transporters are reflected in inability for growth on purines and/or resistance to cytotoxic

nucleobase analogues, making growth tests on different nucleobases a quick and easy way to test transporter function and specificity. Additionally these features allow the use of direct and unbiased genetic screens for selecting mutations that confer altered specificity to a transporter or restore function of inactive mutants [7]. These screens can reveal the importance of residues or interactions between residues that could have not been predicted *a priori*.

A. nidulans also provides easy and rapid functional assays of transport activities by measuring the kinetics of uptake of radiolabeled nucleobases in easily manipulated germinated conidiospores [87]. Similar to enzymes, most transporters have a single substrate binding-site and thus their activity follows Michaelis-Menten kinetics. These assays are used to measure the apparent K_m value, also referred as the Michaelis-Menten constant, that is defined as the substrate concentration that gives half-maximal velocity of an enzymatic/transport reaction and provides information on the affinity by which an enzyme or transporter binds a substrate. The apparent *Vm* value, that is the concentration of substrate accumulated within a specific number/mass of cells in a specific period of time, can also be measured. Finally, *A. nidulans* permits the study of subcellular expression of any transporter functionally tagged with GFP, or other fluorescent epitopes (usually at their C-terminus), in living cells growing under specific, well-defined and fully controllable physiological or stress conditions [7]. Importantly, NATs conserve wild-type kinetics and specificity when tagged with GFP while NCS1 transporters are differentially sensitive to GFP tagging. Thus epifluorescence *in vivo* microscopy can be used to distinguish mutations that affect subcellular trafficking, stability and turnover, from those affecting transport activity *per se*. In recent years, the unique genetic tools of *A. nidulans* have been combined with the tools of uptake assays, GFP-tagging, *in silico* modelling and crystallography, specifically for the study of nucleobase transporters.

1.3 Structure-function relationships in the NAT/NCS2 family

This thesis was focused on the study of transporters of the NAT family, and more specifically, on UapA. As mentioned before the NAT family consists of secondary transporters that are made of 414-650 residues and are highly specific for either nucleobases/H^+ or L-ascorbate/Na^+. Bacteria, fungi, plants and mammals possess nucleobase-specific members while only mammals contain L-ascorbate specific members (SVCT1/2). Mammals also code for

SVCT3 homologues with unknown function. More than 24 members have been functionally characterized via direct assessment of their transport activities or indirectly via genetic and physiological studies (Table 1.1), but the most extensively studied members come from *E. coli* and *A. nidulans*.

Table 1. Characteristics of biochemically characterized NATs

Transporter name	Origin	Physiological substrate	Other substrates	Cation symport	Known structure	Q-H motif (TMS1) Substrate transport	ExxGD motif (TMS8) Substrate bind & transport	NAT *signature* motif (TMS10) Substrate bind & transport & *specificity*
UraA	*E. coli*	**uracil**	5-FU	H^+	yes/dimer bound to uracil	yes	yes (ExxG**H**)	T-**Y**-G-**E**-NxGxxxxTG
XanQ (YgfO)	*E. coli*	**xanthine**	xanthine analogues[#]	H^+	*	yes (**T**-H)	yes	T-**F**-A-**Q**-NxGxxxxTG
XanP (YicE)	*E. coli*	**Xanthine** *low-capacity*	?	H^+	*	yes	yes	C-**F**-G-**Q**-NxGxxxxTG
UacT (YgfU)	*E. coli*	**uric acid** *low-affinity*	xanthine (very low capacity)	H^+	*	yes	yes/no ExxG**M**	S-**F**-S-**Q**-NxGxxxxTG
UapA	*A.nidulans*	**uric acid-xanthine**	uracil, xanthine-uric acid analogues oxypurinol, allopurinol	H^+	yes/ dimer bound to xanthine	yes	yes	T-**F**-A-**Q**-NxGxxxxTR
UapC	*A.nidulans* *A.fumigatus*	**xanthine – uric acid**	uracil, xanthine-uric acid analogues oxypurinol	H^+	*	yes	yes	V-**F**-A-**Q**-NxGxxxxTR/K
Xut1	*C. albicans*	**xanthine – uric acid**	uracil, xanthine-uric acid analogues oxypurinol	H^+	*	yes	yes	V-**F**-A-**Q**-NxGxxxxTK
Lpe1	*Zea mays*	**uric acid-xanthine**		H^+	*	yes	yes	A-**S**-V-**E**-NxGxxxxTG
Nat3	*A. thaliana*	**adenine, hypoxanthine guanine, uracil**	?	H^+	*	yes	yes	A-**S**-V-**E**-NxGxxxxTG
Nat12	*A. thaliana*	**adenine, hypoxanthine guanine, uracil**	?	H^+	*	yes	yes	T-**L**-T-**E**-NxGxxxxTG
rSNBT1 (SVCT4)	Rat, mouse *predominantly in small intestine* *pseudogene in primates*	**hypoxanthine, xanthine guanine, uracil, thymine**	5-FU, oxypurinol	Na^+	*	yes	yes	S-**Y**-S-**E**-NxGxxxxTR
SVCT1	Rat, mouse, human *Epithelia kidney, intestines, liver, lung, skin*	**L-ascorbate**	-	2 Na^+ (need of Ca^{2+} or Mg^{2+})	*	yes	yes	S-**S**-S-**P**-NxGxxxxTK
SVCT2	Rat, mouse, human *Widespread, including Brain*	**L-ascorbate**	-	2 Na^+ (need of Ca^{2+} or Mg^{2+})	*	yes	yes	S-**S**-S-**P**-NxGxxxxTK
SVCT3	Vertebrates *predominantly in kidney*	?	?	Na^+	*	yes	**no** (SxxGC)	S-**S**-F-**P**-NxGxxxxTG

1.3.1 Highly conserved motifs essential for transport function

Previous studies have established the existence of a small number of amino acid residues that are critical for NAT function. Most of these residues fall within three motifs that are highly

conserved in all NATs, in TMS1, TMS8 and TMS10. The first motif in TMS1 is a two-amino acid residue motif and is almost absolutely conserved in all NATs as Gln-His (only few homologues have other versions like Thr-His or Gln-Gln). In the xanthine permease XanQ from *E. coli*, His^{31} in this motif is crucial for affinity and specificity and the presence of an uncharged residue capable of making hydrogen bonds at this position is essential for high affinity binding and transport [88]. Based on homology modelling this crucial hydrogen interaction seems to be with Asn^{325}, whose role will be discussed later. A network of hydrogen bonds between TMS1, TMS3 and TMS10 including the corresponding structurally analogous residues (His^{86} and Asn^{409}) seems to be crucial also for function and stability of UapA [11,89]. Additionally a similarly positioned His residue (His^{51}) affected affinity and specificity of SVCT1 and is important for SVCT2 transport activity [90]. The Gln residue of this motif is also crucial for transport catalysis in both UapA (Gln^{85}) and SVCT1 (Gln^{50}), being probably part of this hydrogen bond network [11,89,90]. Contrary to the above observations, the corresponding residue in XanQ, which possess a Thr at this position (Thr^{30}), does not seem to have a similar role [88].

The motif in TMS8 conforms to the sequence **Glu**-X-X-**Gly**-**Asp**/His/Gly where with bold are the most prominently conserved residues. The first residue of this motif is almost totally conserved in all NATs as a Glu. In UapA this residue (Glu^{356}) is essential for substrate binding and transport and similarly, in XanQ Glu^{272} is irreplaceable in respect to transport [10,88]. In the crystals of both UapA and UraA this residue is one of the residues binding xanthine or uracil, respectively [2,3,6]. The fifth residue of the motif is necessary for xanthine transport in XanQ (Asp^{276}), while mutations to His or Ala in UapA (Asp^{360}) significantly reduce transport activity [11,91]. Several data on UapA supported the involvement of this residue in binding and symport of H^+ and an indirect role in substrate translocation, possibly through an interaction network with Asn^{410} and Thr^{405} that strengthens TMS8-TMS10 association [11]. In UraA, it corresponds to a His residue (His^{245}) that has been proposed to be important for a proton-coupled mechanism of uracil transport [3,6]. The fourth residue of the motif that is almost totally conserved as a Gly has been studied only in XanQ (Gly^{275}), where Cys replacement does not affect transport activity, but confers high sensitivity to inactivation by site-directed alkylation [91]. This means that this residue either plays a role at the periphery of the binding site conferring conformational flexibility, or allows substrate release from the binding pocket. Overall, the **Glu**-X-X-**Gly**-**Asp**/His/Gly motif seems to form an α-helical face at the cytoplasmic end of TM8 that is critical for transport in all studied NATs.

Finally, in TMS10 there is a longer motif that conforms to the sequence Gln/Glu/Pro-**Asn**-X-**Gly**-X-X-X-X-**Thr**-Gly/Arg/Lys, and which is historically referred to as the NAT signature motif. Direct structural evidence has shown that the first residue in this motif interacts strongly with substrates, in both the UapA and UraA crystal structures. In UapA this residue is a Gln (Gln^{408}), whereas in UraA is a Glu (Glu^{290}) [2,3,6]. Mutagenesis in UapA, UraA, but also XanQ (Gln^{324}), confirmed that this residue is critical for substrate binding and transport catalysis [6,8,92]. The well-conserved Asn at the second position is also irreplaceable for UapA (Asn^{409}) and XanQ (Asn^{325}) transport catalysis, but is not involved in substrate binding. As mentioned before, this residue seems to be involved in a network of critical intramolecular interactions in UapA and XanQ [11] . However, in UraA where an interaction of Asn^{291} with TMS1 is also present, the effect of substitutions of this residue are not as dramatic as in UapA [6]. The third residue of the motif, that is variably conserved, is Asn^{410} in UapA and as mentioned previously seems to participate together with Asp^{360} in the interaction network between TMS8 and TMS10. The same residue in XanQ (Asn^{326}) is not important for transport function, but the single-Cys substitution of this residue is highly sensitive to inactivation by site-directed alkylation. The fourth residue of the motif is highly conserved among NATs as a Gly. In UapA (Gly^{411}) and XanQ (Gly^{327}) substitutions of this residue affect both transport kinetics (K_m, V) and specificity, whereas in XanQ Cys substitution is also sensitive to inactivation by site-directed alkylation [8,92,93]. Importantly, the crystallized version of UapA contained the G411V mutation, which does not prevent xanthine binding, but still scores as a loss of function mutant [2,8]. This mutation does not seem to affect the local protein conformation (comparing TMS10 in the UapA crystal with UraA), but probably hinders the core domain from sliding on the dimerization domain (see later). Regarding the two last residues of the motif, Thr is highly conserved, while the last residue is variably conserved as Gly, Arg or Lys. Substitutions of these residues in both UapA (Thr^{416} and Arg^{417}) and XanQ (Thr^{332} and Gly^{333}) affected kinetics or specificity to a varying extend [8,92]. Importantly residue Arg^{417} has been shown to be important specifically for increasing uric acid binding affinity in UapA [94].

Three additional amino acid residues located upstream of the NAT signature motif are also conserved among NATs (Thr/Ser-Phe/Tyr/Ser-Ala/Ser) and can be included in an extended version of the motif. In the UapA crystal two of them, Phe^{406} and Ala^{407}, interact directly with the substrate via π-π stacking and hydrogen bonding (with Ala^{407} main chain), respectively [2]. F406Y was selected as a UapA substitution conferring growth on low concentration of adenine [9]. In a wild-type UapA context, F406Y did not significantly affect

transport activity for physiological substrates, but improved binding of other purines. In XanQ replacement of this residue to Cys (Phe^{322}) did not affect transport function [92]. The crystal structure and mutagenesis of UraA showed similar results concerning the analogous residues (Tyr^{288} and Gly^{289}) [3,6]. Ala^{407} was only studied in UapA in a A407S/Q408P mutant that mimics residues found in L-ascorbate NATs and which has totally lost its capacity for purine uptake [11]. In XanQ, Ala^{323} was found to tolerate replacement with Cys, but showed strong inhibition by site-directed alkylation [92]. Finally, the first residue has only been studied in XanQ, where replacement by a Cys residue (T321C) decreased transport activity [92].

1.3.2 Residues affecting UapA specificity-Gating elements

Previous studies on UapA revealed a number of specificity mutations concerning partially or non-conserved residues that map outside the major substrate binding site. These residues are located either at putative hinge regions distant from the substrate binding site between the two domains or at the interface between the dimerization and the core domain near the substrate-binding site (Figure 1.8) [2,9,10]. Additionally, mutations in residue Arg481 arise often in genetic screens selecting for UapA mutants capable of importing other purines. This residue is even further distant from the major substrate binding site. These substitutions do not affect the apparent K_m, but increase 1.5-2.0-fold the apparent rate of transport of physiological substrates and enlarge UapA specificity by allowing low-affinity transport of several nucleobases, such as hypoxanthine and/or adenine [9,10].

Gln^{113} (TMS2) and Ala^{441} (in the helix 3 between TMS11 and TMS12; Figure 1.8) are located in flexible helical hinges 'distant' from the binding site. Q113L was isolated as a substitution conferring UapA-mediated growth on 2 mM adenine and as a second site suppressor allowing growth of UapA-F528S (its role will be discussed later) on 0.5 mM adenine [9,10]. Q113E did not have the same effect as Q113L, while Q113K scored as a loss of function mutation due to defective plasma membrane localization [10]. Ala^{441} is very close structurally to Gln^{113}, and was also isolated as a suppressor of UapA-F528S on 0.5 mM adenine [2,9]. A441V alone was able to confer growth on 0.5 mM adenine and hypoxanthine as Q113L, and both mutations significantly increased affinity for hypoxanthine and uracil, but not for adenine [9,10]. Subsequently, it has been shown that growth on adenine in these mutants is due to increased transport capacity, rather than substrate binding affinity. Gln^{113} is not conserved, while Ala^{441} is relatively well conserved, although in some NATs is replaced by other residues (Gly, Ser, Thr or Pro).

Substitutions of residues Val^{463} and Ala^{469} in TMS12 were also selected in the genetic screen using UapA-F528S, mentioned previously. V463I and A469E alone affected UapA transport specificity similarly to Gln^{113} and Ala^{441}, as they all conferred growth on high concentrations of adenine and hypoxanthine (2 mM). Overall, these mutants acquired low (adenine and hypoxanthine) or moderate (uracil) affinity for non-native substrates, without affecting significantly the affinity for physiological substrates, although A469E decreased the affinity specifically for uric acid by 8.5-fold. Noticeably, other NAT homologues with no (e.g. XanQ, XanP and PbuX) or moderate (e.g. UapC and Xut1) affinity for uric acid, do not possess an Ala residue at the analogous position of Ala^{469}. Residue Val^{463} of UapA is conserved as a Val or an Ile in other NATs, while Ala^{469} is more variable, with other NATs possessing Arg, Ser or Lys. In conclusion, specific substitutions of these residues confer low affinity transport of nucleobases other than uric acid and xanthine in UapA, leading to the hypothesis that residues at one side of TMS12 might be elements of the substrate translocation pathway [11].

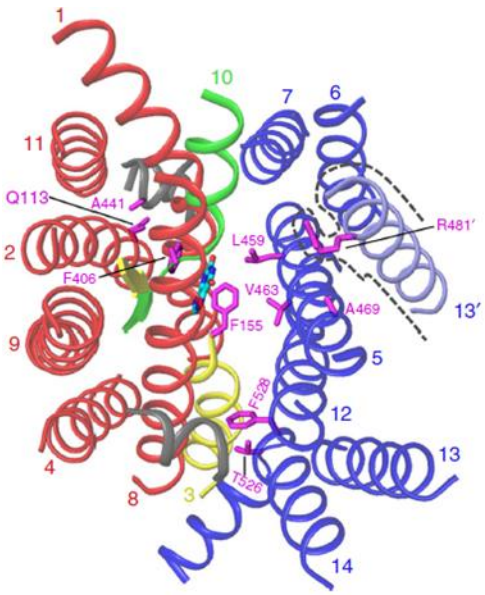


Figure 1.8. Location of residues involved in UapA substrate specificity. Helical regions of the core and dimerization domain are colored red and blue respectively. The half helices and short β-strand regions of TMS 3 and 10 are colored yellow and green, respectively. Xanthine is shown in cyan stick model.TM 13 from the opposite monomer is shown in light blue and its surface is indicated by the black dotted line. Residues involved in substrate specificity are shown in magenta stick model and labelled. The figure is adapted from [2]

Substitutions of residues Thr^{526} and Phe^{528}, located in TMS14 (scaffold/dimerization domain), also enlarge UapA specificity in respect to nucleobases. Mutation F528S was originally selected as a suppressor mutation of the cryo-sensitive UapA-Q408E mutant suggesting that this residue interacts functionally with the NAT signature motif [95]. Thr^{526} substitutions (T526M, T526L) appeared in the genetic screen for UapA mutants allowing growth on 2 mM adenine, together with mutations in Gln^{113}. Substitutions of these two residues conferred UapA-mediated growth on high concentration of adenine or hypoxanthine (2 mM), without significantly increasing the relative affinities for these non-native substrates

[9,10]. Based on these findings and their topology, it has been proposed that these two residues act as outward-facing gating elements that select which purines can have access to the major substrate binding site of UapA, which in turn could explain why specific substitutions of these residues might loosen substrate selectivity [9,11]. The existence of selective gating elements in transporters was later supported by structural studies in other transporters ([7] and references therein).

Phe^{528} and Thr^{526} are only partially conserved, and noticeably also, TMS14 shows significant structural differences in different NATs [11,96]. Substitutions of the corresponding residues in XanQ (Asn^{430} and Ile^{432}) suggested that they contribute to the determination of the transport kinetics of xanthine and only partially to specificity allowing transport of some xanthine analogs with bulky groups. [96]. Also in XanQ, protection from alkylation in the presence of 2-thioxanthine in the N430C mutant, together with the position of this residue distal from the substrate binding site (based on the structural model), suggested that TMS14 possibly has a conformational interaction with TMS10 or other TMSs of the core domain [97]. These observations suggested that in XanQ residues Asn^{430} and Ile^{432} function as a substrate filter rather than a selectivity gate [14].

Finally, specificity mutations via substitutions of Arg^{481}, located in TMS13, appeared with a great frequency, also in the genetic screen using UapA-F528S [9]. Originally, three different substitutions were selected: R481G, R481L and R481Q. These substitutions alone conferred growth on 0.5 mM adenine and hypoxanthine and prominently on 2 mM, without increasing the relative binding affinities. This Arg-residue is conserved only in fungal NATs, whereas bacterial, plant and animal homologues, have Leu, Ile, Thr or Lys residues. Two observations suggested that Arg^{481} might act as an inward-facing gating element. Firstly, deletion of Arg^{481} or relocation of its position by the insertion of two amino acids (Ala-Gly) before it, resulted in UapA instability and increased turnover [9]. Most importantly, combination of Arg^{481} substitutions with substitutions in the proposed outward-facing gating elements (e.g. F528S or T526M) resulted in an additive effect in enlarging specificity. At that time, Molecular Dynamics and the topology of Arg^{481} relative to the substrate binding site, could not explain how Arg^{481} might function as a gating element in the UapA modeled structure. However, in 2016 when the crystal structure of UapA became available, the role of this residue was clarified as it was revealed that UapA forms functional dimers and that Arg^{481} is close to the binding site of the opposite protomer (see also later) [2].

1.3.3 Crystal structures

The crystal structures of two members of the NAT family have been published until now, one is the bacterial uracil transporter UraA (two structures in different conformations) and the other is UapA [2,3,6]. Structures of eight structural homologues have been also reported, which belong either to the SLC4 (including AE members) or the SLC26 (including SulP members) family (Table 1.2). The SLC4 members whose structure has been resolved are the borate efflux transporter Bor1 from *Arabidopsis thaliana*, the Bor1p from *Saccharomyces mikatae*, the human anion exchanger 1 Band3 protein and the human Na^+-coupled acid-base transporter NBCe1 [45,98–100]. The members of the SLC26 family are the proton-coupled putative fumarate symporter SLC26Dg from the bacterium *Deinococcus geothermalis*, the bicarbonate transporter BicA from cyanobacteria *Synechocystis sp*., the Slc26a9 protein from *Mus musculus* and the human SLC26A9 protein [101–104]. UraA, UapA, Bor1p, SLC26Dg, BicA, Slc26a9 and SLC26A9 structures are all available in an inward-open conformation, with UraA being crystalized also in an occluded conformation (all except Bor1p, SLC26Dg and Slc26a9 include a substrate). The crystal structures of Band3 and NBCe1 are available in an outward-open conformation with no substrate. As mentioned before all of these proteins are composed of 14 TMSs that are organized into a pair of 7+7 inverted repeats that constitute the core and dimerization/gate domain.

Table 1.2 Structures of NATs and their structural homologues

Name	Organism	Family	Conformation	Crystal ligands	Resolution (Å)	PDB ID
UapA	*Aspergillus nidulans*	SLC23	Inward	Xanthine	3.70	5I6C
UraA	*Escherichia coli*	SLC23	Inward	Uracil	2.78	3QE7
UraA	*Escherichia coli*	SLC23	Occluded	Uracil	2.50	5XLS
Band3	*Homo sapiens*	SLC4	Outward	-	3.50	4YZF
Bor1	*Arabidopsis thaliana*	SLC4	Inward	-	4.11	5L25
Bor1p	*Saccharomyces mikatae*	SLC4	Inward	-	5.90	5SV9
NBCe1	*Homo sapiens*	SLC4	Outward	-	3.90	6CAA
BicA	*Synechocystis sp.*	SLC26	Inward	HCO_3^- + Na^+	2.81	6KI1
SLC26Dg	*Deinococcus geothermalis*	SLC26	Inward	-	3.20	5DA0
SLC26Dg	*Deinococcus geothermalis*	SLC26	Inward	-	4.20	5IOF
Slc26a9	*Mus musculus*	SLC26	Inward +	-	3.96	6RTC
SLC26A9	*Homo sapiens*	SLC26	Inward	Cl^- + Na^+	2.60	7CH1

For Slc26a9 a potentially intermediate state is also available (+)

The crystal structure of UapA confirmed all the assumptions made based on the genetic, biochemical and *in silico* predicted data [2,11]. Crystallization was not successful for

the purified wild-type UapA, which was rather unstable, but for a more stable mutant including mutation G411V and an N-terminal truncation (Δ1-11). This mutant can bind, but cannot transport, xanthine, and is normally targeted to the plasma membrane. The crystal structure provided a structural rationale concerning the function of residues proposed to be involved in substrate binding and transport, along with the predicted orientation of xanthine within the binding site and its specific interactions. Additionally, the topology of several residues affecting specificity could be justified as they localize in the binding site (Phe^{406}, Gln^{408}), close to it (Phe^{155}), along the putative substrate translocation trajectory (Thr^{526}, Phe^{528}, Val^{463} and Ala^{469}, or Arg^{481} facing the substrate trajectory of the opposite protomer), or in hinge regions connecting the core and scaffold domains (Gln^{113}, Ala^{441}). Main conclusions from the crystal structure was the confirmation of the formation of UapA dimers and the evidence supporting that dimerization is required for function. The dimer has an extensive interface of up to 6 000 $Å^2$ involving only TMSs (12, 13 and 14) of the dimerization/gate domain. TMS13 seems very important for dimerization as it is fitted into a cleft formed in the opposite monomer. Co-expression experiments of the endogenous WT UapA with each one of four kinetically distinct UapA mutants corroborated for the importance of dimerization in transport function. A key feature of the NAT structure, as demonstrated by all similar crystal structures, is the existence of a pair of antiparallel β-strands located between TM3 and TM10. More specifically TMS3 and TMS10 helices only extend halfway through the protein. Before these two helices two short β-strands and random coils are placed in the center of the protein that crossover forming part of the substrate binding site. Additionally, the structure explained the functional role of Arg^{481}. As discussed in the previous section, the effect of Arg^{481} could not be understood considering the monomer form but the structure showed that the side chain of Arg^{481} extends towards the substrate-binding site of the other monomer in a 12 Å distance from xanthine. Molecular dynamics suggested a translocation pathway for the movement of xanthine from the binding site to the inside of the cell in which movement of the substrate is closely associated with reorientation of the side chain of Arg^{481} from the opposite monomer. On the pathway to the cytoplasm xanthine forms transient H-bond and π-π stacking interactions with Arg^{481} reaching a minimum 2 Å distance.

The structure of UraA in the inward-open conformation revealed a monomer in the crystal packing, but the structure published later in the occluded conformation revealed a homodimeric organization [3,6]. UraA dimerization is mediated through the dimerization/gate domains like UapA but the buried surface area is lower (~2 400 $Å^2$). Dimerization was further supported by systematic in vitro and *in vivo* characterizations which showed that UraA exists

in an equilibrium between the monomer and dimer form and that dimerization is required for transport activity [3]. UraA contains substantially shorter loop regions than UapA. The occluded conformation also contained a β-hairpin in the connecting segment between TM13 and TMS14 not seen in the previous structure. Additionally the new structure revealed the importance of two water molecules for substrate coordination through mediation of hydrogen bonds. The presence of these water molecules creates a hydrogen bond network involving Glu^{241} and His^{245} (Glu^{356} and Asp^{360} in UapA) that can provide a path for H^+ translocation. Molecular dynamics performed to investigate the role of these residues in H^+ translocation revealed that protonation of Glu^{241} resulted in disruption of the hydrogen bond network that keeps uracil bound, while protonation of His^{245} further stabilizes it, without interacting directly but by stabilizing the position of Glu^{241}.

Molecular dynamics in both NAT members and their structural homologues coupled with the genetic analyses in UapA and XanQ, supported that these transporters function via the elevator mechanism. The structures showed that the substrate is coordinated predominantly by residues in the core domain, which is a characteristic of this mechanism. Comparing the structures of Band 3 (outward conformation) with UapA (inward conformation) the movements of the core domain against the dimerization domain represent an elevator mechanism. Additionally the tight interaction between the two dimerization/gate domains of the two monomers in UapA suggests that it would probably remain relatively static complying also for an elevator mechanism. However the comparison between the UraA occluded structure and UapA inward structure showed local rearrangements in the dimerization/gate domain. Of course these changes could be a result of sequence differences between bacterial and fungal homologs but at least the changes seen in TMS5 and TMS12 are also shown comparing the two UraA structures suggesting that they are part of the mechanism.

1.3.4 Study of mammalian NAT homologs

Most information about the NAT family comes from the bacterial and fungal members. This is mainly due to the developed tools, mentioned previously, in such systems that allow the quick and easy construction and study of mutants. Studies made until now have followed three major approaches, the forward and reverse genetic analysis of UapA, the systematic Cys-scanning and site-directed mutagenesis of XanQ and other bacterial members and mutagenesis of selected residues based on the crystal structure of UraA. Additionally, the role

of conserved His residues in SVCT1 and SVCT2 have been studied [90], but information on other residues is missing.

The mammalian homologues, SVCT1 and SVCT2, are important for vitamin C (L-ascorbic acid) homeostasis which is used in a variety of biochemical processes ([5] and references therein). Most eukaryotes can synthesize enzymatically vitamin C via the gulonolactone oxidase pathway but some exceptions including humans, monkeys and guinea pigs lack the enzyme L-gulono- γ-lactone oxidase which catalyzes the last step. These organisms have evolved other mechanisms in order to retrieve vitamin C such as taking it from their diet. Vitamin C uptake is mediated by either the sodium-dependent transporters of the NAT family (SVCT1/2) or by glucose transporters (GLUTs) that transport its oxidized form dehydroascorbic acid (DHA) [5]. Knockouts of SVCT1/2 genes in mice resulted in severe health problems or death after birth.

An approach for studying structure-function relationships in transporters of higher organisms is to express them in model organisms, such as *A. nidulans*. Unfortunately long-standing efforts in our lab and many others to functionally express mammalian solute transporters, including hSVCT1/2, rSNBT1 or the NAT homologue from *Drosophila melanogaster* CG6293, in *A. nidulans* or yeast have failed. In all cases, transporters were retained in the fungal ER. This is probably a result of misfolding due to the different membrane environment, such as the presence of ergosterol in fungi instead of cholesterol found in mammalian membranes, or defects in recognition by the COPII packaging machinery for ER exit. Additionally, a study on the human SVCT1 transporter revealed that two residues Asn138 and Asn144 are active consensus glycosylation sites and that glycosylation is essential for functional expression of hSVCT1 [105]. Thus, glycosylation may also be a downstream bottleneck for the functional expression of mammalian transporters in model fungi.

1.4 Transporter-lipid interactions

Membrane transporters fold, traffic and function being in continuous contact with the lipid bilayer. It has long been understood that membrane lipids surrounding transporters and other proteins are not just a two-dimensional fluid in which membrane proteins freely diffuse. However the study of the functional and structural dependence of transporters on specific membrane lipids has recently begun to be explored in detail. For example, the structure of the betaine transporter BetP, which adopts the LeuT-fold, revealed the presence of lipids

specifically bound to residues important in both transport and regulation and it has been proposed that these lipids are required for monomer stabilization and oligomerization [106].

1.4.1 Membrane lipids

There are four major types of lipid in membranes, glycerophospholipids, glycolipids, sphingolipids and sterols (Figure 1.9) [107]. The glycerophospholipids are the most abundant type and they consist of a glycerol moiety esterified with two fatty acids at the 1- and 2-hydroxy positions and with a phosphate group at the 3-hydroxy position (phosphatidic acid or PA) [108]. The phosphate groups can be esterified with many different molecules creating a wide variety of different lipids. Esterification with choline, ethanolamine, inositol, serine, glycerol and phosphatidylglycerol creates phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylinositol (PI), phosphatidylserine (PS), phosphatidylglycerol (PG) and cardiolipin respectively. The glycerophospholipids are often called as phospholipids (PLs) but there are also sphingolipids with a phosphate group [107,108]. Sphingolipids are made of a sphingosine molecule linked to a fatty acid by an amide bond, and esterified to the primary hydroxyl group with a phosphorylalcohol (sphingomyelin) or a carbohydrate unit (glycolipids) [108]. Chloroplasts contain another type of glycolipids called galactolipids that have a similar structure to the glycerophospholipids, although without a phosphate group at the 3-hydroxy position. They have instead a mono or disaccharide that can also be sulfated. Finally sterols consist of four fused cycloalkane rings with a hydroxyl group at the 3-position of the A-ring (Figure 1.10) [109]. Sterols found in plants are called phytosterols while sterols of animals are called zoosterols [108]. The predominant sterol in higher eukaryotes is cholesterol while fungi produce ergosterol that serves many of cholesterol's functions.

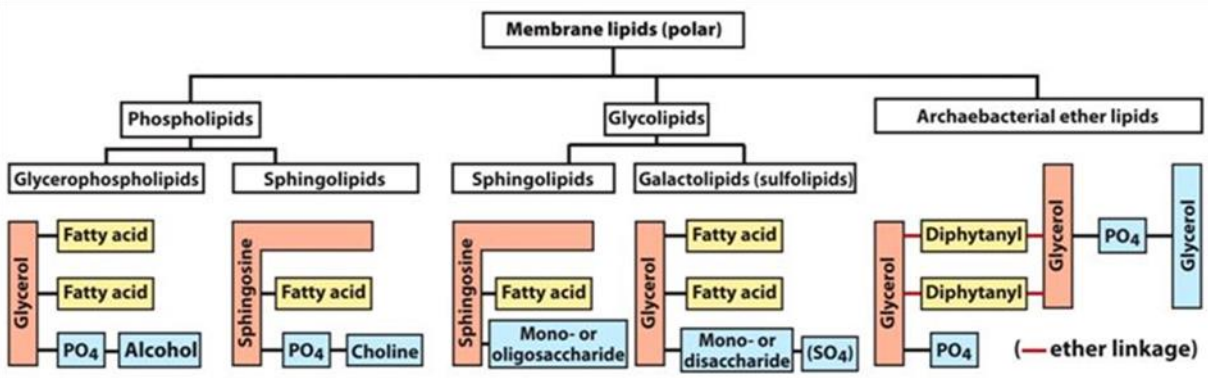


Figure 1.9 Main types of membrane lipids (the picture is adapted from [108]).

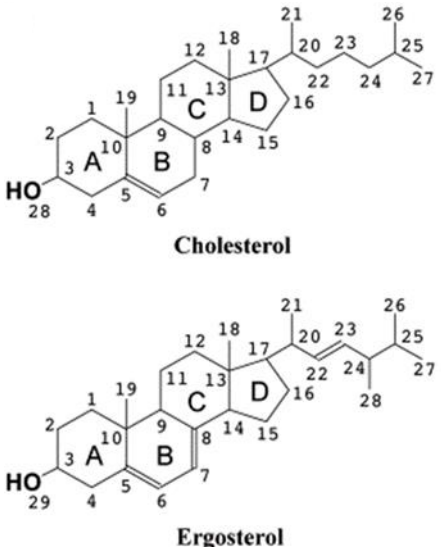


Figure 1.10. Chemical structure of two common sterols with atom numbering (the picture is adapted from [110]).

Lipids can be also categorized based on their position and affinity regarding membrane proteins into three groups: bulk, annular and non-annular lipids [111,112]. Bulk lipids are not in contact with the membrane protein and thus do not interact with it (Figure 1.11) [113]. Annular lipids are in direct contact with the membrane protein mediating between the protein and the rest of the lipid environment [111,112]. Their interactions with proteins are not stable as they exchange very fast with the rest of the lipid environment [114]. On the contrary, non-annular lipids, also referred as structural lipids, bind tightly to membrane proteins so that they often remain bound after purification and crystallization [113]. They are found in protein cavities and between transmembrane helices or protein monomers.

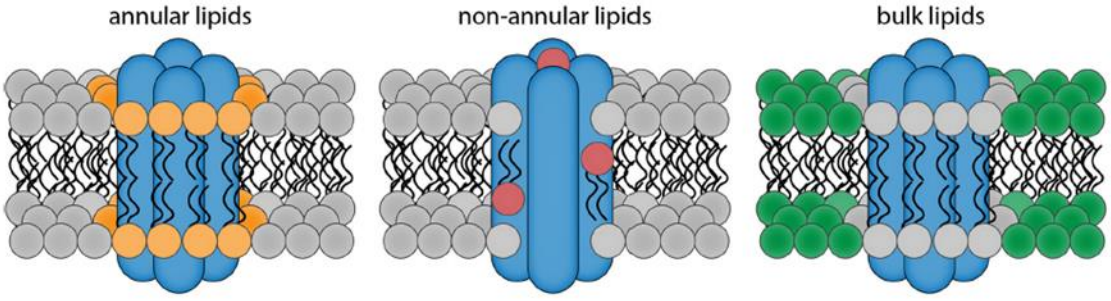


Figure 1.11. Types of lipids based on their affinity to the membrane protein of interest. Annular lipids (orange) form a belt around the protein being in frequent and close contact with it. Non-annular lipids (red) are bound in particular sites on a protein such as protein surface cavities, between transmembrane helices of a protein or at the protein-protein interfaces. Bulk lipids (green) show lower binding affinity for membrane proteins and thus usually do not make contact with them. The figure is adapted from [113].

1.4.2 Functional protein-lipid interactions

The role of protein-lipid interactions has only recently been studied extensively due to technical difficulties resulting from their transient nature. The role of bulk lipids still remains unclear. On the contrary, many studies have shown the importance of non-annular (specific) lipids in both structure and function of many membrane proteins. Proteins may contain specific binding sites for certain lipids [106,115,116]. For example cardiolipin proved to be essential for LeuT dimerization and molecular dynamics predicted binding of cardiolipin and phospholipids at the dimer interface [117]. In this study the authors also correlated the presence of interfacial lipids with the interfacial strength via a bioinformatic analysis of all of the α-helical oligomeric transmembrane proteins with known structures and suggested that lipids are critical for stabilizing weak dimer interfaces. Additionally, resistance to unfolding correlated with specific lipid binding events in the channels MscL from *Mycobacterium tuberculosis* and AmtB from *E. coli* and aquaporin Z AqpZ from *E. coli* [118]. Recent studies also suggested that lipid interactions can also be important for membrane protein stabilization. For example, annular interactions were showed to be critical for stabilization of the elevator-type Na^+/H^+ exchanger NapA from *Thermus thermophilus* [116]. More specifically, it has been shown that NapA did not co-purify with lipids, unlike many other membrane proteins, but addition of increasing amounts of lipids resulted in a concentration-dependent increase in lipid-protein complexes with no apparent saturation, that is indicative of non-specific binding. Lipids seemed to protect the protein structure from destabilization by packing around the dimer domains and the core-dimer interface.

Specific lipid interactions proved also critical for the dimerization of UapA [12]. As mentioned previously, structure analysis and mutagenesis data showed that UapA forms functional dimers [2]. UapA co-purifies with lipids and analysis of the lipid extract showed that these lipids were mostly PE, PI and PC [12]. Delipidation of the purified protein significantly decreased the abundance of the dimer form but addition of either PI or PE or an equimolar mixture of PI and PE restored the UapA dimer. These experiments were performed for a transport inactive, conformationally locked mutant of UapA ($G411V_{\Delta 1\text{-}11}$) that was also crystallized but similar results arose also for wild-type UapA. More specifically the purified wild-type protein showed lower abundance of the dimer form probably due to reduced stability compared to the mutant, but delipidation almost abolished dimerization and this was recovered by PI addition. Moreover MD simulations predicted the existence of a lipid binding site at the dimer interface that is formed by three positively charged residues. These were

Arg^{287} (at the cytoplasmic end of TMS7), Arg^{478}, and Arg^{479} (at the cytoplasmic end of TM13). Alanine substitutions of these residues caused loss of transport function *in vivo* and reduced dimerization suggesting that lipid binding at this site is important for functional dimerization. Finally the MD simulations predicted also a number of other residues involved in lipid binding. These were Lys^{73}, Arg^{133}, Tyr^{137}, Lys^{138}, Lys^{212} and Arg^{421} that locate to the membrane-facing regions of the core domain of the UapA dimer but their potential functional significance has not been examined.

1.5 Goals of the study

As discussed in the previous chapters UapA and some bacterial NATs have been extensively analyzed by a plethora of genetic, biophysical and biochemical studies [2,3,6,7,97]. The determination of the crystal structures of UapA and UraA verified the residues that constitute the major substrate binding site, but provided little insight on the issue of specificity. In NATs, at least two major specificity shifts towards substrates other than nucleobases have occurred; one to L-ascorbate (e.g. SVCT1/2 homologues) and another to unknown substrates (SVCT3 homologues) [5]. The specificity shift from nucleobases to L-ascorbate is a dramatic one as these molecules have different structural and chemical formulas.

The first goal of the present thesis was to investigate the molecular aspects that determine substrate specificity in NATs and address the possible evolution of ascorbate transporters. In particular, the contribution of the NAT signature motif in the evolution of specificity was examined, which is known to be part of the substrate binding site. This was addressed by performing a mutational analysis on UapA introducing NAT motif versions found in other NATs and define their specificity.

In addition, the possible mechanisms through which prominent specificity mutations located distant from the major binding site were addressed [11], for example those concerning residues Arg^{481}, Thr^{526} or Phe^{528}, affect substrate selection and transport.

Based on the recent study on UapA-lipid interactions [12] it was decided to further investigate their role and explore the potential role of other interactions also predicted by MDs. This study showed that lipid binding, probably of PIs and/or PEs at specific residues of UapA at the dimer interface is essential for the formation of functional dimers. In order to further understand the role of these interactions and potentially discover other residues implicated in lipid interactions, a UV mutagenesis was performed in order to select suppressor

mutations that restore UapA-R287A/R478A/R479A growth on uric acid. Additionally, to explore the role of other putative lipid binding residues in the structure and function of UapA, we generated mutants of these residues and explored their potential to affect UapA function, stability and trafficking.

The last goal of the study was the functional heterologous expression of mammalian NAT homologues in *A. nidulans.* This was performed by manipulating a residue that proved critical in UapA stabilization.

Materials and methods

2.1 Strains and growth conditions

2.1.1 Strains used

All strains shown possess the *veA1* mutation which promotes conidiation [119]. *pabaA1, pyroA4, riboB2, argB2, pantoB100* and *pyrG89* are loss of function mutations leading to auxotrophic requirement of para-aminobenzoic acid, pyridoxine, riboflavin, arginine, D-pantothenic acid and uracil/uridine respectively. *AfpyrG* is the *Aspergillus fumigatus* wild type *pyrG* gene (pyrimidine biosynthesis) used as a standard selection marker in *A. nidulans* transformation. *panB* and *riboB* are genes involved in pantothenic acid and riboflavin biosynthesis, used as selection markers for targeted gene knock out. The strains used in this study are included in Tables 2.1-2.4.

Table 2.1 List of strains used in chapter 3.1

Strain	Genotype	References
wt	*pabaA1*	Wild-type reference strain
uapA-	*uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	Pantazopoulou et al., 2007
UapA	*uapA-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	Pantazopoulou et al., 2007
UapA-Q408P	*uapA-GFP-Q408P-argB uapAΔ uapCΔ::AfpyrG azgAΔ argB2 pabaA1*	
UapA-A407S/Q408P	*uapA-GFP-A407S/Q408P-argBuapAΔ uapCΔ::AfpyrG azgAΔ argB2 pabaA1*	Kosti et al., 2012
UapA-F406S/A407S/Q408P	*uapA-GFP-F406S/A407S/Q408P-argB uapAΔ uapCΔ::AfpyrG azgAΔ argB2 pabaA1*	
UapA-A405S/F406S/A407S	*uapA-GFP-T405S/F406S/A407S-argB uapAΔ uapCΔ::AfpyrG azgAΔ argB2 pabaA1*	This study
FurA	*gpdAp-FurA-GFP-panB uapAΔ uapCΔ::pyrG azgAΔ fcyBΔ::argB furDΔ::riboB FurAΔ::riboBΔ cntAΔ::riboB pabaA1*	

Table 2.2 List of strains used in chapter 3.2

Strain	Genotype	References
wt	*pabaA1*	Wild-type reference strain
Δ7	*uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	Krypotou and Diallinas 2014
uapA+	*[gpdAp::uapA-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study

TYAQ	*[gpdAp::uapA-F406Y-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
TFAE	*[gpdAp::uapA-Q408E-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
TYAE	*[gpdAp::uapA-F406Y/Q408E-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
SYSQ	*[gpdAp::uapA-T405S/F406Y/A407S-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
SYSE	*[gpdAp::uapA-T405S/F406Y/A407S/Q408E-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
SSSQ	*[gpdAp::uapA-T405S/F406S/A407S-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
SSSP	*[gpdAp::uapA-T405S/F406S/A407S/Q408P-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
R481G	*[gpdAp::uapA-R481G-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
T526L	*[gpdAp::uapA-T526L-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
F528M	*[gpdAp::uapA-F528M-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
R481G/T526L	*[gpdAp::uapA-R481G/T526L-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
R481G/F528M	*[gpdAp::uapA-R481G/F528M-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
R481G/SYSE	*[gpdAp::uapA-R481G/T405S/F406Y/A407S/Q408E -GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
T526L/SYSE	*[gpdAp::uapA-T526L/T405S/F406Y/A407S/Q408E -GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
F528M/SYSE	*[gpdAp::uapA-F528M/T405S/F406Y/A407S/Q408E -GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study

R481G/T526L/ SYSE	*[gpdAp::uapA-R481G/T526L/T405S/F406Y/A407S/Q408E - GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
R481G/F528M/ SYSE	*[gpdAp::uapA-R481G/F528M/T405S/F406Y/A407S/Q408E - GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
N410V	*[gpdAp::uapA-N410V-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
N410I	*[gpdAp::uapA-N410I-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
N410V/SYSE	*[gpdAp::uapA-N410V/T405S/F406Y/A407S/Q408E-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
N410I/SYSE	*[gpdAp::uapA-N410I/T405S/F406Y/A407S/Q408E-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study

Table 2.3 List of strains used in chapter 3.3

Strain	Genotype	References
UapA-YFPc UapA-YFPn	*alcAp-uapA-YFPc alcAp-uapA-YFPn uapAΔ uapCΔ:: AFpyrG azgAΔ*	Martzoukou et al., 2015
UapA-R287A/R478A/ R479A -YFPc UapA-R287A/R478A/ R479A -YFPn	*alcAp-uapA-R287A/R478A/R479A-YFPc alcAp-uapA-R287A/R478A/R479A-YFPn uapAΔ uapCΔ:: AFpyrG azgAΔ*	Pyle et al., 2018
UapA-GFP or uapA+ or UapA	*uapA-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	Pantazopoulou et al., 2007
UapA-R287A/R478A/ R479A-GFP or RRR	*uapA-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	Pyle et al., 2018
uapA-	*uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	Pantazopoulou et al., 2007
I157F/RRR (in Figure 3.6)	*alcAp-uapA-I157F/R287A/R478A/R479A-YFPc alcAp-uapA-I157F/R287A/R478A/R479A-YFPn uapAΔ uapCΔ:: AFpyrG azgAΔ*	This study
L234M/RRR (in Figure 3.6)	*alcAp-uapA-L234M/R287A/R478A/R479A-YFPc alcAp-uapA-L234M/R287A/R478A/R479A-YFPn uapAΔ uapCΔ:: AFpyrG azgAΔ*	This study
T401P/RRR (in Figure 3.6)	*alcAp-uapA-T401P/R287A/R478A/R479A-YFPc alcAp-uapA-T401P/R287A/R478A/R479A-YFPn uapAΔ uapCΔ:: AFpyrG azgAΔ*	This study
K73A	*uapA-K73A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study

R133A	*uapA-R133A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study
Y137A	*uapA-Y137A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study
K212A	*uapA-K212A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study
R421A	*uapA-R421A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study
Y137A/K138A	*uapA-Y137A/K138A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study
K73A/R421A	*uapA-K73A/R421A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study
K73A/R133A	*uapA-K73A/R133A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study
R133A/R421A	*uapA-R133A/R421A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study
K73A/R133A/R421A or KRR	*uapA-K73A/R133A/R421A-GFP-argB uapAΔ uapCΔ::AfpyrG azgAΔ pabaA1*	This study

Table 2.4 List of strains used in chapter 3.4

Strain	Genotype	Reference
Δ7	*uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	Krypotou and Diallinas 2014
rSNBT1	*[gpdAp::rSNBT1-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
N390T	*[gpdAp::rSNBT1-N390T-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
N390P	*[gpdAp::rSNBT1-N390P-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
G391P	*[gpdAp::rSNBT1-G391P-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
hSVCT1	*[gpdAp::hSVCT1-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
N377T	*[gpdAp::hSVCT1-N377T-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study
N377P	*[gpdAp::hSVCT1-N377P-GFP]pGEM-panB uapAΔ uapCΔ::AFpyrG azgAΔ fcyBΔ::argB furDΔ::riboB furAΔ::riboB cntAΔ::riboB pantoB100 pabaA1*	This study

Strains shown in Figure 3.12 and 3.17 are genetic suppressors of the SYSE and RRR strain derived from UV mutagenesis (see below).

2.1.2 Growth conditions

Standard complete (CM) and minimal media (MM) were used for *A. nidulans* growth (Table 2.6). CM contains all the nutrients and auxotrophies required for fungal growth, whereas MM contains the minimum nutritional supplements possible, depending on the auxotrophic requirements of each strain, and a nitrogen source, according to the desired conditions. Solutions for preparation of *A. nidulans* media are described in Table 2.5. For solid growth mediums, 1-2% agar was added to the liquid medium, before autoclaving. Media and supplemented auxotrophies were at the concentrations given in FGSC (http://www.fgsc.net). Media and chemical reagents were obtained from Sigma-Aldrich (Life Science Chemilab SA, Hellas) or AppliChem (Bioline Scientific SA, Hellas). The pH was adjusted to 6.8 with NaOH 1M or KOH 1M (Figure 3.24 in Na- growth tests). Nitrogen sources were used at final concentrations: $NaNO_3$ 10 mM, ammonium L-(+)-tartrate 10 mM, proline (PRO) 5 mM, uric acid (UA) 0.5 or 2 mM, xanthine (X) 2 mM, adenine (AD) 0.5 or 2 mM, hypoxanthine (HX) 0.5 or 2 mM, inosine (INS) 2 mM, adenosine (ADO) 2 mM and guanosine (GUO) 2 mM. Other nucleobases and analogs were used at the following final concentrations: 5-fluorouracil (5FU) 100 µM, 5-fluorocytosine (5FC) 50 µM, 5-fluorouridine (5FUd) 10µ M, thymine (T) 2 mM and uracil (U) 2 mM.

Table 2.5 Solutions for preparation of *A. nidulans* media

Salt solution		**Vitamin solution**		**Trace elements in 1L H_20**	
KCL	26 g	p-aminobenzoic acid	20 mg	$Na_2B_4O_7$ x 10 H_20	40 mg
$MgSO_4$ $7H_20$	26 g	D-pantothenic acid	50 mg	$CuSO_4$ x 5 H_20	400 mg
KH_2PO_4	76 g	pyridoxine	50 mg	FeO_4P x 4 H_20	714 mg
Trace elements	50 mL	riboflavin	50 mg	$MnSO_4$ x 1 H_20	728 mg
Chloroform	2 mL	biotine	1 mg	Na_2MoO_4 x 2 H_20	800 mg
H_20_{dist}	Up to 1L	H_20_{dist}	Up to 1L	$ZnSO_4$ x 7 H_20	8 mg

Table 2.6 Media used for *A. nidulans* growth

	Complete medium (CM)	**Minimal medium (MM)**	**Sucrose medium (SM)**
Salt solution	20 mL	20 mL	20 mL
Vitamin solution	10 mL	-	-
D-glucose	10 g	10 g	10 g
Casamino acids	1 g	-	-
Bactopeptone	2 g	-	-
Yeast extract	1 g	-	-
Sucrose	-	-	342,4 g
H_20_{dist}	Up to 1L	Up to 1L	Up to 1L

Solid *A. nidulans* cultures were incubated in 37°C or 25°C for 2-4 days and liquid cultures were incubated overnight at 37°C or 25°C, 140 rpm. For the inoculation of cultures, conidiospores were harvested from sporulating culture plates with the use of sterile toothpicks. Repression of protein expression driven by the regulatable alcohol dehydrogenase (*alcAp*) promoter was achieved with the use of 1% glucose, whereas derepression occurred with the use of 0.1% (w/v) fructose as a sole carbon source. Induction of the same proteins was achieved by addition of 0.4% (v/v) ethanol in the non-repressing culture media.

E. coli bacterial cultures (strain DH5a) were grown on Luria-Bertani (LB) medium (Bacto Tryptone 10 g, NaCl 10 g, BactoYeast Extract 5 g for 1L). The pH was adjusted to 7.0 with NaOH 1M. After bacterial transformation, colony selection was achieved with the use of 100 μg/mL ampicillin. Solid cultures were incubated overnight at 37 °C. Liquid cultures were incubated in the same conditions, at 200 rpm.

Agar plates were stored at 4°C, to prevent a serious loss of fungal viability for some weeks. For long-term storage, glycerol stocks were prepared. *A.* nidulans conidiospores were harvested from ¼ of a fresh 100mm CM plate, in 1ml solution of 1:1 glycerol:PBS (NaCl 8 g, KCl 0,2 g, Na_2PO_4 1,44 g, KH_2PO_4 0,24 g, pH 7.4 with 1 N HCl), in a sterile eppendorf tube. The solution was mixed well and the glycerol stocks were stored for long periods at -80°C. For reviving stored fungal cultures, a small quantity of the stock was inoculated on appropriate media and then analyzed with growth tests to verify that no contamination had occurred.

2.2 Genetic crosses and progeny analysis

The two different parental strains, with at least one non-mutual auxotrophy, are inoculated in pairs in Petri dishes containing MM and the appropriate supplements. After incubation at 37°C until the colonies met (2-3 days), small agar pieces containing heterokaryons were transferred on small Petri dishes (30 mm diameter) containing MM with $NaNO_3$ and any mutual auxotrophies. On this plates only heterokaryons were able to grow by producing the missing supplements of each parental strain. Plates were incubated at 37 °C for 2-3 days and were tightly sealed with adhesive tape and incubated at 37 °C for at least 14 days. After that time the fruiting bodies, cleistothecia, usually appeared as black spheres [84].

Single cleistothecia, usually 8, were identified using a stereoscope and isolated with a sterile toothpick. Surrounding cells (Hülle cells, aerial hyphae or any conidiospores) were removed from the cleistothecia by rolling on an agar plate. Each cleistothecium was burst

open against the walls of an eppendorf tube containing 1 mL of sterile H_2O_{dist} where the ascospores were released, creating an ascospore suspension. 10 µl from each suspension were plated on a Petri dish containing only the mutual auxotrophies, in order to identify the crossed cleistothecia, and incubated at 37°C for 2-3 days. 3-10 µl of the suspension from one recombinant cleistothecium were plated in Petri dishes containing MM with the desirable auxotrophies in order to obtain single colonies. Several colonies were selected and were further analyzed by growth tests, PCR or fluorescence microscopy to identify the progeny of interest.

2.3 Epifluorescence microscopy

Images were obtained using a Zeiss Axio Observer Z1 inverted epifluorescent microscope, and the resulting images were acquired with an AxioCamMR3 digital camera using the Zen lite 2012 software. Images were analyzed using Zen 2012 software and Adobe Photoshop CC 2015. Samples were prepared as follows: Germlings incubated in sterile 35 mm Petri dishes, high glass bottom (ibidi, Germany) in 2 ml liquid MM 1% glucose pH 5.2 supplemented with the appropriate supplements and $NaNO_3$ as the nitrogen source. The culture was then incubated at 25°C, for 16-18 h, protected from light.

For the observation of UapA molecules tagged with each of the two halves of the Yellow Fluorescent Protein (YFP), expressed under the control of the *alcAp* promoter mycelia were grown for 16-18 h in derepressing media (minimal media supplemented with $NaNO_3$ and 0.1% (w/v) fructose as a sole carbon source). For the observation of *de novo* synthesized transporters mycelia were grown for 14-16 h in repressing media (minimal media supplemented with $NaNO_3$ and 1% (w/v) glucose as a sole carbon source) and were shifted in derepressing media for 1-4 h. In both cases induction of expression was achieved by addition of 0.4 % (v/v) ethanol in derepressing media.

2.4 DNA manipulations and molecular cloning

2.4.1 PCR, Site Directed Mutagenesis and oligonucleotides

PCR was performed in order to amplify specific regions of DNA for diagnostic or cloning purposes. KAPA Taq (Kapa Biosystems) was used for conventional PCR reactions whereas for

high fidelity amplifications a proof reading polymerase KAPA HiFi (Kapa Biosystems) was used. The reactions were performed according to the manufacturer's instructions, the Tm 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, according to the instructions accompanying the QuikChange® Site-Directed Mutagenesis Kit, using KAPA HiFi polymerase. For each mutation, two long (35-50 bases) 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. If possible, the designed mutations led to the introduction of a restriction site in order to enable diagnostic digestion. To calculate the Tm, the previous formula is transformed to T_m= [69.3 + 0.41 (GC*100/L) – 650/L] – 5 – X, where X is the percentage of mismatches included in the oligonucleotide. Amplification was checked by electrophoresis of 10 µl of the PCR product next to the same amount of template used in the PCR reaction on a 1% agarose gel. If the band intensity of the PCR product was higher than the template's the amplification was considered successful and the rest of the PCR product was digested with 10 u of the restriction enzyme DpnI. This enzyme recognizes and cleaves a methylated 4-nucleotide sequence target (GA^m | TC), so parental non-mutated plasmids that are methylated were fragmented. The mixture was incubated for 2 h at 37°C. After incubation, 2 µl of the DpnI-treated product was used for transformation of *E.coli* competent cells. Single ampicillin-resistant colonies where cultivated in liquid cultures and plasmids were isolated. The mutations were verified by digestion with the appropriate restriction enzyme and/or sequencing. A plasmid with the desired mutation was finally used for transformation of the appropriate *A. nidulans* strain protoplasts.

Composition and conditions of these PCR reactions are described in the Table 2.7 and 2.8. Oligonucleotides used in this study are described in Table 2.9

Table 2.7 Composition of conventional, high fidelity and site-directed mutagenesis PCR reactions

Components	Final concentration		
	Conventional	High-fidelity	Mutagenesis
10x Polymerase buffer (with 1.5 mM $MgCl_2$)	1x	-	-
dNTPs	200 µM each	-	-
2x Polymerase Ready Mix	-	1x	1x
Forward primer	1 µM	1 µM	1 µM
Reverse primer	1 µM	1 µM	1 µM
DNA polymerase	KAPATaq 1 u	-	-
DNA template	10 ng	10 ng	10 ng
$H2O_{dist}$	Up to 25 µl	Up to 25 µl	Up to 25 µl

Table 2.8 Conditions used for conventional, high fidelity and site-directed mutagenesis PCR

Step	Conventional		High-fidelity		Mutagenesis	
	Temperature (°C)	Time	Temperature (°C)	Time	Temperature (°C)	Time
1	95	5 min	95	3 min	95	3 min
2	95	30 sec	98	20 sec	98	20 sec
3	T_m-5	30 sec	T_m	15 sec	60	1 min
4	72	1 min/kb	72	30 sec/kb	72	30 sec/kb
5	72	10 min	72	5 min	72	5 min
6	12	∞	12	∞	12	∞

*Steps 2-4 are repeated for 25 or 35 or 18 cycles in conventional, high-fidelity and mutagenesis PCR respectively

Table 2.9 Oligonucleotides used in this work

	Oligonucleotide name	Sequence 5'→ 3'
1	uapA A407S/Q408P F	CCCCCATGACGACCTTTTCGCCGAACAACGGCGTG
2	uapA A407S/Q408P R	CACGCCGTTGTTCGGCGAAAAGGTCGTCATGGGG
3	uapA F406S/A407S/Q408P F	GACCCCCATGACGACCTCTTCGCCGAACAACGGCGTG
4	uapA F406S/A407S/Q408P R	CACGCCGTTGTTCGGCGAAGAGGTCGTCATGGGGGTC
5	uapA T405S/F406S/A407S F	CAATGACCCCCATGACGTCGTCGTCGCAGAACAACGGCGTG
6	uapA T405S/F406S/A407S R	CACGCCGTTGTTCTGCGACGACGACGTCATGGGGGTCATTG
7	uapA ORF SpeI F	CGCGACTAGTATGGACAACTCCATCCATTCAAC
8	uapA ORF NotI R	CGCGGCGGCCGCAGCCTGCTTGCTCTGATACTC
9	uapA 405TS/FY/AS F (SYSQ)	CAATGACCCCCATGACGAGTTACTCGCAGAACAACGGCGTG
10	uapA 405TS/FY/AS R (SYSQ)	CACGCCGTTGTTCTGCGAGTAACTCGTCATGGGGGTCATTG
11	405TS/FY/AS on Q408E F (SYSE)	CAATGACCCCCATGACGAGTTACAGTGAGAACAACGGCGTGATTG

12	405TS/FY/AS on Q408E R (SYSE)	CAATCACGCCGTTGTTCTCACTGTAACTCGTCATGGGGGTCATTG
13	uapA TFAQ F (SSSP on SSP)	CAATGACCCCCATGACGTCCTCTTCGCCGAACAAC
14	uapA TFAQ R (SSSP on SSP)	GTTGTTCGGCGAAGAGGACGTCATGGGGGTCATTG
15	uapA R481G F	CGCCGTTCACAAGGCGAAATGGGTTTATCCTCACCGCGTC
16	uapA R481G R	GACGCGGTGAGGATAAACCCATTTCGCCTTGTGAACGGCG
17	uapA T526L F	GATCGAGCTCGTGCTTGAGCTCGGGTTTGCGGTCACGGC
18	uapA T526L R	GCCGTGACCGCAAACCCGAGCTCAAGCACGAGCTCGATC
19	uapA F528M F	GAGCTCGTGCTTGAGACGGGGATGGCGGTCACGGCATTGTAAGTCTTC
20	uapA F528M R	GAAGACTTACAATGCCGTGACCGCCATCCCCGTCTCAAGCACGAGCTC
21	uapA N410V F	CATGACGACCTTTGCGCAGAACGTCGGCGTGATTGCCCTCACTC
22	uapA N410V R	GAGTGAGGGCAATCACGCCGACGTTCTGCGCAAAGGTCGTCATG
23	uapA N410I F	GACGACCTTTGCGCAGAACATAGGCGTGATTGCCCTCACTC
24	uapA N410I R	GAGTGAGGGCAATCACGCCTATGTTCTGCGCAAAGGTCGTC
25	N410V on SYSE F	CATGACGAGTTACAGTGAGAACGTAGGCGTGATTGCCCTCACTC
26	N410V on SYSE R	GAGTGAGGGCAATCACGCCTACGTTCTCACTGTAACTCGTCATG
27	N410I on SYSE F	GACGAGTTACAGTGAGAACATAGGCGTGATTGCCCTCACTC
28	N410I on SYSE R	GAGTGAGGGCAATCACGCCTATGTTCTCACTGTAACTCGTC
29	uapA K73A F	CCTTTTTTTGGCCTCAACGAGGCGATTCCCGTGCTGTTGGC
30	uapA K73A R	GCCAACAGCACGGGAATCGCCTCGTTGAGGCCAAAAAAAGG
31	uapA R133A F	GTCGATGGTCCAGATAACGGCCTTTCATATCTACAAGACACCGTAC
32	uapA R133A R	GTACGGTGTCTTGTAGATATGAAAGGCCGTTATCTGGACCATCGAC
33	uapA Y137A F	CCAGATAACGCGATTTCATATCGCCAAGACACCGTACTATATCGG
34	uapA Y137A R	CCGATATAGTACGGTGTCTTGGCGATATGAAATCGCGTTATCTGG
35	uapA K212A F	CGCTTTCGTGCCTCCCGCGGTGATACAGAAAATCTTCCCG
36	uapA K212A R	CGGGAAGATTTTCTGTATCACCGCGGGAGGCACGAAAGCG
37	uapA R421A F	CCCTCACTCGCTGCGCAAACGCATGGGCCGGATACTGCTGCTG
38	uapA R421A R	CAGCAGCAGTATCCGGCCCATGCGTTTGCGCAGCGAGTGAGGG
39	uapA Y137A/K138A F	GATAACGCGATTTCATATCGCAGCGACACCGTAAGTTAGC
40	uapA Y137A/K138A R	GCTAACTTACGGTGTCGCTGCGATATGAAATCGCGTTATC
41	uapA T401P F	GTCGCTGCCCTTGCGACAATGCCCCCCATGACGACCTTTGCG
42	uapA T401P R	CGCAAAGGTCGTCATGGGGGGCATTGTCGCAAGGGCAGCGAC
43	UapA pr	CTTATCTCTCCGCTCCCAG
44	uapA start	CCATCCATTCAACCGACGGCC
45	uapA M	GCCTGCGAGTGCATCGGTG
46	uapA stop	GCCTGCTTGCTCTGATACTC
47	rSNBT1 ORF SpeI F	CGCGACTAGTATGAACTCTGCAGTCTGCAGC
48	rSNBT1 ORF NotI R	CGCGGCGGCCGCCATCTTGGTTTCTGTAACACTCC
49	rSNBT1 N390T F	GGGCCTGGGGAACCGGAACTGGTACCACGTCCTACAGCG
50	rSNBT1 N390T R	CGCTGTAGGACGTGGTACCAGTTCCGGTTCCCCAGGCCC
51	rSNBT1 N390P F	GGGCCTGGGGAACCGGACCGGGTACCACGTCCTACAGCG
52	rSNBT1 N390P R	CGCTGTAGGACGTGGTACCCGGTCCGGTTCCCCAGGCCC
53	rSNBT1 G391P F	GGCCTGGGGAACCGGAAACCCGACCACGTCCTACAGCGAGAAC
54	rSNBT1 G391P R	GTTCTCGCTGTAGGACGTGGTCGGGTTTCCGGTTCCCCAGGCC

55	rSNBT1 seq S R	CACCAGCATCTGGATGCAG
56	rSNBT1 seq M F	CAAAGGCAGCGTTCTGAGC
57	hSVCT1 ORF SpeI F	CGCGACTAGTATGAGGGCCCAGGAGGACCTCG
58	hSVCT1 ORF NS NotI R	CGCGGCGGCCGCGACCTTGGTGCACACAGATGCAG
59	SVCT1 seq M F	CATCATGGCTATTGCACCCTG
60	hSVCT1 N377T F	GGCTATTGGGCACGGGCACCGGGTCCACCTCGTCCAG
61	hSVCT1 N377T R	CTGGACGAGGTGGACCCGGTGCCCGTGCCCAATAGCC
62	hSVCT1 N377P F	GGGCTATTGGGCACGGGCCCCGGGTCCACCTCGTCCAG
63	hSVCT1 N377P R	CTGGACGAGGTGGACCCGGGGCCCGTGCCCAATAGCCC
64	GFP NotI R	CGCGCGGCCGCTTACTTGTACAGCTCGTCC

2.4.2 Plasmid extraction from *E.coli*

For plasmid extraction from *E.coli*, cells from a 4 mL saturated liquid bacterial culture were collected and plasmid preparation was performed using the Nucleospin Plasmid kit (Macherey-Nagel, Lab Supplies Scientific SA, Hellas).

2.4.3 Agarose gel electrophoresis

Agarose gel electrophoresis was performed for the analysis of the size and conformation of DNA in a sample, and the separation and extraction of DNA fragments. The fragments are separated based on their length and conformation by applying an electrical field to move the charged molecules through an agarose matrix where the DNA (negatively charged) moves towards the positively-charged anode during the electrophoresis. Agarose gels were made by addition of 1% w/v agarose in 1x TAE buffer (Tris-acetate 0.04 M, EDTA 0.001 M). The mixture was heated until agarose was completely dissolved and left to cool down for a while before addition of 0.5 μg/mL ethidium bromide which intercalates into the major grooves of the DNA/RNA and fluoresces under UV light, making DNA/RNA visible. A comb was placed in the cast to create wells for sample loading and the gel was completely set before use. Once the gel had set, the gel was placed in the electrophoresis tank filled up with TAE 1x buffer until the gel was covered and the comb was removed. DNA samples were mixed with loading buffer and loaded in the gel. To determine the size of the fragments, a molecular weight marker was loaded along with the samples and gels were run at 100V for 10-40 min. The gel was visualized under a UV transilluminator and DNA bands were either photographed or quickly excised from the gel for cloning purposes and purified as described in the manufacturer's instructions. DNA

bands gel extraction were performed using the Nucleospin Extract II kit (Macherey-Nagel, Lab Supplies Scientific SA, Hellas).

2.4.4 Molecular cloning

In order to assemble recombinant DNA, circular plasmid DNA and insert were cleaved with one or more restriction enzymes creating compatible ends. The insert was amplified by PCR with primers designed to add the desired restriction sites to its termini. Digestion was performed with the appropriate amount of DNA, 1x restriction enzyme buffer, 0.5-1 μL of the restriction enzyme and distilled water to a final volume of 20-50 μL. The digested DNA products were analyzed by agarose gel electrophoresis and were purified using the Nucleospin Extract II kit (Macherey-Nagel). When cloning with one restriction enzyme the digested vector was treated, before gel electrophoresis, with 1 μL alkaline phosphatase for 15 min at 37°C (Takara) in order to reduce re-ligation of the vector without the insert [120]. To generate the recombinant DNA 1 μL of T4 DNA ligase (TaKaRa) was used along with the purified vector and insert at a 1:3 concentration ratio and 1x ligase buffer in 10 μL total volume. The mixture was incubated at 25°C for 1.5 h or at 4°C overnight. After ligation, the mixture was used for transformation of *E.coli* competent cells. Single ampicillin-resistant colonies where cultivated in liquid cultures and plasmids were isolated.

Plasmid pAN510-GFP or its derivative pAN510exp-GFP were used for the complementation of the *argB2* auxotrophy (see later; strains shown in chapters 3.1 and 3.3) [10,121,122]. The vector used for complementation of the *pantoB100* auxotrophy (strains shown in chapters 3.2 and 3.4) was a modified pGEM-T-easy vector carrying a version of the *gpdA* promoter, the *trpC* 3' termination region, and the *panB* selection marker [43]. For bimolecular fluorescence complementation (BiFC) analyses, the N-terminal half of yellow fluorescent protein (YFPn; 154 amino acids of YFP), or the C-terminal half of YFP (YFPc; 86 amino acids of YFP) was amplified from plasmids PDV7 and PDV8 [123] and cloned into pAN510exp-alcAp or pAN520exp-alcAp [124], followed by cloning of the UapA ORF of the mutants. For each UapA, rSNBT1 or hSVCT1 mutant the ORF of the wild-type version was cloned in the corresponding plasmid and the desired mutation was introduced by site-directed mutagenesis.

2.4.5 Preparation of *E.coli* competent cells

In order to pick single colonies from the glycerol stock, a trace of DH5a *E. coli* cells was streaked on an LB agar plate and was incubated at 37°C overnight. 5 ml of liquid LB medium were inoculated by a single colony and were incubated at 37°C, 200 rpm for 16 h. Then, 0.5 ml of the culture were used to inoculate 400 ml LB medium and the culture was incubated at 37°C, 260 rpm. When optical density at 600 nm (OD_{600}) of 0.45-0.55 was reached, cells were collected by centrifugation at 4°C, 4500 g for 5 min and the pellet was re-suspended in 0.4x original volume of sterile ice-cold transformation buffer 1 (30 mM CH_3COOK, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM $RbCl_2$, 15% glycerol, pH 5.8 with 1 M CH_3COOH) and incubated on ice for 5 min. The suspension was centrifuged at 4°C, 4500 g for 5 min and the pellet was re-suspended in 1/25 original volume of sterile ice-cold transformation buffer 2 (10 mM MOPS pH 6.5, 75 mM $CaCl_2$, 10 mM $RBCl_2$, 15% glycerol, pH 6.5 with 1 M KOH) and incubated on ice for 15-60 min. Finally, aliquots of 200 μL competent cells were distributed in sterile eppendorf tubes and stored at -80°C.

2.4.6 Transformation of *E.coli* competent cells

Plasmid DNA (0.01-0.5 μg) was added in 200 μL of defrosted competent *E. coli* cells and the suspension was mixed and incubated on ice for 20-30 min. Then, the cells were subjected to heat shock at 42°C for 90 sec (in order to increase plasma membrane permeability), followed by direct incubation on ice for 2 min. In order for the cells to recover from the heat shock and express the ampicillin resistance gene encoded on the plasmid, 1 mL of liquid LB medium was added in the tube and the cells were incubated at 37°C for 1 h. Finally, cells were collected by centrifugation at 12000 rpm for 1 min re-suspended in 100 μL of the medium and plated on LB agar plated containing 100 μg/ml ampicillin. Plates were incubated at 37°C overnight and colonies were selected and further analyzed.

2.4.7 DNA extraction from *A. nidulans*

Genomic DNA extraction from *A. nidulans* was performed as described in FGSC (http://www.fgsc.net). More specifically, the strain of interest was inoculated in a CM or MM culture plate and incubated for 4 days at 37°C. Conidiospores were harvested from 1/4 of the

plate in 25 mL of liquid MM (containing $NH4^+$ and the required auxotrophies) and cultures were incubated overnight (until full growth and formation of mycelia) at 37°C, 140-150 rpm. The mycelia were collected by a Blutex filter, dried with paper towels and immediately frozen in liquid nitrogen. The mycelia were pulverized in a mortar with a pestle in the presence of liquid nitrogen and ~200 mg of the fine powder were transferred in a 2 mL eppendorf tube and 800 µl of DNA extraction buffer (Tris-HCl 0.2 M pH 8.0, Sodium Dodecyl Sulfate (SDS) 1%, Ethylenediaminetetraacetic acid (EDTA) 1mM pH 8), were added. The mixture homogenized by vortexing and was incubated on ice for 20 min. Then, 800 µl of pure phenol were added and the mixture was shaken vigorously for protein denaturation to occur (and thus separate DNA from proteins) at room temperature (RT) before centrifugation for 5 min at 12000 rpm, RT. 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 upper aqueous phase, containing the DNA, was transferred to a new eppendorf tube with equal volume of chloroform which allows proper separation of the organic and aqueous phase keeping DNA protected into the aqueous phase and denaturation of lipids. The mixture was shaken again vigorously and centrifuged for 5 min at 12000 rpm, RT. 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 upper aqueous phase was recovered and transferred to a new 1.5 mL tube. Then, the DNA was precipitated by addition of equal volume of isopropanol and 1/10 volume of 3 M sodium acetate (pH 5.3) followed by mild agitation and centrifugation at 12,000 rpm for 5 min, RT. The supernatant was discarded and the pellet was washed with 200 µl of 70% EtOH. After spinning for 2 min, EtOH was removed with a pipette and the pellet was dried for 10 min at 50°C. The pellet was finally re-suspended in 50-100 µl of dH2O containing RNase (100 µg/mL) and incubated for 30 min at 37°C. To quantify and check the quality of the extracted DNA, 2-3 µl of the DNA solution were analyzed by agarose gel electrophoresis.

2.5 Transformation of *A. nidulans*

Transformation was performed as described previously [125] using a *ΔazgA ΔuapA ΔuapC::AfpyrG pabaA1 argB2* mutant strain, named Δ3 or a *ΔfurD::riboB ΔfurA::riboB ΔfcyB::argB ΔazgA ΔuapA ΔuapC::AfpyrG ΔcntA::riboB pabaA1 pantoB100* mutant strain, named Δ7. Selection was based on complementation of the arginine auxotrophy *argB2* or the pantothenic acid auxotrophy *pantoB100* respectively. For the selection of UapA-SYSE

suppressors an *nkuA* DNA helicase deficient strain (*nkuAΔ::argB ΔfurD::riboB ΔfcyB::argB ΔazgA ΔuapA ΔuapC::AfpyrG ΔcntA::riboB pabaA1 pantoB100 pyroA4*) was used in order to exclude heterologous recombination events and get single copy transformants, by homologous recombination in the *panB* locus. In this strain selection was based on the complementation of the *pantoB100* auxotrophy.

In particular, the strain of interest was inoculated in a CM-agar plate and incubated at 37 °C for 4 days. Conidiospores were collected, filtered through a blutex and used to inoculate a 200 ml liquid MM with $NH4^+$ as a nitrogen source and the appropriate supplements (vitamins etc.). The culture was incubated at 37°C, 140-150 rpm until germ tubes were observed under the microscope (4-4.5 h). Once the conidia were at the germinative phase the culture was transferred in sterile falcons and centrifuged at 4000rpm, 10min. The pellet was re-suspended in 20 ml of Solution I (1.2 M $MgSO_4$, 10 mM orthophosphate pH 5.8) and the spore solution was transferred in a sterile 250 ml flask with 200mg of lytic enzyme (Glucanex). The lysing enzymes digest the cell wall and protoplasts are generated. The suspension was incubated at 30°C, 60 rpm for 2 h. Protoplast were concentrated by centrifugation (4000 rpm, 10 min) and washed with 10 ml of Solution II (1 M Sorbitol, 10 mM Tris-HCl pH 7.5, 10 mM $CaCl_2$) and were re-suspended in the same solution (1-10 mL, depending on the density of the culture). ¼ of the total volume of Solution III (60% (w/v) PEG6000, 10 mM Tris-HCl pH 7.5, 10 mM $CaCl_2$) was added and protoplasts were distributed in eppendorf tubes. At this step protoplasts could be frozen at -80 °C after addition of 1% v/v DMSO.

Plasmid DNA (1.5 -2 µg) was added in an eppendorf tube with protoplasts and incubated on ice for 15 min. A control tube without plasmid DNA was included in order to evaluate whether the protoplasts and the solutions used were free of contaminations. 500 µl of Solution III were added, mixed and incubated for another 15 min at RT, followed by centrifugation (6000 rpm, 10 min). Protoplasts were, then, washed and re-suspended in 1 ml and 200 µl of Solution II, respectively. Finally, protoplasts were mixed with 4 ml of melted Top SM in sterile falcons (Sucrose MM, 0.35% agar) and quickly used to inoculate solid SM plates (1% agar). The SM-bottom was supplemented with $NH4^+$ and vitamins depending on the selection marker carried in the plasmid; e.g. for an *argB2* auxotroph strain transformed with an *argB* carrying plasmid, the growth medium must not contain arginine. The plates were incubated for 4-5 days at 37°C, transformants were isolated by streaking on MM, analyzed by growth tests and verified by PCR and Southern analysis.

2.6 RNA extraction from *A. nidulans*

The strain of interest was inoculated in a CM or MM culture plate and incubated for 4 days at 37°C. Conidiospores were harvested from the plate in 100 mL of liquid MM (containing $NH4^+$ and the required auxotrophies) and cultures were incubated for 14-16 h (until full growth and formation of mycelia) at 25°C, 140-150 rpm. The mycelium was collected and grinded with liquid nitrogen similarly to the DNA extraction protocol. About 100 mg of the powder were collected in an eppendorf tube and 10 mg acid-washed glass-beads (particle-size 425-600 μm) and 1 mL of RNAtidy G (Applichem) were added. The mixture homogenized by vortexing and was incubated at RT for 2-3 min. Then, 200 μL of chloroform were added, followed by vortexing and incubation at RT for 10 min before centrifuged at 4°C, 12000 rpm for15 min. The upper aqueous phase was transferred to a new eppendorf tube and equal volume of isopropanol was added. The mixture was mixed by inversion and incubated for 10 min on ice, before centrifuged at 4°C, 12000 rpm for15 min. The supernatant was discarded using a pipette and the pellet was washed with 200 μL 70-80% v/v EtOH (with DEPC; diethyl pyrocarbonate; water). Then, the sample was vortexed, centrifuged at 4°C, 10000 rpm for 3 min and again the supernatant was discarded using a pipette followed by a second wash with 200 μL 100% EtOH. The mixture was mixed with vortex and centrifuged at 4°C, 10000 rpm for 3 min. Ethanol was discarded with a pipette and the pellet was left to air dry for 15-20 min, but some moisture should be maintained in order to be soluble again. Finally, the pellet was resuspended in 20-30 μL DEPC water and if needed incubated at 55°C to completely dissolve RNA. The quality and concentration of the RNA were estimated by measuring OD at 260 nm and 280 nm (OD260/OD280= ~2 corresponds to pure RNA). The above can be checked also with agarose gel electrophoresis (in 2% agarose).

Gloves were worn during the RNA extraction procedure and every solution is prepared with DEPC (deactivates RNases) water so that the samples are protected from RNases. DEPC is added in water and incubated overnight at 37°C and then deactivated by autoclaving.

2.7 Northern blot analysis

RNA that was extracted as described above, was analyzed according to an adjusted protocol based on [120]. First of all, V μL that corresponded to 10 μg RNA was denaturated with 3V denaturation solution [DMSO (67.44% v/v), glyoxal (19.35% v/v), Orthophosphate (11,87%

v/v)]. The mixture was incubated at 50°C for 1h and then transferred on ice for 5 min. The samples were analyzed in a 1% w/v agarose gel made with orthophosphate buffer 10 mM pH 7 that was also used as a running buffer. Denaturation and electrophoresis on a denaturing gel was needed to avoid the formation of secondary and tertiary RNA structures that can change the migration rate. 0.5V µL of loading buffer (Orthophosphate 10 mM pH 7, Bromophenol-blue 0.2% w/v, glycerol 50% v/v) was added to the samples (4V µL) that were loaded to the gel. Electrophoresis was performed in an electrical field corresponding to 48 mA for 5 min and then a peristaltic pump with continuous flow was set in order to assure homogeneity of the running buffer (orthophosphate is dissociated during the electrophoresis and the pump transfers solution all over the gel to avoid concentration gradient). When the dye of the loading buffer had run ~9 cm from the wells, the electrophoresis was stopped and the gel was cut keeping only the part with the samples.

RNAs were transferred to a positively charged nylon membrane by capillary action from a region of high water potential to a region of low water potential making the following construction: a wad of paper that dips into the tray of 20xSSC buffer (NaCl 3 M, $Na_3C_6H_5O_7$ 0.3 M pH 7 with NaOH), the gel, the positively charged membrane, a stack of paper towels and a weight on the top to succeed equal distribution of the weight. The transfer was performed overnight and after the completion, the membrane was exposed to UV light for 5 min. This step was performed in order for the RNA to be immobilized on the membrane through UV cross-linking so that it does not come off of the membrane during the subsequent hybridization steps. Then a dye (Methylene-blue 0.03% w/v, CH_3COONa 0.5 M pH 5.2) was used to detect ribosomic RNAs to confirm equal loading and success of the transfer. After detection the membrane was discolored with SDS 1% w/v and 0.2xSSC.

The next step is detection of the RNA of interest with an appropriate probe complementary to a region of the RNA target. The probe is synthesized by PCR using [^{32}P] nucleotides. Before the reaction, 200-400 ng of DNA of interest what was acquired by PCR (~ 1 kb) were mixed with 10 ng random hexamers (used as primers) and dH_2O up to 21 µL final volume and boiled for 2 min to make DNA single-stranded. Then, the mixture was moved directly on ice to remain single-stranded and nucleotides (10 mM; except of the radiolabeled one), Klenow polymerase, Klenow buffer and dH_2O up to 50 µL were added. Finally the radiolabeled nucleotide was added and the sample was incubated at 37°C for 1 h. In order to remove unincorporated radiolabeled nucleotides the sample was then added in a sepharose column and the probe in the resulting sample was denaturated with NaOH. Before hybridization of the probe, pre-hybridization of the membrane is important to eliminate non-

specific background. For pre-hybridization, the membrane was placed in Church solution (BSA 1% w/v, NaH_2PO_4/NaH_2PO_4 0.5 M pH 7, SDS 7% w/v, EDTA 1 mM) and incubated at 65°C for 2 h under mild agitation, so that non-specific molecules bind to the sites of the membrane without RNA. After the pre-hybridization, the probe was added in the solution and the membrane was incubated at an optimal temperature for the probe and the RNA, in this case 65°C, overnight under mild agitation. The probe can be collected in a centrifuge tube, stored at -20°C and be re-used after boiling for 10 min. Hybridization must be followed by four washing stages, three with 2xSSC solution and 0.1 w/v SDS and one last with 0.2xSSC and 0.1% w/v SDS for 30 min, to remove non-specifically bound probes and reduce non-specific background. Finally the membrane was placed in an autoradiography cassette with an X-ray film (Kodak X-omat) on top and incubated at -80°C until bands appear during film development (20 min to 48 h) using Developer and Fixer solutions (Kodak).

2.8 Native MS of UapA

Expression and purification of the UapA K73A/R133A/R421A and K73A/R133A/R421A/T401P mutants was carried out as described previously [12]. In brief, the mutants were individually expressed as C-terminally GFP8His-tagged constructs in *Saccharomyces cerevisiae* FGY217 cells (12L), using vector pDDGFP2. Yeast cells were incubated at 30°C with shaking at 300 rpm to an OD_{600} of 0.6. Galactose was then added to the cultures to a final concentration of 2% to induce UapA expression. After incubation for a further 22 h, the cells were harvested by centrifugation and resuspended. Cells were broken in a Constant Systems cell disruptor and the membranes isolated by centrifugation. The membranes were resuspended, flash-frozen, and stored at -80°C. Membranes were solubilized for 1 h in n-dodecyl-β-D-maltoside. Unsolubilized membranes were removed by centrifugation. The supernatant was incubated with Ni^{2+}-NTA resin for 2 h. The His-tagged UapA bound to the resin was then washed with buffers containing 10 and 30 mM imidazole to remove contaminants. UapA was eluted with buffer containing 250 mM imidazole before overnight dialysis to dilute the imidazole from the protein sample. During dialysis, the protein was cleaved using a His-tagged TEV protease. The sample was run through a His-trap column, from which UapA was eluted in the flow through, to remove the His-tagged GFP and TEV. The sample was then loaded onto a size-exclusion chromatography column. Fractions containing monodisperse UapA were analyzed by SDS-PAGE and concentrated to 20 µM, flash frozen, and stored at -80°C.

Native MS of UapA K73A/R133A/R421A/T401P was carried out as described previously [12]. In brief, UapA was buffer exchanged into MS buffer [250 mM EDDA (pH 6.3), 0.014% DDMLA (v/v), and 10 mM L-serine] to a UapA concentration of 20 μM using Micro Bio-Spin 6 columns (Bio-Rad, Hercules, CA). UapA was loaded into gold-coated capillaries and the protein sprayed into a Synapt G2-Si (Waters Associates, Milford, MA) by nanoelectrospray ionization. The following conditions were used in the mass spectrometer for optimal peak resolution: capillary voltage +1.3-1.5 kV, sampling cone voltage 150 V, trap collision energy (CE) 200 V, transfer CE 0 V, backing pressure 3.88 mbar, trap and transfer pressure (argon) 1.72e-2 mbar, and ion mobility cell pressure (nitrogen) 2.58 mbar. The mass spectrometer was calibrated using cesium iodide. Spectra were recorded and processed using Masslynx 4.1 software (Waters Associates). The relative abundances of each oligomeric state were quantified by UniDec [126] as described previously [12].

2.9 Phylogenetic analysis datasets and methods

Representative NAT sequences from organisms across all domains of life were selected, representing all major taxa, with particular emphasis in characterized family members of known specificity. For the trees shown in Figure 3.1 and 3.3, which address in detail the plant and animal NATs (SVCT clades) and the evolution of L-ascorbate transporters, at first the NAT domain containing sequences in complete sequenced genomes was detected, coming from Ensembl release 91, Ensembl Metazoa release 38, and selected species from Ensembl Plants release 38. To detect NAT sequences in the genomes HMMER and the PF00860 (*Xan_ur_permease*) Pfam HMM profile were used, and performed an hmmsearch using the GA (gathering) curated threshold (cut_ga parameter). For the 728 detected sequences, a multiple sequence alignment was built using MAFFT v6.861b in the L-INS-i mode, designed for "set of sequences containing sequences flanking around one alignable domain". The resulting alignment was trimmed using trimAl v1.3 [127] using a gap score cut-off of 0.1. The best-fit evolutionary model selection was selected prior to the phylogenetic inference using ProtTest 3 [128]. 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 [129]. An ML tree was inferred with RAxML v8.2.11 [130] 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). We

used the exact same methods for the phylogenetic tree of only the SVCT clade, a subset of the previous tree, consisting of 674 sequences, apart from the MAFFT algorithm used, which was the G-INS-I in this case (assumes that entire region can be aligned). ETE toolkit [131] was used for the execution of the phylogenetic pipelines and the visualization of all phylogenetic trees in Figure 3.1 and 3.3.

2.10 Molecular simulations shown in chapter 3.2

Protein Model Construction. Models of UapA-TFAE, UapA-SYSE and UapA-SYSQ were generated by mutating the specific residues on Maestro platform (Maestro, version 2018-4, Schrödinger, LLC, New York, NY, 2018), on the basis of the structure of UapA crystallized in the inward-open conformation (PDB ID: 5i6c). In order to construct the model of UapA outward conformation Band3 transporter (4yzf) was used as template [45]. Band3 is an anion exchanger of 14 transmembrane helices, crystallized in the outward conformation. The model was constructed using Prime Homology in Prime 2018-4 (Schrödinger, LLC, New York, NY, 2018), by superimposing helices one by one and aligning all the important residues.

Protein Preparation. The protein was prepared using the Protein Preparation Protocol implemented in Schrodinger suite (Schrödinger Suite 2018, Protein Preparation Wizard) and accessible within the Maestro program (Maestro, version 2018-4, Schrödinger, LLC, New York, NY, 2018). Specifically, at first hydrogen atoms were added followed by an optimization of the orientation of hydroxyl groups of Asn, Gln and of the protonation state of His in order to maximize hydrogen bonding. The final step was that of minimization of the protein, using the OLS3 force field [132].

Ligand preparation. Ligand preparation for docking was performed with LigPrep application (LigPrep, version 2018-4, Schrödinger, LLC, New York, NY, 2018), which consists of a series of steps that perform conversions, apply corrections to the structure, generate ionization states and tautomers, and optimize the geometries. The force field chosen was OPLS3 [132].

Induced Fit Docking. Schrodinger Suite protocol was used (Schrödinger Suite 2018-4 Induced Fit Docking protocol; Glide, Schrodinger, LLC, New York, NY, 2016; Prime, Schrödinger, LLC, New York, NY, 2018), taking into account the side chain or backbone movements, or both, upon ligand binding, thus circumventing an inflexible binding site. In the first softened-potential docking step, of the protocol, 20 poses per ligand were retained. In

the second step, for each docking pose, a full cycle of protein refinement was performed, with Prime 2018-4 (Prime, version 3.0, Schrödinger, LLC, New York, NY, 2018) on all residues within 5 Å of any out of the 20 ligand poses. The Prime refinement starts by performing conformational search and by minimizing the side chains of the selected residues. After convergence to a low-energy result, an additional minimization of all selected residues (side chain and backbone) is performed with the Truncated-Newton algorithm using the OPLS3 parameter set [132] and a surface Generalized Born implicit solvent model. The ranking of the obtained complexes is implemented according to Prime calculated energy (molecular mechanics and solvation), and the complexes within 30 kcal/mol of the minimum energy structure are used in the last step of the process, redocking with Glide (Schrödinger, LLC, New York, NY, 2016) using standard precision, and scoring. Finally, the ligands used in the first docking step are re-docked into each of the receptor structures retained from the Prime refinement step. The final ranking of the complexes is performed by a composite composite score which takes into account the receptor-ligand interaction energy (GlideScore) and receptor strain and solvation energies (Prime energy).

Molecular Dynamics (MD). For the construction of the protein-ligand complex CHARMM-GUI platform was used. Each model was inserted into a heterogeneous fully hydrated bilayer 120 Å x 120 Å x 120 Å, consisting of DPPC lipids and ergosterol. The membrane embedded system was solvated with TIP3P water molecules, neutralizing with counter ions, and adding 150 mM Na^+ and Cl^-. CHARMM36m [133] force field was used for protein and lipids, while the ligand was prepared using Antechamber [134] and the general Amber force field [135]. The protein orientation into the membrane was calculated using the PPM server [136]. The assembled simulation system consisted of ~130,000 atoms. The systems were simulated using GROMACS software [137]. The models were minimized and equilibrated to obtain stable structures. Minimization was carried out for 2,000 steps with a step size of 0.001 kJ/mol applying a steepest descent followed by a conjugate gradient algorithm, and the system was equilibrated for 20 ns by gradually heating and releasing the restraints to expedite stabilization. Finally, the system was further simulated free of restraints at a constant temperature of 300 K for 100 ns, using Nose-Hoover thermostat [138] and Parrinello-Rahman semi-isotropic pressure coupling [139] and compressibility at 4.5e-5 bar-1. The van der Waals and electrostatic interactions were smoothly switched off at 1.2 nm by switching function, while long-range electrostatic interactions were calculated using the particle mesh Ewald method [140].

2.11 Molecular simulations shown in chapter 3.3 and 3.4

UapA (wild-type or mutated when discussed) or rSNBT1 homolog dimers were inserted into a lipid bilayer using the CHARMM-GUI tool ([141], http://www.charmmgui.org). The resulting system was explicitly solvated using the transferable intermolecular potential with 3 points (TIP3P) water model [142] and neutralized by the addition of Na^+ and Cl^- counter ions at concentrations of 0.15 cM. The lipid bilayer utilized was composed of 20% ergosterol, 9% POPC, 12% DYPC, 9% YOPC, 6% POPE, 3% DYPE, 5% YOPE, 3% DOPE, 19% POPI, and 14% PYPI, as described previously (http://www.charmm-gui.org/?doc=archive&lib=lipid, [12]). All UapA mutations were constructed by utilizing the CHARMM-GUI's initial step "PDB Manipulation Options." The N-terminal residues were always methylated and the C-terminal residues were always amidated. MD simulations were performed with GROMACS 2018 [137] using the all-atom force field CHARMM36 [143]. Periodic boundary conditions were used. Long-range electrostatic interactions were treated with the Particle Mesh Ewald method. Nonbonded interactions were described with a Lennard-Jones potential, with a cut-off distance of 1 nm and an integration step of two femtoseconds (fs) was implemented. The system was progressively minimized and equilibrated using the GROMACS input scripts generated by CHARMM-GUI, and the temperature and pressure were held at 303.15 Kelvin and 1 bar, respectively [144]. The resulting equilibrated structures were then used as an initial condition for the production runs of 100 ns with all the constraints turned off. Production runs were subsequently analyzed using GROMACS tools, and all images and videos were prepared using VMD software [145].

For the *in silico* mutation of I157F, L192F, and L431F on UapA we started from the crystal structure of UapA and mutations were performed using the "mutation" command on Maestro v11.5 (Schrödinger Release 2018-1). Each resulting structure was inserted to Protein Preparation Wizard Workflow as implemented on Maestrov11.5. Restrained minimization was converged when heavy atom root-mean-square deviation (RMSD) was >1 Å.

Homology modeling of rSNBT1. The construction of a structural model of rSNBT1 was based on the crystal structure of UapA in the inward-open conformation [Protein Data Bank (PDB) entry 5I6C]. The final model was built using PRIME software with an energy-based algorithm [146]. A loop refinement routine was also implemented.

Induced-fit docking of uracil on rSNBT1. Protein preparation using OPLS2005 force field [147] and molecular docking was performed with the Schrödinger Suite 2018. After

protein structure alignment with the crystal structure of UapA (PDB 5I6C), the binding pocket was defined by residues F124, E347, Y395, S396, and E397. Uracil was docked on the final structure from Homology Modeling, using the induced-fit docking protocol (Schrödinger Release 2018-1: Schrödinger Suite 2018-1 Induced Fit Docking protocol; Glide, Schrödinger, LLC, New York, NY, 2018; Prime, Schrödinger, LLC, New York, NY, 2018), which is intended to circumvent the inflexible binding site, and accounts for the side chain and backbone movements upon ligand binding [148].

Results

3.1 Evolution of substrate specificity in the Nucleobase Ascorbate Transporter (NAT) protein family

Adapted from Kourkoulou, A., Pittis, A. A. & Diallinas, G. ***Microb. Cell*** *5, 280–292 (2018).*

3.1.1 Rationale of the study

As mentioned before, transporters that belong to the NAT family are highly specific for either nucleobases or L-ascorbate. More than 24 members of the NAT family have been functionally characterized via direct assessment of their transport activities or indirectly via genetic and physiological studies [1,5,7,97,149]. Bacteria, fungi, plant and mammals (i.e. rSNBT1) have nucleobase specific homologues, whereas L-ascorbate specific members are only described in mammals (e.g. SVCT1/2). Nucleobases are planar molecules with relatively low solubility, whereas L-ascorbate is non-planar and extremely soluble. Such a difference should necessitate prominent changes in the architecture of the substrate-binding site and the substrate translocation trajectory in the relevant NATs. In addition, a distinct homologue clade of NATs, identified in vertebrates (SVCT3) seems unrelated to either nucleobase or L-ascorbate transport. In other words, there are at least two major specificity shifts in NATs. We decided to investigate the molecular aspects of NAT substrate specificity and address the evolution of ascorbate transporters.

3.1.2 L-ascorbate transporters belong to a phylogenetically distinct NAT clade

We performed an extensive phylogenetic analysis using different datasets and methodologies. We chose NAT sequences that, with very few exceptions, were made of 414-650 amino acid residues and we tested selected representatives from all major taxa with the HHpred modelling program (https://toolkit.tuebingen.mpg.de) which showed that all conformed to the 7+7-helix inverted repeat structure of UapA (analysis not shown). Firstly, we reconstructed a phylogenetic tree using identified NAT homologues in all domains of life. Figure 3.1, 3.2 and 3.3 highlight our major findings in respect to the distribution of NAT sequences. Several gene losses were identified in mesozoa, various pathogenic protozoa such as the genera Trypanosoma, Plasmodium, Leishmania and Giardia, in many basal fungi including Microsporidia, Cryptomycota and Blastocladiomycota, in several Saccharomycetales (after the

whole genome duplication), in basal Metazoa (e.g. Porifera), in some Ctenophora, in various jaw worms (Gnathostomulida and Xenacoelomorpha) and in jawless fish, lampreys and hagfish. The first evolutionary split was between homologues from archaea, bacteria, algae and fungi (UapA-like group) and NATs from plants and metazoa (Figure 3.1a). In the UapA-like group, archaea, algae and fungi form distinct sub-clades, while bacterial homologues were clustered based on specificity. *Dictyostelium discoideum* was also clustered in a bacterial clade suggesting a possible horizontal gene transfer (Figure 3.2). The second evolutionary split was between homologues from plants and metazoa. Plants form a clear monophyletic group including bryophytes, monocots, dicots and some green algae (Figure 3.1a and 3.3). Metazoan NATs were distributed in several distinct branches that can be grouped in four major branches, named as SVCT1 (Slc23a1), SVCT2 (Slc23a2), SVCT3 (Slc23a3) and SVCT4 (Slc23a4) based on the terminology of the mammalian homologues.

To shed more light into the evolution of metazoan NATs, we deepened our taxon sampling and we used HMMER [150] and the NAT Pfam domain (PF00860) to extract 728 NAT sequences from 171 plant and metazoan fully sequenced genomes (see methods). After reconstructing a phylogenetic tree using these sequences, we detected and extracted only the 674 sequences of metazoan NATs (Figure 3.1a and 3.3). Most metazoa (early diverging animals such as Placozoa, several Cnidaria and flatworms, non-vertebrates and most vertebrates) include one or more Slc23a4 homologues that are clustered into an “extended” SVCT4 group (>45% identity among its members). Most species include a single Slc23a4 homologue, but there is also significant genus-specific expansion of the group, apparently by gene duplications (e.g. 6-14 paralogues in Caenorhabditis, 8 in Ciona intestinalis, 5-6 in some fish; Figure 3.1b). The only member of this group that is functionally characterized is the nucleobase transporter from rat, rSNBT1 [4]. Based on the conservation of specific residues in the NAT signature motif, that is critical for substrate specificity, present in other nucleobase-specific NAT members from bacteria fungi and plants, we could hypothesize that the other members of this group are also specific for nucleobase-related substrates. Moreover, all Haplorhini (anthropoid primates or simians and the tarsier *Carlito syrichta*) have lost functional Slc23a4 homologues, due to gene truncations (i.e. generation of pseudogenes; [5]). Slc23a4 losses were also detected in other species corresponding to genus-specific cases rather than loss in major groups (Figure 3.1b).

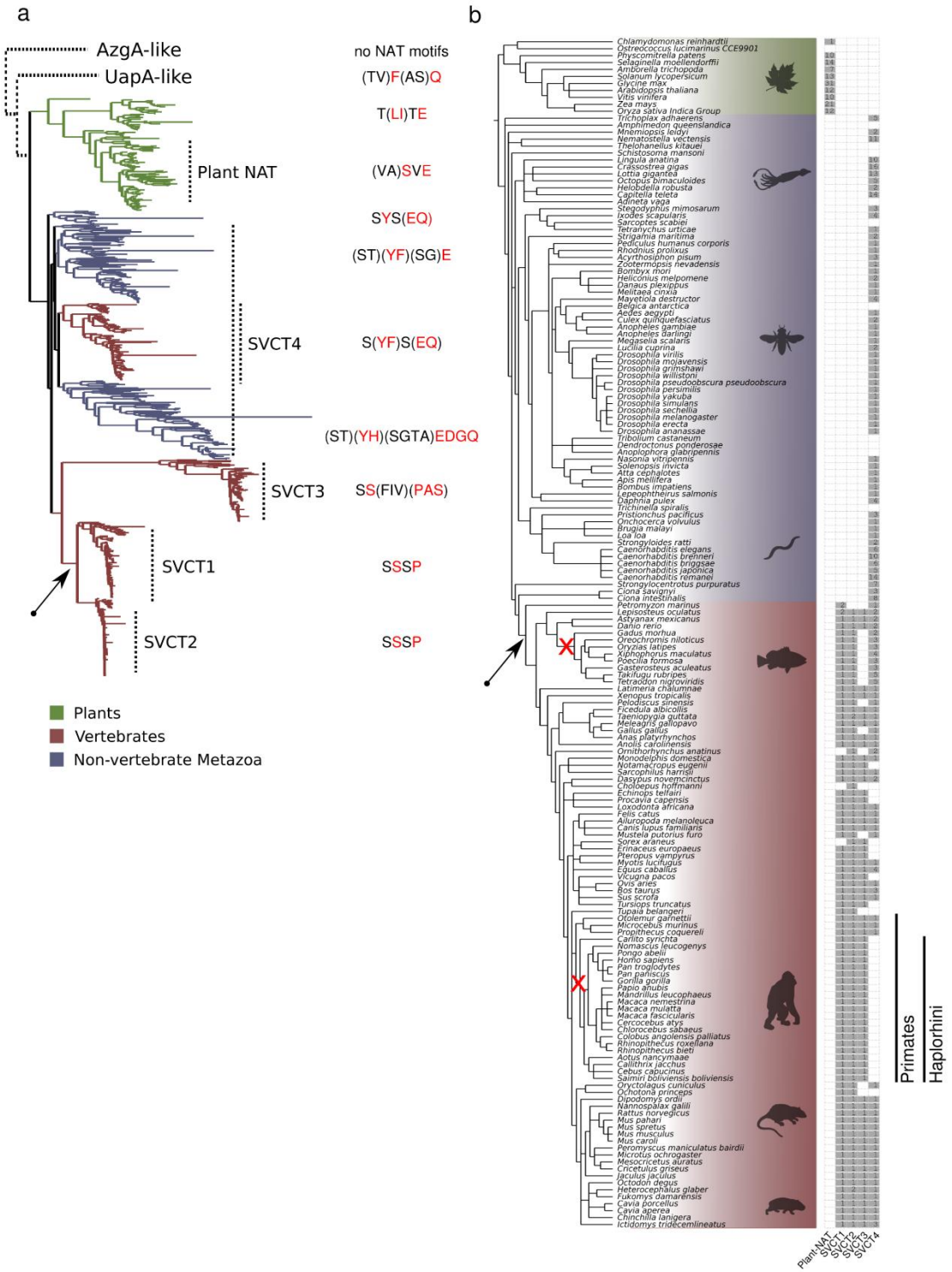


Figure 3.1 Evolution of the SVCT clade in NATs. (A) ML phylogenetic tree of the SVCT clade of NATs and the distinct NAT signature motifs. The different SVCT plant and animal NAT subfamilies, as well as a variable consensus of the four first amino acids in the NAT signature motif, are indicated next to the tree. The tree was midpoint rooted, as shown, and the other two clades, UapA-like and AzgA-like, are placed schematically above with dotted lines. The UapA-like group corresponds to canonical microbial NATs and is shown in more detail in Figure 3.3. The AzgA-group is used as an outgroup of structurally homologous proteins, which however do not conserve the functional motifs of canonical NATs. For the complete version of the plant and animal NATs see also Supplementary Figure S2. (B) Phylogenetic profile of the SVCT clade subfamilies for the 171 plant and animal species. The frequency of members in each subfamily is shown next to the NCBI taxonomy tree structure of the 171 species that were searched for NAT domains. The sub-families were defined based on (A) and in reference to the known

mammalian transporters (SVCT1-4). Two early losses in two lineages, the SVCT3 in the Teleostean clade Eurypterygii and SVCT4 in the primate suborder Haplorhini, are indicated in the corresponding branches with a red "x". The emergence of the L-ascorbic acid transport specificity is indicated with a black arrow both in the phylogenetic tree (A), as well as in the taxonomy tree (B). The colors (green, red, blue) in both (A) and (B) are as indicated in the figure legend. Both trees were visualized using the ETE toolkit [131].

Apart from Slc23a4, vertebrates also include a SVCT1-like and a SVCT2-like homologue (~50% identity with Slc23a4 in the same species) and these are clustered in a separate clade that includes the well-characterized mammalian L-ascorbate transporters. Most species maintain one member of SVCT1 and SVCT2 (Figure 3.1b). Thus, these two sub-families probably emerged after a duplication event of a NAT transporter in a vertebrate ancestor. Due to the conservation of the NAT signature motif present in L-ascorbate transporters and the high similarity found in this clade (>68%) together with the fact that all functionally characterized homologues of this clade are L-ascorbate transporters, we can hypothesize that it consists of L-ascorbate transporters and thus L-ascorbate specificity evolved in the vertebrate ancestor and was retained in fish and tetrapods. However, whether all members retain L-ascorbate specificity, or whether some homologues can also transport nucleobases or other substrates, is not formally known. Most vertebrates also include SVCT3-like sequences (Figure 3.1b). The only studied transporter of this clade, hSVCT3 of human, does not transport nucleobases or L-ascorbate, thus these transporters remain of unknown specificity [5]. Noticeably, SVCT3 has been lost in some lineages including the Acanthomorpha, a very diverse clade of teleost fishes. The SVCT3 clade is the most divergent NAT group in the metazoan clade, as it shows the lowest sequence similarity with the other clades (< 34% identity with SVCT1/2 and SVCT4), and does not conserve the TMS8 motif present in "canonical" NATs necessary for function and specificity. This might explain the unstable position of this clade using different algorithms and parameters (not shown).

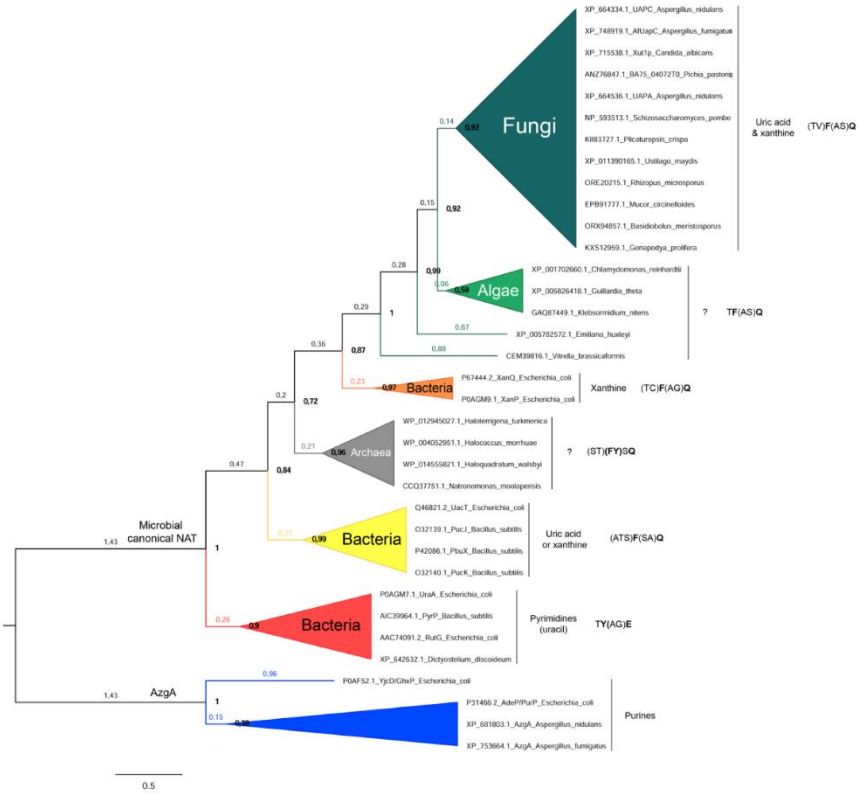


Figure 3.2. Phylogeny of microbial NATs. The tree shown includes all functionally characterized NATs from bacteria and fungi and their closest homologues from other prokaryotic and eukaryotic microorganisms, obtained by blastP (https://blast.ncbi.nlm.nih.gov). The AzgA sequences at the bottom of the tree are functionally characterized purine transporters of the so-called AzgA-like family, which includes proteins structurally similar with NATs, but show very little primary sequence similarity and lack NAT-specific functional motif (closest out-group). Multiple sequence alignments was built from the selected NAT sequences (accession number shown in the figure) using MUSCLE v7.0.26 (http://www.megasoftware.net/). MEGA was also used for testing the aligned sequences for optimal amino acid substitution model. According to the AIC, the LG+G+F model [151] was selected and a tree was created using a maximum likelihood (ML), and visualized by FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). Bootstrap values are shown in bold at the nodes of the clades. Evolutionary distances are also shown on the branches of the tree. The variable, four-amino acid, part of the NAT signature motif is shown at the right side of the tree (in bold the residues critical for specificity, see main text). Note that the microbial NAT sequences shown in this tree are collapsed and referred as "UapA-like sequences" in Figure 3.1.

3.1.3 The NAT signature motif is differentially conserved in nucleobase and L-ascorbate transporter clades

The phylogenetic analysis performed herein confirmed that, except from the SVCT3 group, the motifs in TMS1 and TMS8 are highly conserved in all NATs. Moreover, it showed that the critical NAT signature motif (TMS10) exists in two major versions especially considering three additional upstream residues (Fig 3.1a and 3.3). In the UapA-like, plant and SVCT4 clade which all include nucleobase-specific NATs, the motif conforms to the sequence Thr/Ser[1]-Phe/Tyr[2]-

Ala/Ser^{3}-Gln/Glu^{4}-Asn^{5}-X-Gly^{7}-X-X-X-X-Thr^{12}-Gly/Arg/Lys^{13} (numbers shown are arbitrary). In the SVCT1/2 clade that contains the L-ascorbate specific members the motif is Ser^{1}-Ser^{2}-Ser^{3}-Pro^{4}-Asn^{5}-X-Gly^{7}-X-X-X-X-Thr^{12}-Arg^{13}. The most prominent changes between the two versions concern the second and the fourth residues where nucleobase specific members have an aromatic (Phe/Tyr) and a polar (Gln/Glu) residue respectively whereas L-ascorbate specific NATs conserve a Ser and a Pro residue in these positions. Most other conserved residues in the NAT signature motif (Thr/Ser^{1}, Ala/Ser^{3}, Asn^{5}, Gly^{7}, Thr^{12} and Gly/Arg/Lys^{13}) are known to contribute to transport kinetic parameters rather than substrate specificity or are not differently conserved depending on specificity [8].

Previous structural studies confirmed the importance of these two residues in substrate binding as in UapA Phe^{2} interacts with the nucleobase ring of xanthine via a pi-pi stacking interaction while Gln^{4} makes a strong bipolar interaction [2]. Similarly, in UraA the phenyl ring of Tyr^{2} is parallel to the pyrimidine ring of uracil and Glu^{4} anchors uracil by making a hydrogen bond with it [3,6]. Thus, substitution of Phe/Tyr and Gln/Glu with a Ser and a Pro respectively, might contribute to the generation of L-ascorbate transporters. It should be noticed that no mutation up to date converted a nucleobase specific NAT into a transporter capable of recognizing L-ascorbate. Importantly, substrates interact also with the Glu of the TMS8 motif that is conserved in both nucleobase and L-ascorbate transporters, which further indicates the possible critical role of the NAT signature motif in this specificity shift.

3.1.4 The differentially conserved residues in the NAT signature motif are essential for function and specificity

There have been no genetic screens, so far, to select a change of NAT specificity from nucleobases to L-ascorbate. This is mainly due to the fact that L-ascorbate cannot be used as a C or N source to select mutants importing and growing on this substrate. Notice that, L-ascorbic acid is not taken up by *A. nidulans* cells as there is neither a specific endogenous transporter, nor evidence for passive diffusion (G. Diallinas, unpublished observations). In an effort to shed more light on the evolution of NATs and test whether the NAT signature motif has a critical role to this change of specificity we rationally designed specific variations of the motif. We constructed and functionally analyzed the following UapA mutants according to the NAT motif found in L-ascorbate specific NATs (SSSP): Q408P [11], A407S/Q408P, T405S/F406S/A407S and F406S/A407S/Q408P. In other words, we modified the first part of the NAT signature motif from TFAQ (wild-type UapA) to TFA**P**, TF**SP**, **SSS**Q or T**SSP** (changes in

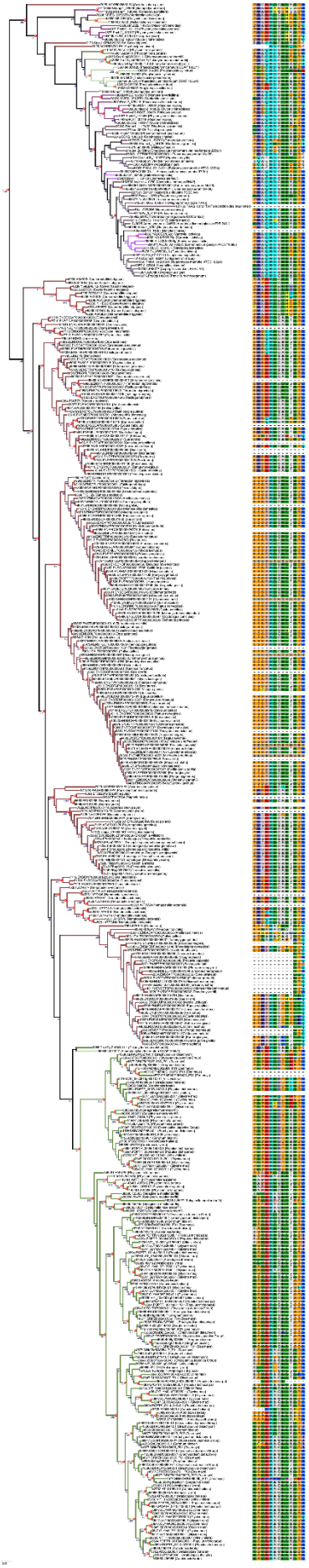


Figure 3.3. Phylogenetic tree of the SVCT clade and the NAT signature motif. The same ML phylogenetic tree as in Figure 3.1 is shown, here extended and with the sequence - followed by the species - names shown. The protein names correspond to the sequence ids as of Ensembl genomes, or the Uniprot name, for those cases that mapping was successful. Red "spheres" indicate duplication nodes (according to the species overlap algorithm), and the numbers next to the nodes correspond to aLRT support values. On the right, the NAT signature motif is shown for each protein sequence. The ETE toolkit was used for tree visualization [130].

bold). We expressed these alleles in a background genetically lacking the three major endogenous nucleobase transporters UapA, UapC and AzgA (ΔACZ strain, referred as uapA-), via the native *uapA* promoter. UapC is a paralogue of UapA exhibiting similar specificity for xanthine and uric acid, albeit distinct transporter kinetics (i.e. lower affinity for uric acid compared to xanthine, and lower transport capacity in general;[87]). AzgA is a high-affinity hypoxanthine/adenine/guanine transporter [152]. Finally, all mutants analyzed were made in a fully functional UapA protein tagged with GFP.

We functionally analyzed these versions in respect to their capacity to transport or bind nucleobases or L-ascorbate, and for proper plasma membrane sorting and stability. Purines can be used as N sources, therefore their accumulation by UapA mutants can be scored directly in growth tests. Thus, a control strain expressing a functional UapA (i.e. *uapA*$^{+}$), in a ΔACZ background grows on uric acid or xanthine, but not on other purines. Figure 3.4a shows that all UapA mutants tested lost (T405S/F406S/A407S and F406S/A407S/Q408P) or had dramatically reduced (e.g. Q408P and A407S/Q408P) capacity for growth on UapA physiological substrates, uric acid and xanthine. Additionally, none of the mutants acquired capacity for growth on hypoxanthine (Figure 3.4a) or adenine (not shown). *In vivo* epifluorescence microscopy showed that all mutant versions were properly and stably localized at the plasma membrane (right panel in Figure 3.4a), which in turn implies that mutations do not result in problematic sorting to the plasma membrane but affect transport function per se. The above results were also confirmed by direct uptake assays measuring the transport rate of radiolabeled xanthine (Figure 3.4b; [87]).

As L-ascorbate cannot be used as C or N source, we tested whether the mutations introduced result to any growth phenotype when L-ascorbic acid was added in the medium. Figure 3.4c shows that none of the UapA mutations analyzed, or the total genetic absence of UapA (*uapA*$^{-}$) alters the growth phenotype of *A. nidulans* on 0.1% L-ascorbic acid (5.7 mM). We also performed HPLC or biochemical assays to detect cytoplasmic accumulation of externally supplied L-ascorbic acid in the mutants or a wild-type strain. In all cases, there was no evidence of accumulation and thus transport by any UapA mutant (results not shown).

We also tested whether the presence of L-ascorbate induced substrate-elicited endocytosis as an indirect measure of L-ascorbate recognition by wild-type UapA and mutants. This *in vivo* cellular assay is based on the fact that UapA physiological substrates trigger endocytic turnover and the appearance of early endosome and/or vacuoles due to sorting into these organelles of GFP-tagged UapA. Thus, based on the fact that all mutants

showed stable and proper localization to the plasma membrane, we tested whether L-ascorbate promotes UapA endocytosis, which would be an indication that this metabolite is recognized by wild-type UapA or its mutant versions [153]. As a negative control we used a strain expressing a transporter unrelated to NATs, namely FurA tagged with GFP. FurA is a highly specific allantoin transporter of *A. nidulans* that belongs to the NCS1/FUR family that does not recognize either nucleobase or L-ascorbate [43].

Figure 3.4d shows that after 15 min, the presence of 0.1% L-ascorbate, resulted in sorting of UapA mutants into endosome-like structures, but has no similar effect on wild-type UapA or FurA. This effect became more prominent after prolonged presence of L-ascorbate (1h), mostly evident in T405S/F406S/A407S and F406S/A407S/Q408P, but also became apparent, albeit at a much lower level, in wild-type UapA. Importantly, in all conditions tested no endosome-like structures were detected in the strain expressing FurA-GFP. Additionally, after prolonged incubation with L-ascorbate the morphology of some hyphae altered to a phenotype often seen under severe cell stress.

The changes include increase of the number and the size of vacuoles and hyphal width reduction along with total loss of fluorescence. The fraction of hyphae with this phenotype was dependent on the number of mutations introduced as the most prominent effect was observed in the triple mutants F406S/A407S/Q408P (in 55% hyphae) and T405S/F406S/A407S (in 68% hyphae) (Figure 3.4d and 3.4e). A similar but reduced effect was seen in A407S/Q408P (40%) and Q408P (27%) (Figure 3.4d and 3.4e). Notably, this morphological effect caused by L-ascorbate was much lower in strains expressing wild-type UapA or FurA. Thus, both the presence of endocytic-like structures and the L-ascorbate-dependent morphological effect were UapA specific. Moreover, as these effects were more visible in F406S/A407S/Q408P and T405S/F406S/A407S, we can hypothesize that the presence of the tandem Ser residues within the NAT substrate binding site and the removal of the aromatic Phe residue, could favor L-ascorbate recognition.

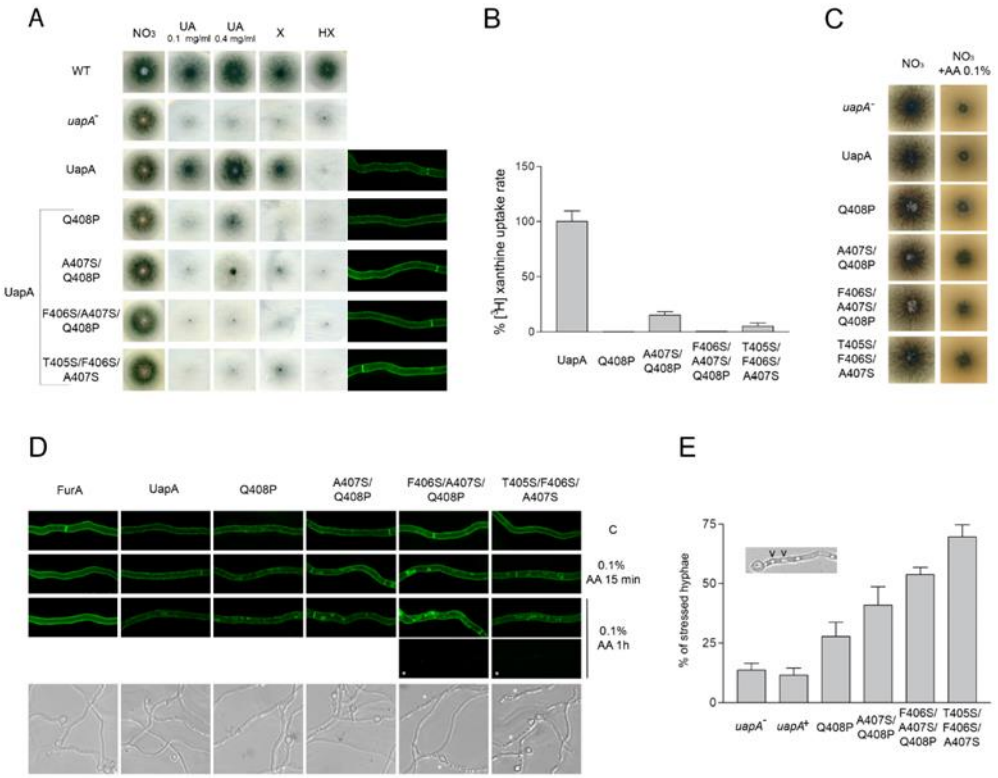


Figure 3.4. Specific residues within the NAT signature motif are critical for UapA function and specificity. (A) Growth tests, on standard MM, of isogenic strains expressing wild-type (UapA) or mutant UapA versions (Q408P, A407S/Q408P, T405S/F406S/A407S and F406S/A407S/Q408P), expressed in a genetic background genetically lacking other major purine transporters (i.e. ΔuapC ΔazgA). A strain lacking a functional UapA (uapA-) and a standard wild-type A. nidulans strain (WT) are also included in the test as controls. The test depicts growth on 10mM sodium nitrate (NO_3) or 0.1 mg/ml uric acid (UA), xanthine (X), hypoxanthine (HX), after 48 h at 37°C. A 4-fold higher concentration of UA (0.4 mg/ml) was also used in order to distinguish UapA mutants with highly reduced affinity for substrates from mutants with an apparent total loss of transport activity. Subcellular localization of wild-type or mutant UapA-GFP versions, as revealed by *in vivo* epifluorescence microscopy, is also shown at the right side of the panel. For details of sample preparation see Material and methods. (B) Comparative UapA-mediated uptake rates of radiolabeled ^{3}H-xanthine in strains ex-pressing wild-type (UapA) or mutant versions (Q408P, A407S/Q408P, T405S/F406S/A407S and F406S/A407S/Q408P) of UapA-GFP. Results are averages of three measurements for each concentration point. SD was < 20%. (C) Growth tests, at 37°C, on standard MM supplemented with 10 mM sodium nitrate (NO3) as nitrogen source in the absence or presence of 0.1% L-ascorbate (AA) of isogenic strains lacking UapA (uapA-), or expressing wild-type (UapA) or mutant UapA versions. (D) Upper panel: *In vivo* epifluorescence microscopy following the effect of L-ascorbate (0.1% AA) on UapA-GFP or FurA-GFP subcellular localization. 1 h in AA triggers "loss" of GFP fluorescence in a significant fraction (> 50%) of hyphae in the triple mutants T405S/F406S/A407S and F406S/A407S/Q408P (lowest black boxes marked with a red asterisk). The same effect, albeit significantly reduced, is also seen in A407S/Q408P (36%) and Q408P (22%), while in the wild-type UapA-GFP or FurA-GFP fluorescence loss is < 10% (not shown). Lower panel: L-ascorbate (1% AA for 1 h) triggers differential hyper-vacuolarization and reduction of hyphal width (i.e. cell stress) in UapA mutants and controls. Stressed hyphae are marked with a red asterisk. (E) Quantification of the effect of L-ascorbate on hyphal morphology, as recorded for 100 hyphae, of each strain analyzed. In all cases, hyper-vacuolarization and reduction of the width of hyphae was associated with loss of fluorescence (not shown). The results depicted in the graph confirm the apparent cytotoxicity of L-ascorbate in the triple mutants T405S/F406S/A407S and F406S/A407S/Q408P, where 50-70% of hyphae seem stressed, followed by progressively lower effects on A407S/Q408P (40%), Q408P (27%) and wt UapA (11%).

3.1.5 UapA can bind L-ascorbate with extremely low affinity: evidence for evolution of L-ascorbate transporters via optimization of a NAT sub-function

The observation that the prolonged presence of L-ascorbate can trigger a degree of wild-type UapA-GFP endocytic turnover made us re-examine whether UapA can recognize L-ascorbate. To test this possibility, we performed transport competition assays using increasing concentrations of L-ascorbate (1-60 mM). Figure 3.5 shows that L-ascorbate inhibits the uptake of radiolabeled ^{3}H-xanthine (1 μM) with a K_i value of ~17 mM. Thus, wild-type UapA can indeed bind L-ascorbate, but with very low affinity. We repeated the same experiment using a UapA mutant that recognizes and transports all nucleobases (R481G/T528M)[9] and we obtained a moderately increased binding affinity for L-ascorbate (K_i ~ 11 mM). This mutant has an intact NAT signature motif, but includes two mutations in the outward and inward gating elements that seem to loosen UapA specificity [2,9]. So, UapA can bind L-ascorbate with an affinity approx. 2500-fold lower than the one for its physiological substrates (xanthine and uric acid), or and 10- to 100-lower than its affinity for other nucleobases [8]. This finding supports the hypothesis that also other nucleobase-specific NATs might recognize L-ascorbate with very low-affinity and this in turn supports a model where L-ascorbate NATs evolved via a progressive improvement of a sub-function of nucleobase-specific NATs.

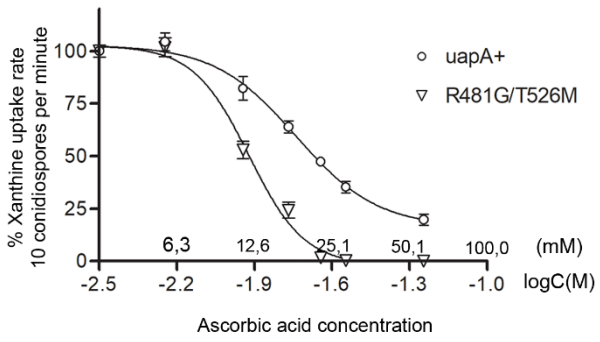


Figure 3.5. L-ascorbate inhibits UapA transport activity. The graph shows dose response curves of [^{3}H]-xanthine uptake rate in strains expressing either wild-type UapA or the isogenic mutant UapA-R481G/T526M, in the presence of increasing concentration of non-radiolabeled L-ascorbate (0-56 mM). Results are averages of three measurements for each concentration point. SD was < 18%. For de-tails of measurements and *Ki* estimation see Materials and methods.

3.1.6 Conclusions

Gene duplication and divergence is a major mechanism of evolutionary novelty, providing protein families with an expanded repertoire of activities and novel specificities. This analysis

supports that a duplication of a NAT ancestral vertebrate transporter gave rise to the L-ascorbate transporters (NAT neo-functionalization). In the course of evolution one copy retained nucleobase specificity while the other gained via a progressive improvement of a sub-function (i.e. very low affinity binding of L-ascorbate) L-ascorbate specificity. Finally, the NAT signature motif that is critical for substrate recognition probably plays a determinative role in this dramatic specificity shift.

3.2 Context-dependent cryptic roles of specific residues in substrate selectivity of the UapA purine transporter

Adapted from Kourkoulou, A., Zantza, I., Foti, K., Mikros, E. & Diallinas, G. ***J. Mol. Biol. (2021).***

3.2.1 Rationale of the study

As mentioned previously, a prominent feature of the NAT family members is a motif in TMS10, historically referred as NAT signature motif [8,92]. This includes residues critical for substrate binding and specificity (Gln^{408}, Arg^{417} as well as Phe^{406} which contributes moderately to specificity) or transport catalysis (Asn^{409}, Gly^{411} and Thr^{416}). In the UapA crystal, Gln^{408} forms two hydrogen-bonds (H-bond) via its side chain with xanthine while Ala^{407} forms a H-bond via the nitrogen of the main chain and F406 interacts via a π-π stacking interaction [2]. Asn^{409} and Thr^{416} seem to be involved in a network of dynamic intramolecular interactions with specific residues in TMS1, TMS3 and TMS8. Gly^{411} is a key element in movements associated with UapA-mediated transport as G411V is capable of substrate binding, similar to the wild-type, but lacks detectable transport activity [8]. Finally, Arg^{417} has been shown to be important for uric acid binding affinity.

In the previous chapter, we showed that the NAT motif appears in two major versions depending on specificity. In ascorbate-specific NATs the critical differences are two substitutions in the first part of the motif: the aromatic Phe/Tyr residue (Phe^{406} in UapA) is substituted by Ser, while the polar residue Gln/Glu (Gln^{408} in UapA) is replaced by a Pro. Mutant versions of UapA introducing residues present in the L-ascorbate transporters proved to have practically no transport activity for either purines or L-ascorbate. Thus, possibly other residues, outside the major substrate binding site, also determine specificity. Indeed, most specificity mutations in UapA, selected by direct genetic screens, concern residues that map outside the major substrate binding site and enlarge UapA specificity so that it can transport,

besides xanthine and uric acid, other purines and uracil. The most interesting one, Arg^{481}, lies in close proximity to the xanthine-binding site of the opposite protomer [2]. Other examples are Thr^{526} and Phe^{528} that locate in the interface between the dimerization and core domains near to the substrate-binding site. Mutations in these residues conferred UapA-mediated growth on hypoxanthine and adenine showing no or very little measurable affinity for these purines [9,10,154]. Here we developed a novel strategy to analyze specificity of variable NAT signature motif versions via mutational analysis on UapA and we run Molecular Dynamics to shed some light into questions and conclusions drawn from the functional analysis of the mutants.

3.2.2 Context-depended role of the NAT signature motif in determining UapA specificity

In order to explore an extended specificity profile of the various NAT signature motif versions we used, for the first time (regarding UapA), a different genetic system. In this system we did not use the previously described ΔACZ strain which lacks three nucleobase transporters (UapA, UapC and AzgA), but a Δ7 strain which genetically lacks all seven major endogenous nucleobase-related transporters (UapA, UapC, AzgA, FcyB, FurD, FurA and CntA) and has no detectable transport activity for purines, pyrimidines, nucleosides, allantoin and nucleobase toxic analogues [155]. FcyB is a high-affinity cytosine transporter that also contributes to purine uptake [13,156], FurD and FurA are high-affinity uracil and allantoin transporters [43,157] and CntA is the single general nucleoside transporter [158]. Thus, any UapA mutant introduced by genetic transformation in this strain could be analyzed for its ability to transport any purine-related substrate in a 'clean' background. Moreover we used the strong *gpdAp* promoter to drive the transcription of uapA alleles so that UapA mutants with too little transport activity (via the native uapA promoter) can be measured. This system has been previously used to reveal cryptic functions of NCS1 transporters [43,155,156]. In all cases *uapA* substitutions were made in a fully functional UapA protein tagged with GFP.

Using the above system, we functionally analyzed seven UapA versions with substitutions in the first four variable residues of the NAT signature motif to examine their role in substrate recognition and transport. These are: UapA-T**Y**AQ, -TFA**E**, -T**Y**A**E**, -**SYS**Q, -**SYSE**, -**SSS**Q and -**SSSP**, where in bold are substitutions corresponding to the TFAQ (Thr^{405}-Phe^{406}-Ala^{407}-Gln^{408}) motif present in the wild-type UapA. Among them, UapA-TFA**E**, -T**Y**AQ

and -**SSS**Q have been previously analyzed via expression from the uapA native promoter and in the ΔACZ background that did not permit testing of nucleobase-related substrates other than purines [8,9](and Figure 3.4). Figure 3.6a shows the variable NAT signature motif sequences present in different functionally characterized NATs (except from SVCT3 which remains of unknown function). **SYSE** is the sequence found in many metazoan NATs that belong to the SVCT4-clade, including rSNBT1, the only characterized nucleobase-specific metazoan NAT (Figure 3.1a, 3.3 and 3.6a). **SYS**Q is also a sequence found commonly in metazoan SVCT4 NATs (Figure 3.1a and 3.3), while **SSSP** sequence is found only in ascorbate-specific NATs (Figure 3.1a, 3.3. and 3.6a). The other single, double and triple mutants were used to examine the individual role of the corresponding residues. Notice that fungal and bacterial NATs that transport uric acid and xanthine possess at the critical positions 406 and 408, a Phe and a Gln residue respectively, while bacterial uracil transporters have a Tyr and a Glu at these positions (Figure 3.2 and Figure 3.6a).

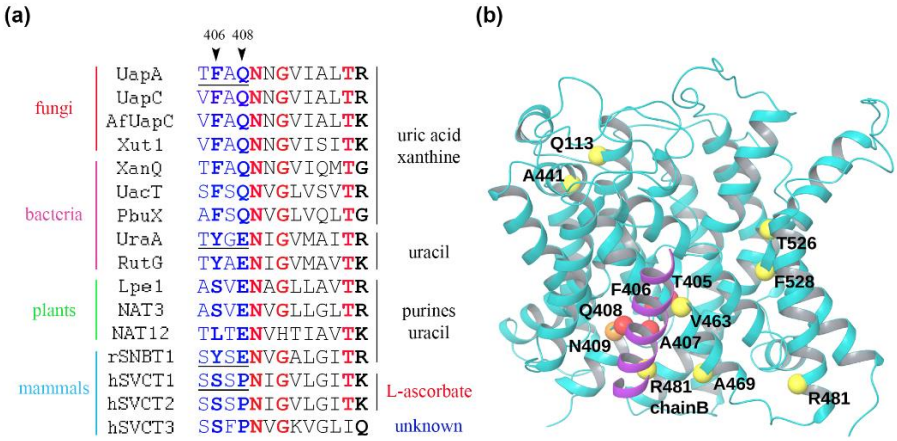


Figure 3.6. Sequence of the NAT signature motif and topological model of UapA highlighting residues critical for specificity. (A) Alignment of UapA, UapC, AfUapC, Xut1, XanQ, UacT, PbuX, UraA, RutG, Lep1, NAT3, NAT12, rSNBT1 (SVCT4), SVCT1, SVCT2 and SVCT3. UapC (XP_664334), AfUapC (XP_748919) and Xut1 (XP_715538) are xanthine-uric acid transporters from *A. nidulans*, *Aspergillus fumigatus* and *Candida albicans*. XanQ (P67444), UacT (Q46821) and PbuX (P42086) are xanthine and/or uric acid transporters from *E. coli* and *Bacillus subtilis*. RutG (AAC74091) and UraA (P0AGM7) are uracil transporters from *E. coli*. NAT3 (NP_180219) and NAT12 (NP_850108) are purine-uracil transporters from *Arabidopsis thaliana*. Lpe1 (Q41760) is a maize xanthine-uric acid transporter also showing low affinity for uracil. rSNBT1 (SVCT4/SLC23A4) is a rat purine-uracil transporter. hSVCT1 (Q9UHI7) and hSVCT2 (NP_005107) are human L-ascorbate transporters. hSVCT3 (NP_001138361) is a human transporter of unknown specificity, unrelated to nucleobases, nucleosides or L-ascorbic acid. (B) Topology of the UapA monomeric subunit showing residues affecting (enlarging) UapA specificity: Val^{463}, Ala^{469} located in TMS12, Thr^{526} and Phe^{528} at end of TMS14 (outer gate), Arg^{481} in start of TMS13 (inner gate), and Glu^{113} and Ala^{441} at two flexible helical hinges linking the core and dimerization domains. UapA function as a homodimer (monomers A and B). Notice that Arg^{481} of monomeric subunit B (only part of TMS13 of this subunit shown in purple) affects specificity by regulating the translocation of substrates via the trajectory of the opposite monomeric subunit A. All other specificity mutations are located in the interface of the core and dimerization domain in each monomer, and thus affect the sliding of the core, including the substrate binding site, along the dimerization domain (for details see [2]).

Figure 3.7 summarizes growth tests of mutants and control strains on purines, purine-nucleosides or allantoin that were used as N sources, and on media containing the toxic pyrimidine and pyrimidine-nucleoside analogues 5-fluorouracil (5FU), 5-fluorocytosine (5FC) and 5-fluorouridine (5FUd). As expected, the negative control strain Δ7 does not grow on purines, purine-nucleosides or allantoin and is resistant to 5FU, 5FC and 5FUd, which is nearly the mirror image of the growth phenotype of a standard wild-type *A. nidulans* strain. The positive control strain (Δ7::*uapA*$^{+}$) grows on uric acid (UA) and xanthine (X), but not on other nucleobase-related compounds tested, and is resistant to 5FU, 5FC and 5FUd. T**Y**AQ and TFA**E** mutant versions led to growth on substrates similar to the wild-type. **SYS**Q, besides growth on UapA physiological substrates, conferred growth on high adenine (AD) concentration (2 mM) and very weak growth on hypoxanthine (HX) compared to the negative control (Δ7), while the single substitution T**Y**AQ led to very weak growth on 2 mM of HX and AD. In addition, T**Y**AQ and more prominently **SYS**Q, led to increased sensitivity to 5FU. **SYSE** showed no apparent transport capacity for the purines tested (except from very weak growth on 2 mM AD) but conferred high sensitivity to 5FU. The loss of transport capacity is mainly due to the combination of Y406 and E408, as T**Y**A**E** could not confer growth on any purine tested. T**Y**A**E** also conferred sensitivity to 5FU although to a lower degree. Finally, UapA versions carrying NAT sequences mimicking the motif of L-ascorbic acid transporters showed no apparent capacity for nucleobase transport (i.e. **SSSP**), although the one that conserves Gln^{408} (i.e. **SSS**Q) seemed to confer very weak growth on X and 2 mM UA. Both, however, seem to confer some low level accumulation of 5FU not seen in the wild-type UapA control. None of the mutants conferred growth on allantoin or purine-nucleosides and all were resistant to 5FC and 5FUd. Importantly, the mutations studied did not affect subcellular trafficking as all mutant versions showed proper localization at the plasma membrane (Figure 3.7, lower panel), suggesting that growth reflects UapA activity.

Comparative uptake assays, using radiolabelled xanthine were in line with growth tests as T**Y**AQ and **SYS**Q showed >70% of wild-type UapA transport activity, while all the other mutants with no or dramatically reduced growth, showed very low relative transport (Figure 3.8a). TFA**E** as expected from previous studies [8] exhibited significantly reduced transport activity despite similar to the wild-type growth. The difference between the assays can be explained if we consider that substrate concentration in growth tests is 1000-fold higher than the concentration used in transport assays, and that growth is a result of substrate accumulation and therefore does not indicate transport rates. Next, we performed transport

kinetic assays for the **SYS**Q mutant that conserved considerable transport levels of radiolabeled xanthine.

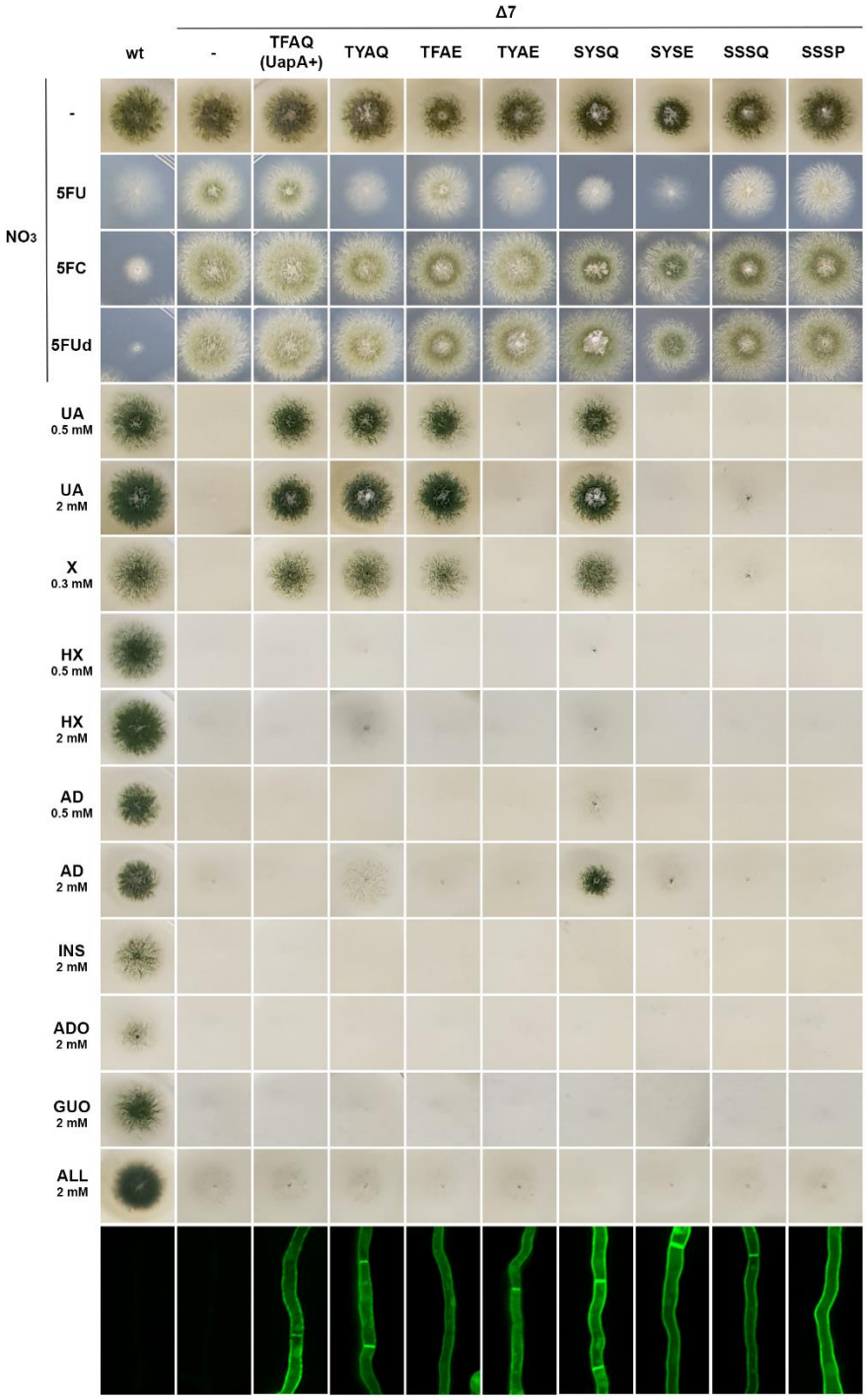


Figure 3.7. Context-depended role in UapA specificity of variably conserved residues in the NAT signature motif. The image shows growth tests of UapA mutants and a control strain (referred by the sequence of their NAT motif, e.g. wild-type UapA is TFAQ, etc.) on toxic nucleobase/nucleoside analogues at 100 μM 5FU, 50 μM 5FC, 10 μM 5FUd (5FU is 5-fluorouracil; 5FC is 5-flurocytosine; 5FUd is 5-fluorouridine), or on different purines (UA is uric acid; X is xanthine, HX is hypoxanthine; AD is adenine), nucleosides (INS is inosine; ADO is adenosine, GUO) is guanosine) or allantoin (ALL), as sole N sources. Concentrations of purine-related substrates used as N sources are in the range of 0.3–2.0 mM.

Toxic analogues are scored in the presence of 10 mM sodium nitrate (NO_3) as nitrogen source. The lower panel shows the subcellular localization of UapA mutants, as evidenced via a GFP tag, in all strains analyzed. All UapA mutant versions shown are analyzed in the isogenic background of Δ7 strain (see text). A standard wild-type (wt) strain possessing all relative endogenous nucleobase-related transporters is also shown for comparison in the first column. Growth tests were performed at 37°C and pH 6.8. Lower panel: Inverted fluorescence microscopy images showing localization of the GFP-tagged UapA constructs. Samples were grown for 18 h at 25°C on MM with NO_3 as N source.

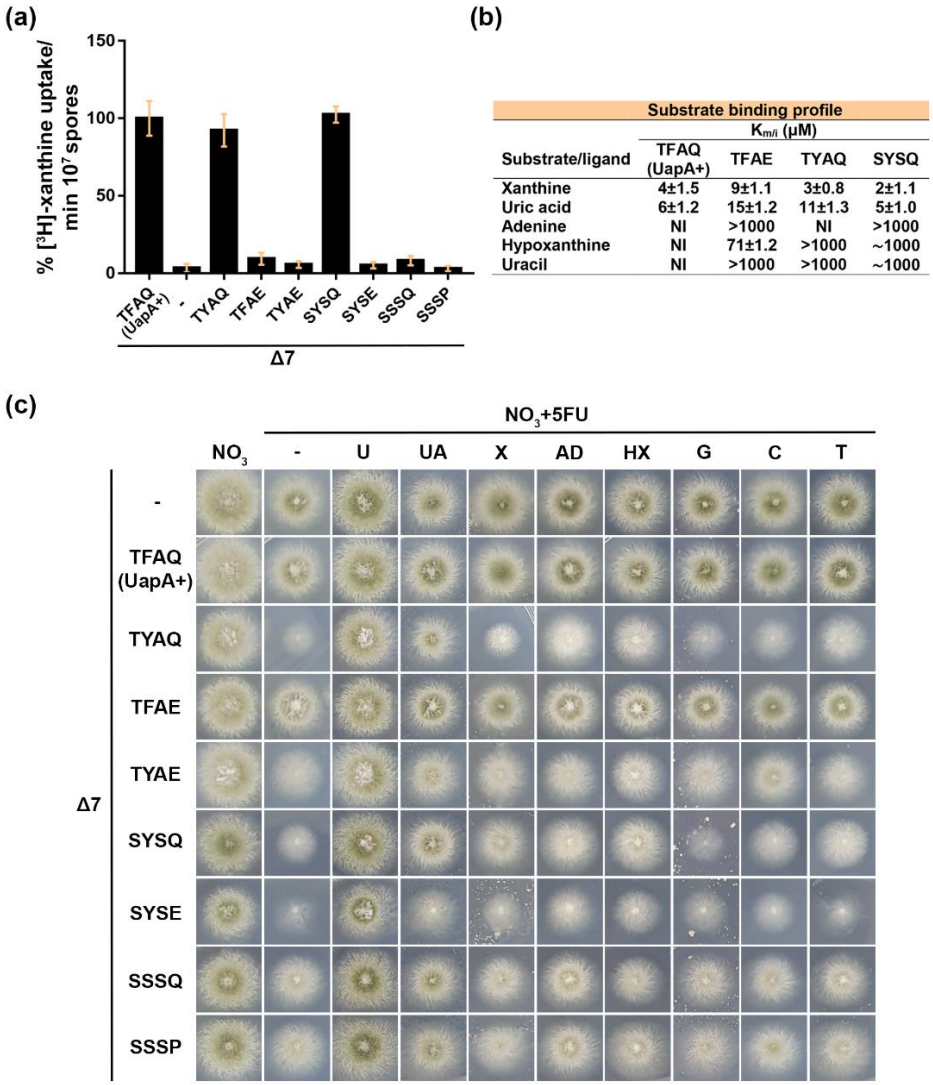


Figure 3.8. Transport activity and specificity of UapA NAT motif mutant versions (A) Comparative transport of 0.1 μM of radiolabelled ^{3}H-xanthine expressed as percentages of initial uptake rates (V) compared to the wild-type (UapA) rate taken as 100%, performed at 37°C (see Materials and methods). Results shown are the average of triplicate measurements from 3 independent assays with a SD shown as error bars. (B) K_m (xanthine) or K_i (uric acid, adenine, hypoxanthine, uracil) values estimated via competitive inhibition assays of radiolabelled xanthine uptake ([87]; see also Materials and methods) in selected mutants. NI stands for no evidence of radiolabeled xanthine transport inhibition in the simultaneous presence of excess 'competitor' nucleobase supplied to up to 1 mM. >1000 means detectable inhibition, but the *Ki* value could not be determined given that nucleobases are little soluble at concentrations >2 mM. Results shown are averages of triplicate measurements from 3 independent assays with SD values shown. (C) *In vivo* competition assays scoring the reversion of 5FU (100 μM) growth inhibition in the presence of excess purines or pyrimidines (2 mM) in the growth medium. Strains and growth conditions are as described in Figure 3.7. U is uracil, G is guanine, C is cytosine and T is thymine.

Data for T**Y**AQ and TFA**E** mutants [9,10] are shown for comparison. Figure 3.8b shows that **SYS**Q did not affect K_m/K_i values for UapA physiological substrates, but gained better K_i values for AD, HX and U, compared to the wild-type. We also performed uptake assays using radiolabelled uracil for the **SYS**Q and **SYSE** mutants that conferred the highest sensitivity to 5FU, but none of the mutants showed measurable accumulation of uracil (see also later). We also performed *in vivo* competition assays that score the ability of excess purines or pyrimidines to compete with 5-FU uptake, and thus suppress 5FU sensitivity (Figure 3.8c). This tests showed that 5FU sensitivity of all mutants was reversed by 2 mM U. **SYS**Q sensitivity was reversed, mostly by UA, X, HX and AD, while in the **SYSE** mutant competition was moderate. For T**Y**AQ and T**Y**A**E** results were quite similar with **SYS**Q, except from the less efficient reversion by xanthine in the T**Y**AQ mutant. Low level accumulation of 5FU by **SSS**Q and **SSSP** mutants was partially reversed by UA. There was no significant competition of guanine (G), cytosine (C) and thymine (T) in all strains, considering the effect of these substrates in the negative control (Δ7).

To summarize, UapA-**SYS**Q seems to be the most efficient and promiscuous UapA version, being able to recognize and transport UA, X, AD and 5FU, and to at least bind with low affinity also HX and U. In contrast, UapA-**SYSE**, which differs only at a single residue with UapA-**SYS**Q (i.e. Q408E), has lost its transport capacity for purines, but acquired higher capacity for 5FU accumulation. The fact that loss of transport capacity is mainly due to the combination of Tyr^{406} and Glu^{408}, while these two mutations alone do not lead to total loss of purine uptake, suggests that there is negative epistatic interaction between these two residues. These results clearly support that UapA specificity is determined by complex interactions of residues in the NAT signature motif that cannot be predicted.

3.2.3 Combination of the SYSE sequence with selected UapA mutations leads to distinct transport profiles

As mentioned before, the **SYSE** sequence in the NAT signature motif is naturally present in several metazoan NATs of the SVCT4 branch (Figure 3.1a and 3.3), including the only characterized member, rSNBT1, that functions as a promiscuous nucleobase transporter [4]. However, this sequence in the context of UapA resulted in loss of transport capacity for purines and acquisition of a capacity for 5FU accumulation. This finding strongly suggests that the residues in the NAT signature motif interact functionally and in a context-dependent manner with other residues in NATs. To address this issue, we combined the **SYSE** sequence

with mutations in residues Arg^{481}, Thr^{526}, and Phe^{528} that are known to enlarge UapA specificity [9,10,154], and may also cooperate with the NAT motif. Figure 3.9a shows growth tests of combinations of **SYSE** with mutations R481G, T526L, F528M, R481G/T526L and R481G/F528M. Notice that F528M is a novel mutation that was designed based on the conservation of a Met residue in this position by metazoan NATs of the SVCT1/2 and SVCT4 branch. As expected mutations in Arg^{481}, Thr^{526}, and Phe^{528} in a wild-type NAT context (i.e. TFAQ instead of **SYSE**) increased UapA promiscuity in respect to purine accumulation, but most of them did not suppress loss of transport capacity by **SYSE**. Only F528M/**SYSE** resulted in detectable growth on purines, especially on high concentrations of adenine. The simultaneous presence of **SYSE** with T526L, F528M or R481G/F528M increased sensitivity to 5FU compared with **SYSE** alone, while R481G led to partial suppression of the ability of **SYSE** to accumulate 5FU. Also, F528M when combined with **SYSE** showed increased sensitivity compared to the F528M alone, while **SYSE** partially suppressed R481G/T526L and R481G/F528M, to a lower degree, sensitivity. Moreover, the current genetic system revealed the ability of UapA-R481G, T526L, F528M, R481G/T526L and R481G/F528M to accumulate different levels of 5FU. Finally, notice that none of the combined mutations affected proper localization of UapA in the plasma membrane (right column).

Radiolabeled xanthine transport assays further confirmed the above observations with all mutants containing the **SYSE** sequence, except F528M/**SYSE**, showing practically no xanthine transport (Figure 3.9b). F528M/**SYSE** showed a very low but considerable transport rate (15-20% of the wild-type; Figure 3.9b) that was apparently sufficient for conferring growth (Figure 3.9a). Importantly, mutations in residues Arg^{481}, Thr^{526}, and Phe^{528} in a wild-type NAT context led to increased xanthine transport rates (Figure 3.9b). We also performed uptake assays using radiolabelled uracil to test whether strains with increased capacity for 5FU accumulation, also have a capacity for uracil transport. Indeed mutants T526L, R481G/T526L and R481G/F528M that possess a wild-type NAT motif showed detectable uracil transport. Interestingly, the **SYSE** sequence apart from showing no uracil transport, suppressed transport capacity of T526L, R481G/T526L and R481G/F528M. Thus, UapA mutants carrying the **SYSE** sequence show an apparent inability to transport radiolabelled uracil, despite their capacity to lead to significant 5FU accumulation. One explanation is that this UapA version has high-affinity specifically for 5FU, but low affinity for uracil. Similar cases are known to exist, such as the Lac permease which shows higher affinity for various chemical analogues of D-galactopyranosides than its physiological substrates due to additional interactions of substituted positions in the analogues [159,160].

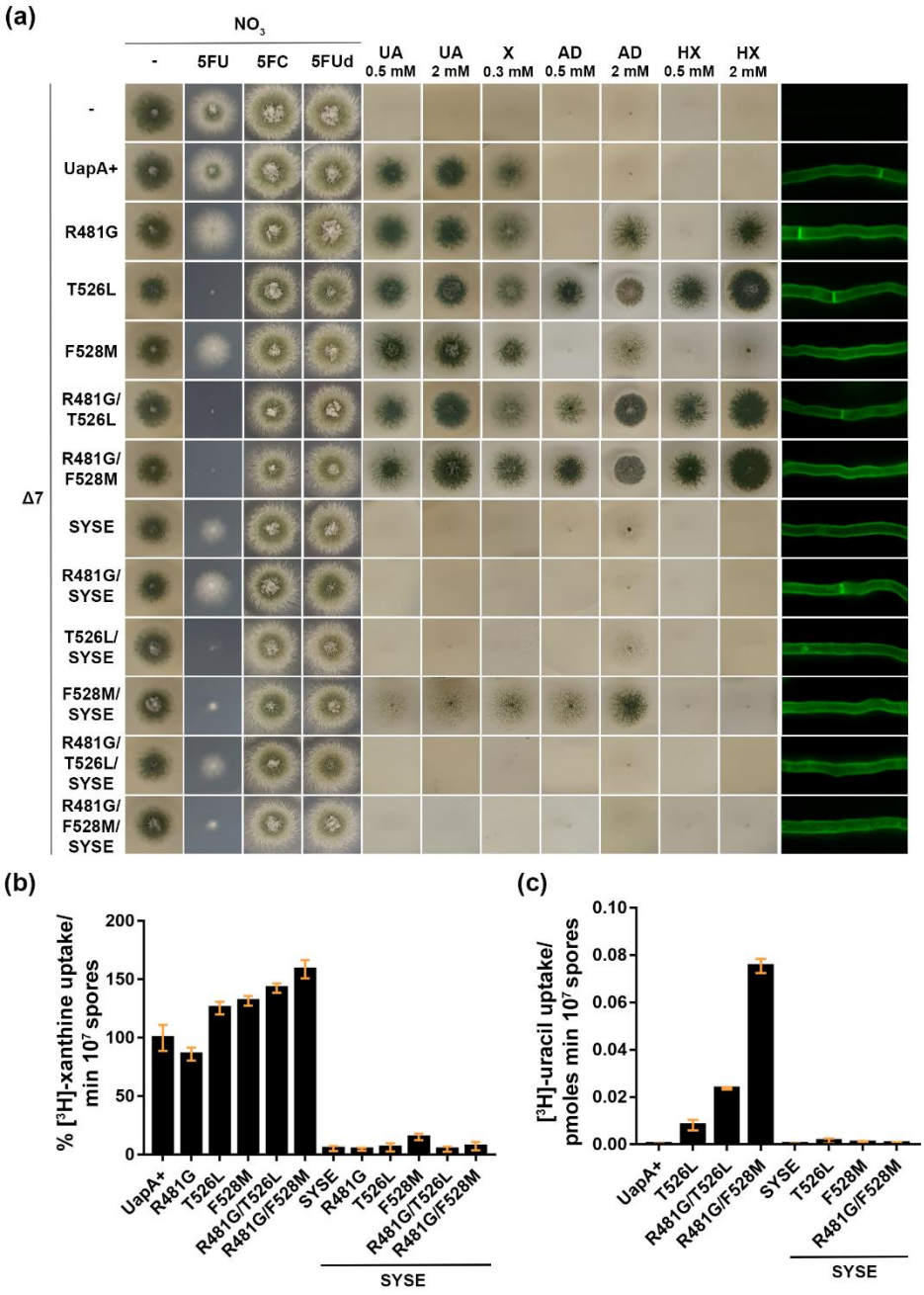
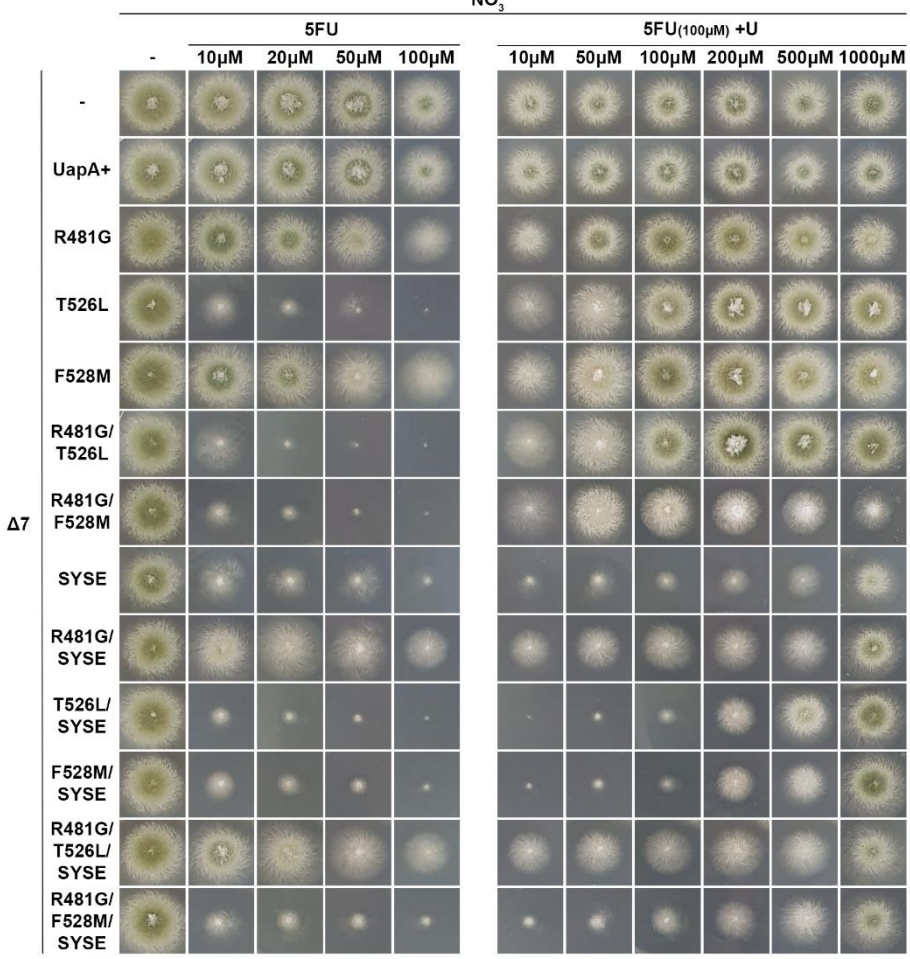


Figure 3.9. Combination of the SYSE NAT motif sequence with specificity mutations leads to functionally distinct UapA versions (A) Growth tests of strains expressing UapA versions carrying the **SYSE** sequence in the NAT motif combined with selected specificity mutations (F528M, T526L, R481G, R481G/F528M or R481G/T526L). Growth conditions and control strains are as described in Figure 3.7. Inverted fluorescence microscopy images confirming the proper localization of all UapA mutant versions to the PM is shown on the last right column. (B) Comparative ^{3}H-xanthine (0.1μM) transport rates in strains expressing UapA versions combining **SYSE** with selected specificity mutations (F528M, T526L, R481G, R481G/F528M or R481G/T526L). Details are as in Figure 3.8A. Standard deviation is depicted with error bars. (C) Comparative ^{3}H-uracil (0.1 μM) transport rates in strains expressing UapA mutant versions that conferred increased capacity for 5FU accumulation. Details are as in Figure 3.8A. Standard deviation is depicted with error bars.

To test this possibility we performed again *in vivo* competition assays to follow in this case 'reversion' of 5FU toxicity by increasing concentrations of uracil. Firstly, we performed

growth tests with a 5FU concentration range (10-100 μM). All mutants except R481G, F528M, R481G/**SYSE** and R481G/T526L/**SYSE**, that showed significantly reduced growth only at 100 μM, were very sensitive to 5FU at concentrations as low as 10 μM (Figure 3.10a left panel). Using 100 μM 5FU as a standard concentration, we performed the *in vivo* competition assays with a uracil concentration range of 10-1000 μM (Figure 3.10a right panel). In UapA versions with mutations outside the NAT motif uracil suppressed toxicity at a concentration range of 50-100 μM, regardless of the toxicity level, which suggested that they recognize 5FU and uracil with comparable affinities. Notice that R481G/F528M 5FU toxicity seems to be suppressed up to 50-100 μM of uracil, but as uracil increased above that threshold toxicity became more prominent. On the contrary, in mutants with the **SYSE** sequence 5FU toxicity was suppressed only by 10-fold excess uracil (1000 μM) which suggested that the presence of the **SYSE** sequence might lead to either low affinity for uracil or higher affinity for 5FU or both. To address this issue directly, we estimated the K_i values for uracil and 5FU in F528M/**SYSE**, which was the only combination of **SYSE** with other mutations that acquired a considerable transport of radiolabeled xanthine. Figure 3.10b shows that F528M/**SYSE** recognized uracil and 5FU with quite similar affinities (102 and 271 μM, respectively), while F528M alone did not alter K_i values for these two substrates compared to the wild-type. Thus the presence of **SYSE**, at least in the context of F528M does not make a significant distinction between uracil and 5FU. F528M/**SYSE** contributed also to better binding of adenine (275 μM) but not hypoxanthine and conserved high affinity for physiological substrates similar to wild-type UapA (Figure 3.10b), in line with growth tests (see Figure 3.9a). F528M alone did not affect significantly K_i values also for adenine and hypoxanthine (Figure 3.10b), despite very low apparent accumulation in growth tests (see Figure 3.9a). K_i values for novel mutations outside the NAT motif were also measured (data for R481G and T526L mutants from [9,10] are shown for comparison). Single mutants R481G and T526L did not affect significantly K_i values for adenine, hypoxanthine and uracil whereas the double mutant R481G/T526L recognized uracil with a moderate affinity (162 μM; Figure 3.10b). Similar results have been previously shown for double mutants of Arg^{481} and Phe^{528} residues [9]. In the present analysis, 1000 μM of uracil did not compete with the uptake of radiolabeled xanthine in the R481G/F528M mutant, but the K_m value for uracil, measured using 0,1 μM radiolabelled uracil, was 52 μM (shown in Figure 3.10b with an asterisk). This unexpected finding may be interpreted by the presence of a second binding site. In a previous study, docking calculations also indicated the existence of a secondary binding site which includes Thr^{526} and Phe^{528}, but the genetic analysis did not support the docking findings [13].

(a)



(b)

Substrate binding profile							
	Km/i (μM)						
Substrate/ligand	UapA	R481G	T526L	F528M	R481G/T526L	R481G/F528M	F528M/SYSE
Xanthine	4	18	5	5	6	4	2
Uric acid	6	19	10	24	8	28	15
Adenine	NI	NI	>1000	NI	>1000	NI	275
Hypoxanthine	NI	NI	>1000	NI	>1000	NI	NI
Uracil	NI	~1000	>1000	NI	162	NI/ 52*	102
5FU	NI	ND	ND	ND	ND	ND	271

Figure 3.10. Relative recognition of 5FU versus uracil as established by growth tests in UapA mutants. (A) Left panel: Concentration-dependent 5FU toxicity in the mutants. Right panel: *In vivo* competition of 100 μM 5FU toxicity by increasing concentration of uracil (range 10–1000 μM) in the mutants. Growth conditions are as described in Figure 3.7. (B) K_m (xanthine) or K_i (uric acid, adenine, hypoxanthine, uracil, 5FU) values estimated via competitive inhibition assays of radiolabelled xanthine uptake ([87]; see also Materials and methods) in selected mutants. The K_m for uracil of the R481G/F528M mutant using radiolabelled uracil is shown with an asterisk, next to the K_i value that was measured using radiolabeled xanthine. Details are as described in Figure 3.7B. Results shown are averages of triplicate measurements from 3 independent assays with a SD of < 20%.

In an effort to find additional residues that interact functionally with the NAT signature motif we investigated whether mutation of the equivalent residue to 408 from a Glu to a Gln coincides with other mutations in NATs in the course of evolution. More specifically, we explored if the presence of a Glu residue at position 408 "needs" the presence of a specific residue in another position, in order for the **SYSE** sequence to be functional in the context of UapA. For this we used correlated sequence conservation algorithms (BIS2Analyzer, see Materials and methods) and we found that the analogous residue to 410 in NAT proteins is conserved as an Asn only when the residue at position 408 is occupied by a Gln and when at position 408 is a Glu (e.g. **SYSE**), 410 is occupied by variable aliphatic amino acids (see also Figure 3.6a). Thus we combined **SYSE** with mutations N410I and N410V to test if this residue has a 'cryptic' role in determining specificity in a context-dependent manner. Figure 3.11 shows that none of the mutations suppressed loss of transport capacity by **SYSE** but instead resulted in a significant reduction in 5FU toxic accumulation. Notably, single N410V mutation led to loss of apparent UapA transport activity of purines, while N410I had no effect on UapA activity and specificity in respect to growth tests. The role of this residue has not been tested before in UapA, whereas the equivalent residue in the bacterial homologue XanQ, Asn^{326}, was not essential for xanthine transport [92,93,161]. Additionally, studies in the bacterial xanthine/uric acid transporters, UacT and SmXUacT, that conserve a Gln at the position 408 showed that the "opposite" mutations V312N and V320N (equivalent residue to 410) resulted to functional inactivation [161,162]. Thus, re-establishment of the combination present naturally in UapA is not compatible with transport activity in the context of those transporters, further showing the complexity of epistatic interactions in NATs.

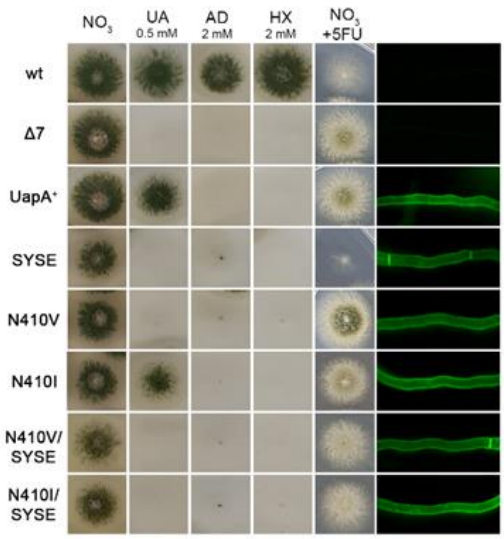


Figure 3.11. Growth tests of strains expressing UapA carrying the SYSE NAT motif sequence genetically combined with N410I and N410V. Details are as described in Figure 3.7.

3.2.4 Suppressor mutations conferring uric acid transport in UapA-SYSE support the functional importance of Phe^{528} and reveal new residues critical for specificity

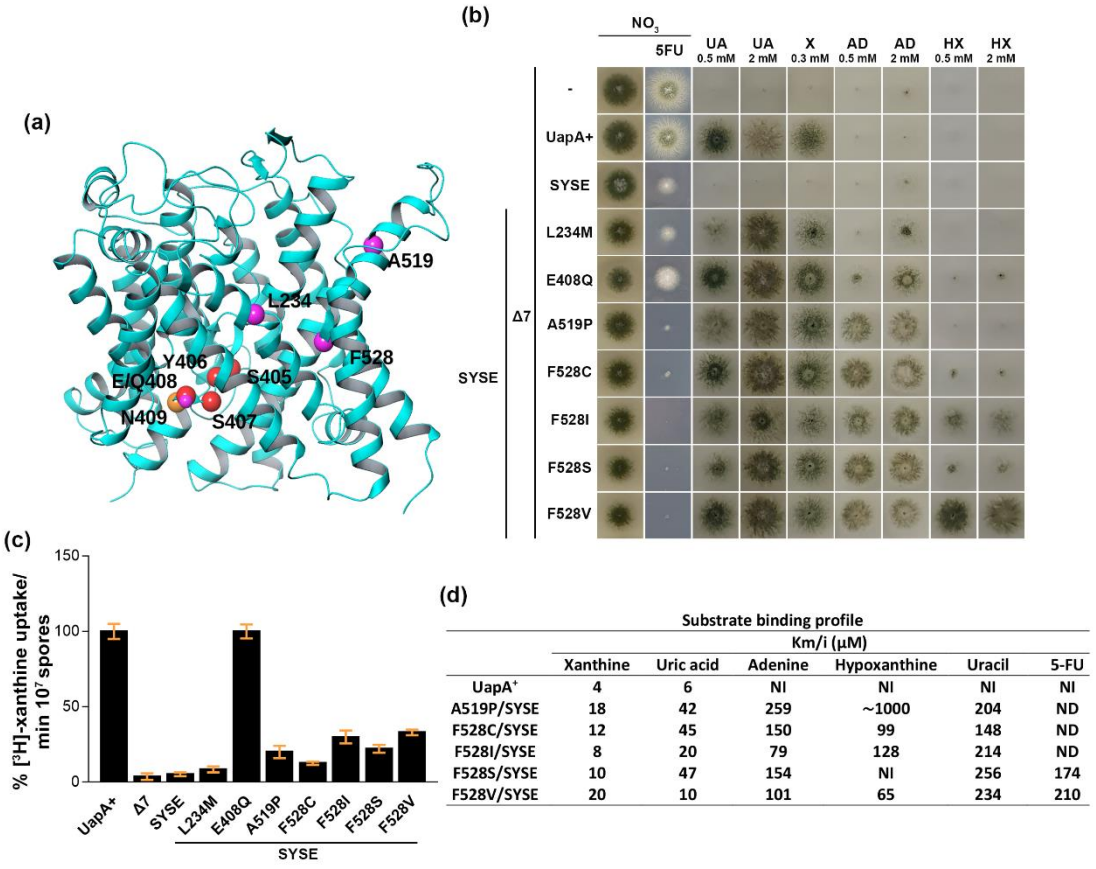
To further understand why **SYSE** leads to inability for purine transport and explore more about the epistatic interactions that determine specificity in NATs, we performed standard UV mutagenesis of the **SYSE** mutant. We selected genetic suppressors that restore UapA-mediated transport activity and thus growth on media containing UA as the sole N source. We obtained >50 suppressors and UapA ORF of 33 of them was amplified by PCR and sequenced. All revertants included mutations within the uapA orf and their profile and position is summarized in Table 3.1 and Figure 3.12a. Overall we got 7 different amino acid substitutions that concerned four residues: Leu^{234}, Glu^{408}, Ala^{519} and Phe^{528}. Mutation of Glu^{408} reverted the **SYSE** sequence to **SYS**Q, which was expected to be isolated as it was earlier shown to confer growth on UA (Figure 3.7a). Mutations in Phe^{528} were: F528C, F528I, F528S and F528V. Mutations in this residue were also expected as F528S was selected previously as a suppressor mutation of the cryosensitive Q408E [95]. Moreover, replacements of this residue with aliphatic (Ala, Met) or polar amino acids (Ser, Thr, Gln or Asn) converted UapA into a promiscuous purine-uracil transporter ([154] and Figure 3.9a). However we also got unexpected mutations. A mutation that concerns residue Leu^{234} located in TMS5 (L234M) that was also isolated as a suppressor mutation of a mutant with inefficient dimerization (see Chapter 3.3) and A519P that has never appeared in previous genetic screens. Leu^{234} is close to residues Phe^{460}, Ala^{407}, Phe^{155}, Ile^{230} and Phe^{528} and might be a part of the substrate translocation trajectory. Ala^{519} locates in TMS14, same as Thr^{526} and Phe^{528} that both affect specificity, and has not been studied before. Introduction of a Pro residue at this position might modify the positioning of these residues and thus affect specificity.

We analyzed the role of those mutations in function and specificity by growth tests and uptake assays. All revertants grew on UA and X, albeit with different levels, with L234M being the least efficient (Figure 3.12b). E408Q, as expected, grew exactly as in Figure 3.7a. In addition, all mutants, except from L234M, grew well on adenine and resulted in increased 5FU accumulation compared to the original **SYSE** mutant. Finally, only F528V and F528I (moderately) showed significant growth on hypoxanthine. Uptake assays using radiolabelled xanthine revealed that all revertants except from E408Q, have reduced transport rates (~15-28% of wild-type levels), with L234M showing the lowest rate (Figure 3.12c). L234M did not seem to significantly increase **SYSE** transport rate, despite conferring growth on xanthine. It should be noted again that in growth tests, purines were added at higher concentrations (1–

2 mM) to be used as nitrogen sources, compared with the uptake experiments (0.3–0.5 mM) for technical reasons. Additionally, growth tests depict accumulation and not rate. Thus, any mutation that reduces the affinity or the transport rate might exhibit very low transport rates

Table 3.1. Profile of SYSE suppressors

Mutation	TMS	Domain	Codon change	Number of isolates
L234M	5	Dimerization	CTG → ATG	2
E408Q	10	Core	GAG → CAG	5
A519P	14a	Dimerization	GCG→ CCG	3
F528S	14b	Dimerization	TTT →TCT	10
F528S	14b	Dimerization	TTT →TCA	1
F528C	14b	Dimerization	TTT → TGT	6
F528V	14b	Dimerization	TTT → GTT	5
F528I	14b	Dimerization	TTT → ATT	1



(d)

Substrate binding profile Km/i (μM)	Xanthine	Uric acid	Adenine	Hypoxanthine	Uracil	5-FU
UapA+	4	6	NI	NI	NI	NI
A519P/SYSE	18	42	259	~1000	204	ND
F528C/SYSE	12	45	150	99	148	ND
F528I/SYSE	8	20	79	128	214	ND
F528S/SYSE	10	47	154	NI	256	174
F528V/SYSE	20	10	101	65	234	210

Figure 3.12. Characterization of mutations allowing uric acid transport in the UapA-SYSE context (A) Topological model showing the position of suppressor mutations relative to the NAT motif in UapA (**B**) Growth tests of suppressors. Growth conditions and control strains are as described in Figure 2. (**C**) Comparative ^{3}H-xanthine (0.1μM) transport rates in suppressors. Details are as described in Figure 3.8A. Standard deviation is depicted with error bars. **(D)** $K_{i/m}$ values (μM) determined using ^{3}H-xanthine uptake competition in suppressors. Details are as described in Figure 3.8B. ND indicates not determined. Results shown are the averages of triplicate measurements from 3 independent assays with a SD of < 20%.

in uptake experiments, but can still allow growth on the corresponding substrate. Next, we estimated the K_i values for all suppressors with significant transport activity (Figure 3.12d). All suppressors showed reduced but relatively high affinity for UapA physiological substrates and moderate affinities for other purines, uracil and 5FU with few exceptions of certain mutants towards specific substrates. Given that mutations in Phe^{528} do not alter significantly the affinity of substrates [154], we could assume that the **SYSE** sequence is, in principle, the reason of acquisition of binding affinities for substrates. Thus, the **SYSE** sequence transforms UapA into a promiscuous transporter, recognizing, apart from uric acid and xanthine, adenine hypoxanthine, uracil and 5FU. However, combinations of mutations in Phe^{528} with Q408E modified K_i values for adenine and hypoxanthine, compared to the single Q408E, revealing possible epistatic interactions between those two residues [154].

3.2.5 Computational analysis gives insights on the specificity modifications of the various NAT signature motif versions

To better understand the previous results we performed molecular simulations of protein-substrate interactions. We constructed models based on the structure of UapA in the inward conformation for TFA**E**, **SYS**Q and **SYSE** and we focused on the differences in the binding cavity. Firstly, we observed that in the presence of F406Y the pKa of Glu^{356} and Asp^{360} (TMS8), that are critical for substrate binding, were modified. More specifically, in **SYSE** (and **SYS**Q similarly) pKa of Glu^{356} rises from 7.5 (wild-type UapA) to 10.0, while pKa of Asp^{360} drops from 9.4 to 7.4. Considering that the previous experiments were performed at pH 6.8 Glu^{356} is considered protonated, while Asp^{360} negatively charged in **SYSE**. Additionally, the phenolic hydroxyl of Y406 forms a hydrogen bond with the Val^{153} backbone carbonyl (Figure 3.13a,b), which stabilizes the main part of the binding cavity between TMS3 and TMS10, which is known to form the characteristic beta-sheet structure in the middle of the transmembrane domain where the substrate is embedded. This stabilization may contribute to the increased transport capacity for 5FU seen in T**Y**AQ and **SYS**Q and the slightly better binding of hypoxanthine and adenine compared to the wild-type TFAQ (Figures 3.7 and 3.8).

Next, we used Induced Fit Docking (IFD) calculations to perform simulations of **SYSE** interactions with substrates. We showed that the most energetically favorable binding structure of uracil, exhibited a remarkable similarity to the crystal structure of uracil bound by its natural transporter UraA [3] (Figure 3.13a) and a quite similar positioning resulted also

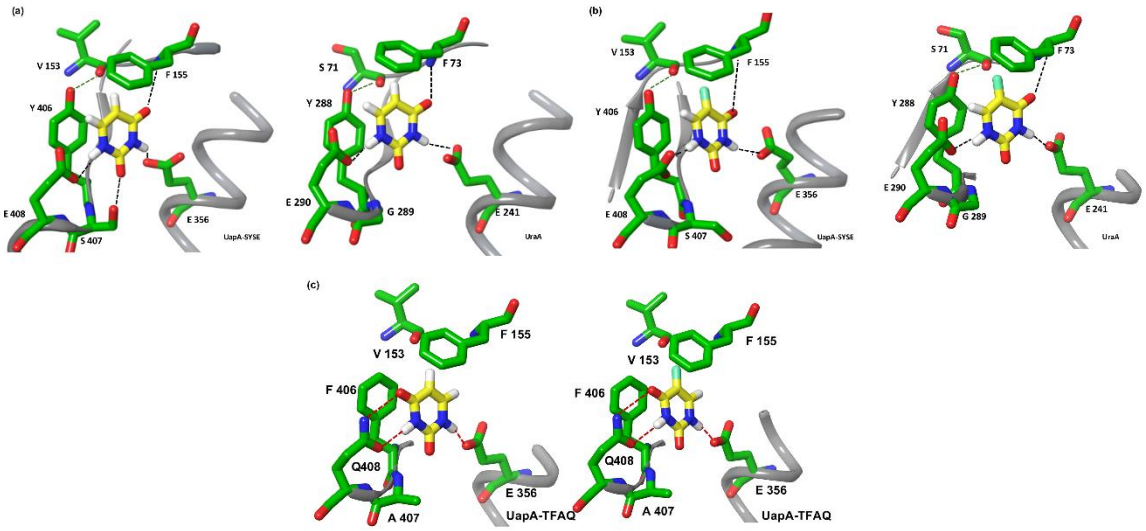


Figure 3.13. Molecular Simulations of UapA-SYSE interactions with substrates (**A**) UapA-**SYSE**-uracil and UraA-uracil interactions. Left panel: The UapA-**SYSE**-uracil interaction is stabilized by the formation of four H-bonds (depicted as dashed lines) between N1H and N3H with Glu^{408} and Glu^{356} carboxylates, respectively, C2=O with Ser^{407} hydroxyl group and C4=O with Phe^{155} backbone, as well as, a p-p stacking formed between uracil with both Phe^{155} and Tyr^{406}. Notice also that Tyr^{406} hydroxyl group forms a H-bond with Val^{153} backbone. Right panel: The UraA-uracil structure as resulted from X-ray crystallography is presented for comparison [6]. (**B**) UapA-**SYSE**-5FU (left) and UraA-5FU (right) interactions. In both structures 5FU is oriented similarly to uracil with the F atom located in a hydrophobic cleft and C-F bond parallel to both Phe^{155} and Tyr^{406} rings. (**C**) Molecular Simulations of UapA interactions with uracil (left) or 5FU (right).

for 5FU (Figure 3.13b). More specifically, uracil formed hydrogen bonds with the side chains of Glu^{408}, Glu^{356} and Ser^{407} and π-π stacking interactions with Phe^{155} and Tyr^{406}. In UraA, uracil formed exactly the same interactions except from the hydrogen bond with the equivalent to Ser^{407} residue, as UraA conserves a glycine at this position. Notice that the F atom is located in a hydrophobic cleft of the binding pocket where the C-F bond is parallel to both Phe^{155} and Tyr^{406} rings, without affecting proper orientation of the uracil ring. Notably, these ligands interact via a mirror-image orientation with wild-type UapA (Figure 3.13c), which does not transport uracil or 5FU thus, uracil/5FU docking in **SYSE** and UraA is different from wild-type UapA. However, UraA recognizes uracil with higher affinity than **SYSE**, suggesting that other residues outside the major binding site may be also crucial for recognition or binding affinities cannot be well predicted by modeling approaches.

We also performed IFD calculations to understand specificity changes towards hypoxanthine in the NAT mutant versions. In all cases we present the lowest energy binding modes of hypoxanthine in the various versions. It was shown before that in TFA**E**

hypoxanthine inhibited xanthine transport with a relatively low K_i value (70 μM) but this mutant could not transport hypoxanthine. In wild-type UapA Gln^{408} forms two H-bonds with N1H and C2=O of xanthine, Glu^{356} forms one H-bond with N7H, Phe^{155} forms one H-bond with N9 and Ala^{407} forms one H-bond with C6=O ([2] and Figure 3.14a). Hypoxanthine got a completely different orientation interacting via C6=O with Gln^{408} and via N9H with Glu^{356} justifying why wild-type does not bind hypoxanthine (Figure 3.14b left panel). However, in TFA**E** N9H of hypoxanthine interacted strongly with the two negative charges of Glu^{408} while the C6=O group formed an H-bond with Ser^{154} and N1H interacted with the two positive charged of Glu^{356} (Figure 3.14b central panel). The bond between N9H and Glu^{408} is so robust that may inhibit transport explaining together with the topological disorientation within the binding site why hypoxanthine is recognized with relatively high binding affinity but is not transported. Finally, we performed simulations also for the **SYS**Q mutant that recognizes hypoxanthine with low affinity (~1000 μM). **SYS**Q made more interactions than the wild-type UapA that does not bind hypoxanthine via the protonated Glu^{356} with N3 and N9H of hypoxanthine and via Gln^{408} with C6=O, possibly explaining low affinity (Figure 3.14b right panel).

3.2.6 Phe^{528} dynamically interacts with the substrate binding site

Mutations in Phe^{528} were isolated several times in this work, but also previously, and Phe^{528} proved to be a key residue affecting specificity alone and in combination with mutations in the NAT signature motif. Previous studies proposed that this residue is part of the outer gating elements, controlling the selective passage of substrates [7,13]. In order to expand our understanding on the functional role of this residue we constructed the model of the outward conformation of UapA by homology modelling to the anion exchanger domain of Band 3 crystal structure. Band 3 or anion exchanger 1 (AE1), is a structural homologue of UapA, found in human red blood cell membrane where it carries out chloride/bicarbonate anion ($HCO3^-$) exchange across the plasma membrane [163]. The crystal structure of the anion exchanger domain has been recently resolved in an outward conformation and represents an excellent model for the outward conformation of transporters operating by the elevator mechanism [45]. In the inward conformation Phe^{528} is close to the substrate binding site at a ~6.00 Å distance from Glu^{356} (Figure 3.15b). However in the outward modelled conformation Phe^{528} got closer at a ~3.50 Å distance interacting with Glu^{356} through p-electrons (Figure 3.15a). Additionally, this distance remained stable through the 100 ns simulation, suggesting that is a

quite stable interaction. This suggests that Phe^{528} interacts dynamically with the substrate binding site and may explain why Phe^{528} mutations affect specificity in UapA.

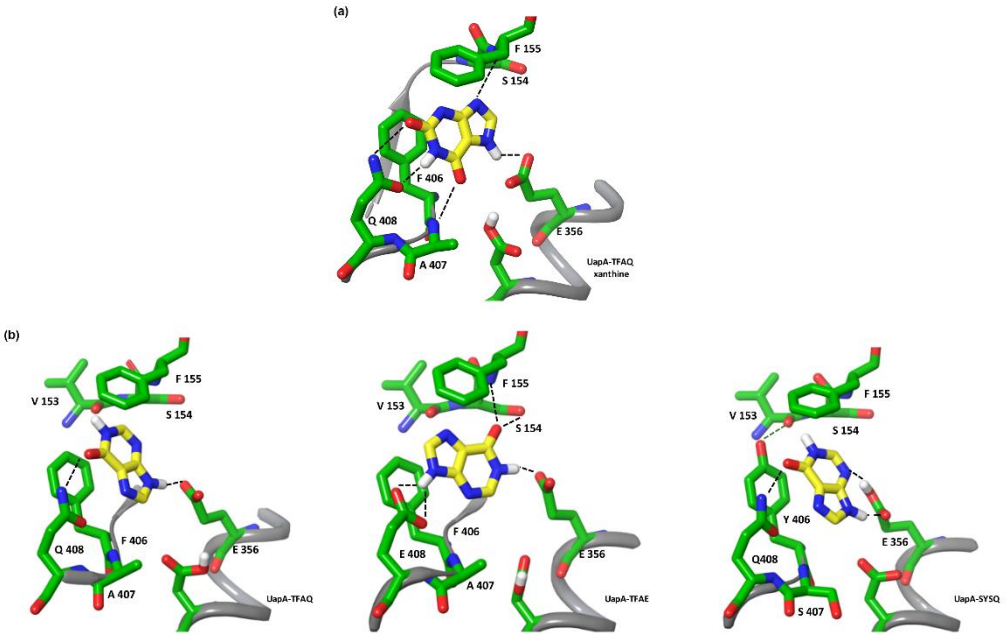


Figure 3.14. Molecular Simulations of UapA mutants-substrate interactions. (A) Wild-type UapA-xanthine interactions [2]. (**B**) Hypoxanthine interactions with UapA, UapA-TFA**E** and UapA-**SYS**Q. Hypoxanthine forms two H-bonds with the UapA binding cavity, one between N9H and Glu^{356} carboxylate and the second between C6=O and Gln^{408} amide group. In UapA-TFA**E**, hypoxanthine N9H and N1H both form H-bonds with Glu^{408}, and Glu^{356} carboxylates, respectively, while C6=O forms H-bond with Ser^{154} hydroxyl group and Phe^{155} backbone. In UapA-**SYS**Q hypoxanthine interaction Glu^{356} is protonated. Hypoxanthine H-bond interactions with the binding cavity are between N9H, N3 and Glu^{356}, as well as, between C6=O and Gln^{408}.

3.2.7 Conclusions

The NAT signature motif in TMS10 is known to be critical for substrate binding and specificity. We showed in the previous chapter that this motif might contribute to the evolution of substrate specificity in the NAT family. Here we decided to functionally analyze NAT motif versions found in several NATs by constructing a series of UapA mutants. Those mutants were expressed in a novel genetic system that allows the characterization of an extended specificity profile and revealed context-dependent cryptic roles of residues in the NAT motif. We showed that **SYS**Q is a promiscuous UapA version while **SYSE** leads to loss of transport capacity for purines but gained a capacity for toxic 5FU accumulation in the context of UapA. In UapA-**SYSE** uracil is well stabilized getting a docking position, similar to the UraA, not seen in wild-type UapA. UV mutagenesis of the **SYSE** mutant revealed new residues that affect specificity, those are Leu^{234} and Ala^{519}. Combination of **SYSE** with mutations in Leu^{234}, Ala^{519} and Phe^{528} led to a

transporter able to translocate purines, uracil and 5FU. Finally, computational analysis proposed that Phe^{528} interacts dynamically with the substrate binding site and more specifically with Glu^{356}.

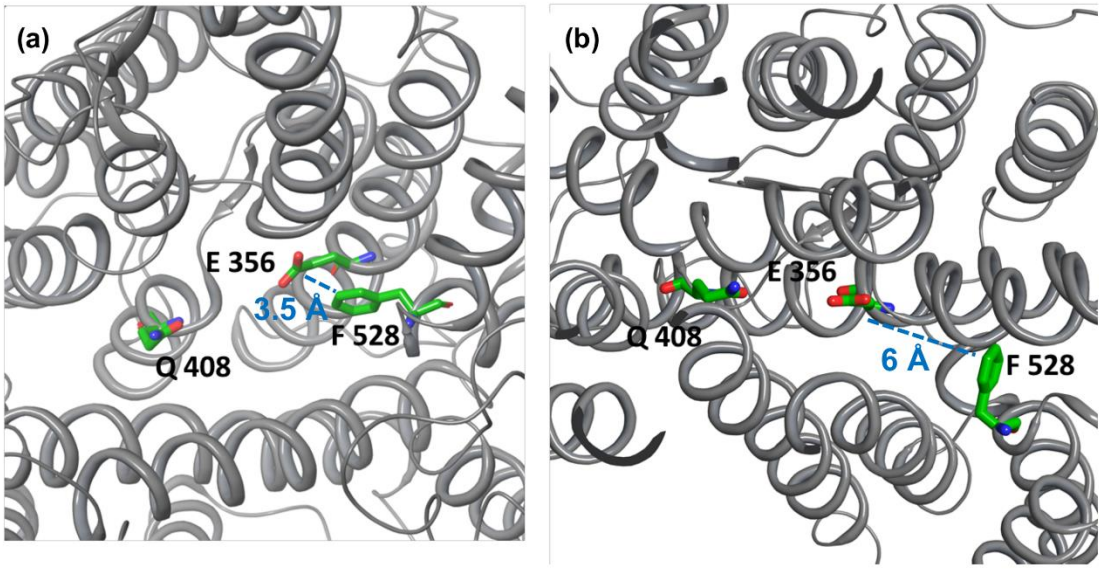


Figure 3.15. UapA outward conformation. (**A**) The model constructed using Band3 as template is shown mainly focusing on the interaction between Phe^{528} and Glu^{356}. The Glu^{356} carboxylate group remains almost parallel to the Phe^{528} phenyl ring (3.5 Å) during the 100 ns MD simulation. (**B**) The corresponding distance in the inward conformation is much higher (pdb structure 5i6c).

3.3. Specific residues in a purine transporter are critical for dimerization, ER exit, and function

Adapted from Kourkoulou, A., Grevias, P., Lambrinidis, G., Pyle, E., Dionysopoulou, M., Politis, A., Mikros, E., Byrne, B. & Diallinas, G. ***Genetics****. 213, 1357-1372 (2019).*

3.3.1 Rationale of the study

Functional and structural dependence of transporters on specific membrane lipids has only recently been explored in detail. Recently, native mass spectrometry (MS) showed that the UapA dimer co-purified with lipids, mostly phosphatidylcholine (PC), phosphatidylinositol (PI) and phosphatidylethanolamine (PE), but delipidation caused dissociation into monomers [12]. Addition of either PI or PE restored the UapA dimer suggesting that these lipids have a stabilizing effect on this form of the transporter. Molecular dynamics predicted a specific lipid binding site in the dimer interface that is formed by three positively charged residues, Arg^{287}, Arg^{478} and Arg^{479}. Triple-alanine replacement of these arginines revealed that these residues,

and pottentially lipid binding at this site, are critical for the formation of functional dimers. Molecular dynamic simulations predicted that, apart from the dimer interface, lipids might also bind on the opposite side of the mobile core domain of UapA. Here, we performed genetic screens in order to investigate further the role of residues Arg^{287}, Arg^{478}, and Arg^{479} in UapA dimer formation and/or stability, and study the role of the predicted lipid binding residues that map in the core domain.

3.3.2 Arg^{287}, Arg^{478}, and Arg^{479} are crucial for *ab initio* dimerization of UapA in the ER network

Previous studies proposed that UapA dimers are formed at the ER and dimerization is essential for subsequent membrane trafficking [124]. However, residues Arg^{287}, Arg^{478} and Arg^{479} are critical for dimer formation but not trafficking to the plasma membrane (PM). The triple R287A/R478A/R479A mutant could either result to inability for dimerization at the ER or dimer disruption at the PM. In other words, these residues could either be essential for the *ab initio* formation of dimers at the ER or UapA dimerization occurs at the ER, dimers travel to the PM and dissociate due to loss of specific lipid interactions. To test those scenarios we used a previously described BiFC assay [124] adapted to follow the trafficking of de novo-made UapA by performing time-course experiments in strains expressing UapA under the regulatable $alcA_p$ promoter [164]. Conidiospores of a strain containing two copies of the $alcA_p$-uapA gene, tagged at either the N- or the C-part of the yfp ORF, were allowed to germinate overnight under conditions that repress UapA transcription (Glucose MM, see also Materials and Methods) until young hyphae developed. Then culture shifted to a derepression medium (Fructose MM) and hyphae were observed at specific time-points for 1-4 h. Figure 3.16a shows that in 3 h after derepression we detected reconstitution of split-YFP fluorescence at the ER network in the strain expressing wild-type UapA. At the same point for the triple R287A/R478A/R479A mutant we detected fluorescence at the ER network but the intensity was very weak. After 4 h of derepression both wild-type UapA and R287A/R478A/R479A marked the PM, but again the signal from R287A/R478A/R479A was weak as expected from previous studies [12]. Therefore, the triple mutant significantly reduced UapA dimerization at the ER, but the transporter was able to travel and reach the PM as judged by the normal PM localization of R287A/R478A/R479A tagged with intact GFP (Figure 3.16b). Thus, in this case either UapA monomers can be secreted to the PM or R287A/R478A/R479A do not lead to

total loss of dimerization but results to inability for functional dimerization that is needed for proper reconstitution of split-YFP but is redundant for the trafficking mechanism.

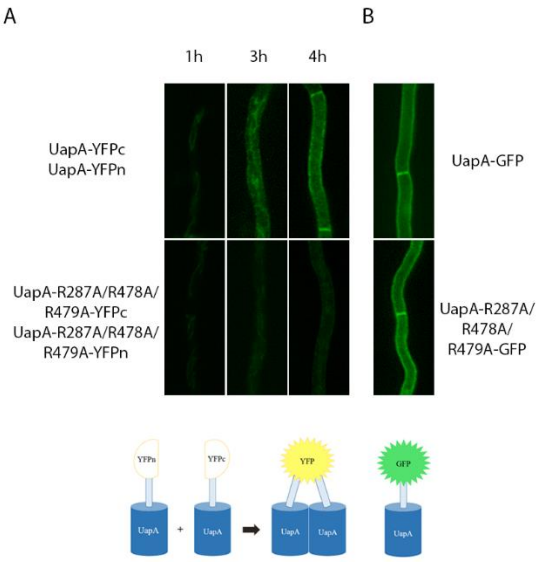


Figure 3.16. Residues Arg^{287}, Arg^{478} and Arg^{479} are crucial for *ab initio* dimerization of UapA in the ER. (A) Bimolecular complementation (BiFC) analysis of de novo expressed UapA and the R287A/R478A/R479A mutant, expressed via the *alcAp* promoter in the presence of non-repressive carbon (fructose) and NO_3 as nitrogen source, for 1, 3 and 4h. Mutant constructs tagged with the individual YFP domains were co-expressed in *A. nidulans*. Upon UapA dimerization the YFP is reconstituted. YFP fluorescence was visualized by epifluorescence inverted microscopy. (B) Localization of the GFP-tagged UapA constructs.

3.3.3 Genetic suppressors of the triple R287A/R478A/R479A mutant map in the core or dimerization domains and re-establish UapA dimerization

In order to expand our understanding on the molecular basis of how these residues affect dimerization, we performed standard UV mutagenesis of triple mutant R287A/R478A/R479A in order to select genetic suppressors that restore UapA-mediated transport activity and thus growth on media containing uric acid as the sole N source. We obtained >50 suppressors and the UapA ORF of 38 of them was amplified by PCR and sequenced. Table 3.2 summarizes the profile of the suppressor mutations obtained. Overall we got 13 different amino acid substitutions, in 10 different residues, 8 in the core domain of UapA, and 2 in the gate domain (Figure 3.17a). Growth tests showed that all revertants were able to grow on uric acid and xanthine, both in 25°C and 37°C, and also some of them showed moderate growth on adenine or hypoxanthine (compared with the strain expressing wild-type UapA; Figure 3.17b). In addition, almost all suppressors were sensitive, albeit at a different degree, to oxypurinol, a toxic xanthine analogue. Comparative uptake assays further showed that most suppressors have a positive effect on the low transport rate of the genetic background of triple mutant R287A/R478A/R479A with E286Q (~70%) and E286K (~51%) obtaining the highest transport rates (Figure 3.17c). However, the effect of some suppressors (L192F and T401F) was marginal despite conferring growth on xanthine, again probably due to the differences between the two assays explained before (see Chapter 3.2.2 and 3.2.4). We also estimated the

Table 3.2. Mutations suppressing the inability of UapA-R287A/R478A/R479A to grow on UA.

Mutation	Location	Domain	Number of isolates	Codon change	Number of changed nucleotides
S119T	TMS2	Core	1	TCG→ACG	1
V150I	TMS3	Core	9	GTT→ATT	1
V153T	TMS3	Core	1	GTC→ACC	2
I157F	TMS3	Core	3	ATC→TTC	1
I157L	TMS3	Core	3	ATC→CTC	1
L192F	TMS4	Core	2	TTG→TTT or TTC	1
L234M	TMS5	Dimerization	2	CTG→ATG	1
E286K	TMS6	Dimerization	6	GAA→AAA	1
E286Q	TMS6	Dimerization	2	GAA→CAA	1
A396P	TMS9	Core	1	GCC→CCC	1
T401F	TMS10	Core	1	ACC→TTC	2
T401P	TMS10	Core	1	ACC→CCC	1
L431F	TMS11	Core	6	TTG→TTT or TTC	1

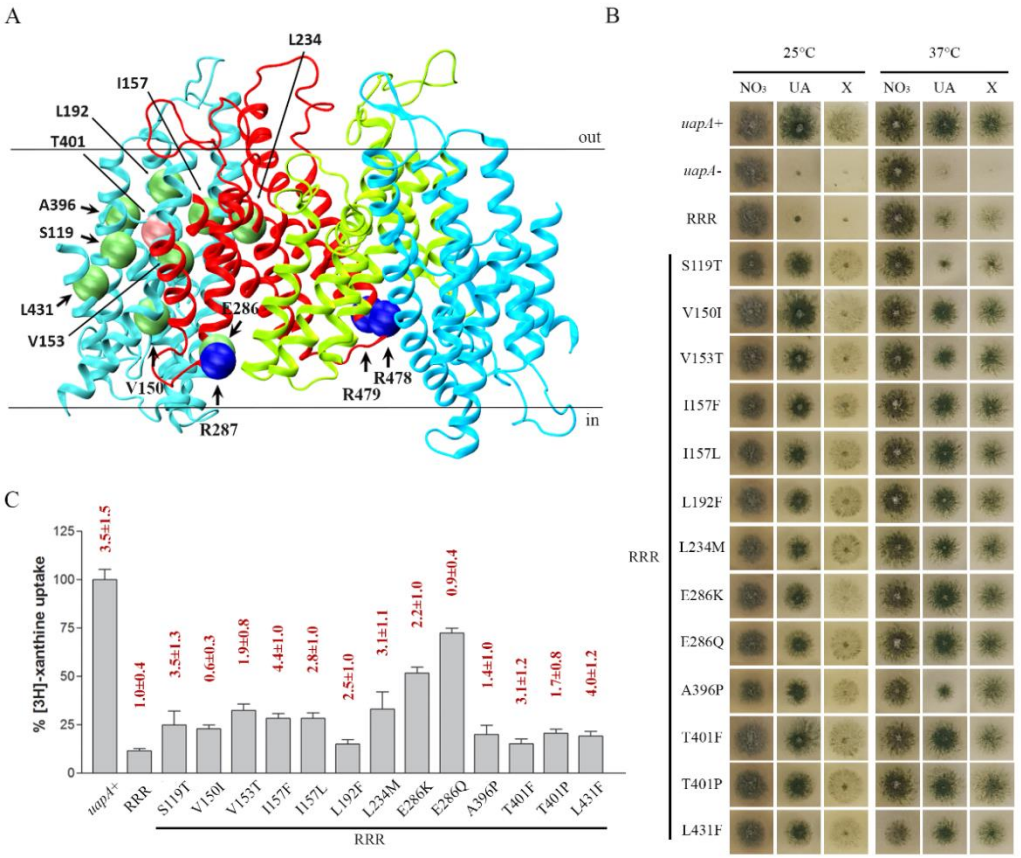


Figure 3.17 Genetic suppressors of R287A/R478A/R479A localized in the core and the dimerization domain of UapA partially restore UapA function. (A) Topology of amino acids modified in R287A/R478A/R479A suppressors. Core domains are colored light blue, and dimerization domains red and green for clarity. Substituted amino acids in the original mutant strain are shown with blue spheres, and in the suppressors with green and pink (Thr^{401}) spheres. (B) Growth tests of R287A/R478A/R479A suppressors on UapA physiological substrates. Control strains are strains with total genetic deletions in all major purine transporters (ΔACZ; negative control), referred in the figure as *uapA*$^-$, and a ΔACZ transformant expressing wild-type uapA–gfp (*uapA*$^+$; positive control). All suppressor strains and the original R287A/R478A/R479A strain are isogenic to the negative and positive control strains, and express UapA from single-gene copies of uapA tagged with GFP. All strains were grown in minimal media containing 10 mM nitrate (NO_3), 0.5 mM uric acid (UA), or 1 mM xanthine (X) as a nitrogen source at 25°C (left panel) or 37°C (right panel). RRR depicts the R287A/R478A/R479A original genetic background. (C) Relative ^{3}H-xanthine transport rates of R287A/R478A/R479A and R287A/R478A/R479A suppressors expressed as percentages of initial uptake rates (V) compared to the wild-type (*uapA*$^+$) rate. K_m values (μM) for xanthine are shown at the top of histograms. Results are averages of three measurements for each concentration point. SD was less than 20%.

approximate K_m values of several suppressors (Figure 3.17c, on top of histograms) and we showed that none of the suppressors affect significantly xanthine affinity. K_m values measured before for the original R287A/R478A/R479A mutation and wild-type UapA are shown for comparison [12].

Suppressors of R287A/R478A/R479A could either restore functional dimerization or lead to functional UapA monomers. To test these two alternatives, we used again the previously described adapted BiFC assay (see Chapter 3.3.2 and Materials and Methods). We selected three R287A/R478A/R479A suppressors, one of each domain (I157F and L234M) and T401P that was isolated again as a suppressor (see later on Chapter 3.3.5), and we followed UapA dimerization of *de novo* made UapA. Figure 3.18 shows that all suppressors studied increased reconstitution of split-YFP fluorescence, at both the ER and the PM, compared to the original R287A/R478A/R479A mutant supporting that suppressors restored function by restoring dimerization.

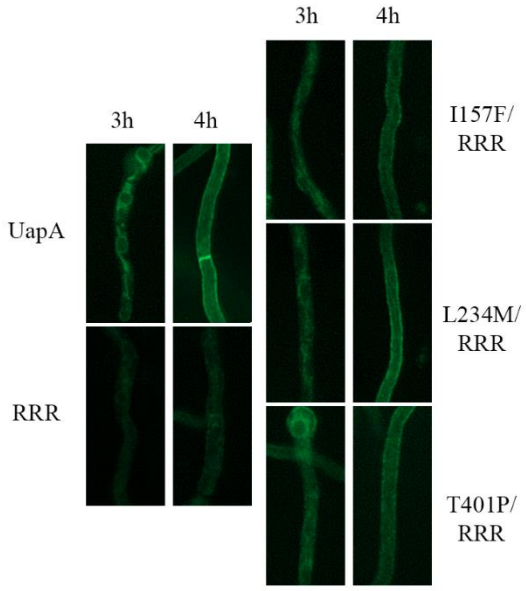


Figure 3.18. Genetic suppressors R287A/R478A/R479A re-establish UapA dimerization. Bimolecular complementation (BiFC) analysis of de novo expressed UapA, R287A/R478A/R479A and selected R287A/R478A/R479A suppressors, I157F, L234M and T401P, expressed via the *alcAp* promoter in the presence of non-repressive carbon (fructose) and NO_3 as nitrogen source, for 1, 3 and 4h. Mutant constructs tagged with the individual YFP domains were co-expressed in *A. nidulans*. Upon UapA dimerization the YFP is reconstituted. YFP fluorescence was measured by epifluorescence inverted microscopy. RRR depicts R287A/R478A/R479A.

3.3.4 MD provides a structural rationale for the effect of suppressor mutations on UapA stability and function

In order to understand the role of the suppressor mutations in UapA dimerization we examined their involvement in interactions and performed short MD simulations based on the UapA crystal structure. Most of the substitutions concerned residues located in the movable core domain (in TMS2, TMS3, TMS4, TMS9, TMS10 or TMS11). The core domain consists of two layers of transmembrane helices: TMS1, TMS3, TMS8, and TMS10 in the inner part in

contact with the dimerization domain, and TMS2, TMS4, TMS9, and TMS11 in the outer part in contact with the membrane lipids. These two layers are stabilized through an extended network of hydrophobic interactions and certain key polar interactions, mainly related to the residue Asn^{388}. Most suppressor mutations are located between these two transmembrane layers (Figure 3.19a and b). Ser^{119}, Val^{153}, Ile^{157}, Leu^{192}, Thr^{401}, and Leu^{431} interact with each other and with many residues of this network or are located between them. More specifically, Leu^{192} interacts with Tyr^{189} (TMS4), Thr^{401} (TMS10), Ile^{157} (TMS3), Phe^{165} (TMS3), Ile^{101} (TMS1), Ile^{193} (TMS4), and Ile^{346} (TMS8). Similarly, Ile^{157} interacts with Ile^{101} (TMS1), Thr^{401} (TMS10), Pro^{402} (TMS10), Val^{349} (TMS8), Ile^{346} (TMS8), and Leu^{192} (TMS4). Thr^{401} is located in the center between the two TMS layers of the core domain and interacts with Tyr^{189} (TMS4), Leu^{192} (TMS4), Ile^{193} (TMS4), Val^{153} (TMS3), Ile^{157} (TMS3), and Ile^{101} (TMS1). Val^{153} is surrounded by Thr^{401} (TMS10), Pro^{97} (TMS1), Val^{94} (TMS1), Met^{400} (TMS10), Ser^{119} (TMS2), Met^{403} (TMS10), and Cys^{123} (TMS2). On the opposite side, the Ser^{119} side chain is located in the middle between Val^{153} (TMS3), Val^{94} (TMS1), Met^{400} (TMS10), and Cys^{123} (TMS2). Finally, Leu^{431} is located between Ala^{87} (TMS1), Met^{90} (TMS1), and Leu^{120} (TMS2). The suppressor mutations, and more particularly S119T, V153T, L192F, T401F, and L431F, seem to enhance the above-mentioned interactions and mainly stabilize the core domain. To confirm this hypothesis, we constructed models of selected suppressor mutations (I157F, T401F, L192F and L431F) that were subjected to geometry optimization and we performed short MD calculations. Figure 3.19, c–f, shows that the phenyl moieties of the mutated residues I157F, T401F, L192F and L431F are placed in the space between the other lipophilic residues, increasing hydrophobic interactions between TMS8, TMS3, TMS4, and TMS10. Regarding S199T, introduction of a methyl group will also enhance, albeit to a lower degree, interactions with Val^{153} (TMS3) and Met^{400} (TMS10). Finally, V153T introduces a new hydrogen bond with Ser^{119}, that interacts with the backbone of Val^{94}, creating a hydrogen bond network between TMS3, TMS2, and TMS1. These findings suggest that suppressor mutations by strengthening hydrophobic interactions between the two transmembrane layers, stabilize the motile core domain. Thus, the stabilization of the moving part of the monomeric unit probably stabilizes also the dimer and partially restores function.

Apart from the core domain, suppressor mutations concern also residues, Leu^{234} and Glu^{286}, in the dimerization domain. Glu^{286} is located at the end of the cytoplasmic-facing part of TMS6 next to the mutated Arg^{287}. Suppressors E286K and E286Q, as they introduce positively charged residues probably replace directly the interactions with lipids of the nearby mutated Arg^{287} (i.e., in R287A). This is also supported by the fact that the double mutant

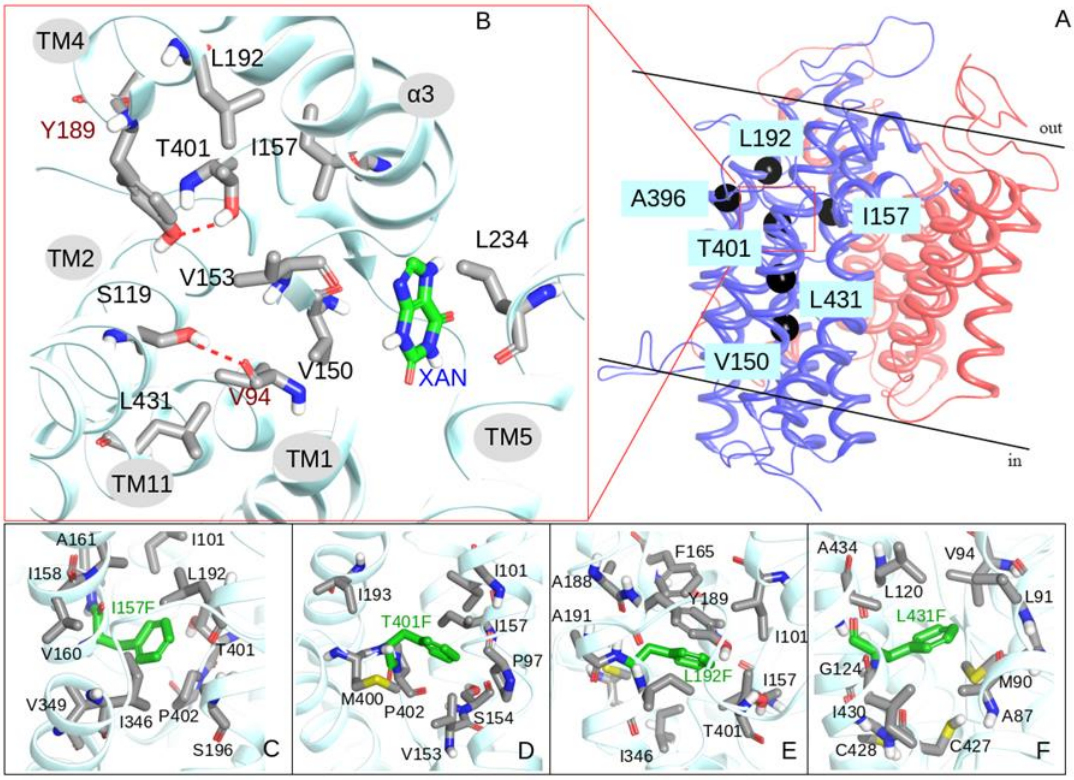


Figure 3.19 Molecular dynamics provide a structural rationale for the effect of suppressor mutations on UapA stability and function. (A) Ribbon representation of UapA monomer. The core domain is colored blue and the gate domain red. Topologies of type I suppressors on UapA crystal structure are depicted with black spheres. (B) Detailed view of the topology of Thr^{401} and amino acids within 4 Å, including residues mutated in suppressors in black lettering (see text). The substrate (xanthine) is also shown. (C)-(F). Detailed view of I157F, T401F, L192F, and L431F mutations and amino acids within 4 Å.

R478A/R479A is functional at 37°C and that these suppressor mutations confer the highest transport rates (Figure 3.17c). Alternatively, these mutations could restore UapA-mediated transport activity by enhancing interactions between the monomers. Finally, Leu^{234} faces the core domain and interacts mainly with Ile^{158} that is located in the core domain (TMS3). Thus L234M could increase the strength of interactions between the dimerization and core domains.

3.3.5 Arg^{133} and Arg^{421} are essential for ER exit, function, and lipid-dependent stability

MD simulations predicted that, apart from the dimer interface, lipids might also bind to other the cytosolic-facing residues located on the outside, membrane-facing part of the core domain of the UapA dimer. These residues could interact with non-annular/specific or annular lipids

and are Lys^{73}, Arg^{133}, Tyr^{137}, Lys^{138}, Lys^{212} and Arg^{421} [12]. Lys^{73} is in the N-terminus just upstream of TMS1, Arg^{133}, Tyr^{137}, and Lys^{138} in the L2 loop, Lys^{212} in helix 1H of L4 and Arg^{421} in the border of L10 with TMS11 (Figure 3.20a). Arg^{421} is highly conserved in all NATs, whereas Lys^{73}, Arg^{133}, and Lys^{212} are conserved only in fungal NATs. On the contrary Tyr^{137} and Lys^{138} are not conserved. To explore the potential role of these interactions we generated a range of single, double and triple mutants of the predicted binding residues with C-terminal GFP tags (K73A, R133A, Y137A, K212A, Y137A/K138A, K73A/R421A, K73A/R133A/R421A). We transformed these constructs into the ΔACZ *A. nidulans* strain. Most mutants grew normally on uric acid or xanthine, except from the double R133A/R421A and the triple K73A/R133A/R421A mutants that were unable to grow on either uric acid or xanthine, at either 25°C (not shown) or 37°C (Figure 3.20b). Fluorescence microscopy revealed that loss of transport activity of the double R133A/R421A and the triple K73A/R133A/R421A mutants is a result of impaired UapA trafficking as they showed almost complete and total ER-retention respectively (Figure 3.20b right panel). Additionally, among the functional mutant versions the single R133A and the double K73A/R133A and K73A/R421A mutants exhibited moderate ER-retention. Comparative uptake assays also confirmed the results of the growth tests and the microscopy as R133A/R421A and K73A/R133A/R421A showed no transport activity (Figure 3.20c). All the other mutants, except from the K73A/R133A which exhibited significantly reduced transport, showed transport rates similar to the wild-type. For the K73A/R133A mutant we also measured the K_m value for xanthine that was not affected suggesting that reduced transport is not assigned to reduced substrate binding. Overall, Arg^{133} and Arg^{421} proved very important for UapA function and sorting to the PM, while Lys^{73} was critical for trafficking, when present in the context of R421A, and function, when present in the context of R133A.

For the K73A/R133A/R421A mutant that exhibited total ER-retention we investigated its ability to dimerize. For this, we employed the BiFC assay, as described before for the R287A/R478A/R479A mutant and its suppressors. Figure 3.20d shows that 3 h after expression a weak YFP fluorescence signal was detected suggesting that the K73A/R133A/R421A mutant could apparently dimerize at the ER and thus dimerization is not sufficient for the trafficking process. After 4 h a stronger signal was visible at the ER network confirming again problematic sorting of this mutant. This finding together with the observation that reduction of dimerization of the R287A/R478A/R479A mutant at the ER did not affect proper ER exit and sorting to the PM suggests that functional dimerization of UapA and recognition by the trafficking mechanism are not necessarily related. We also tried to purify the

K73A/R133A/R421A protein in order to investigate the oligomeric states of UapA by native MS as performed for the R287A/R478A/R479A mutant [12] but the protein was prone to aggregation and highly unstable (Figure 3.21). The instability observed upon purification in contrast with the stability in total extracts may suggest that instability is a result of lipids removal.

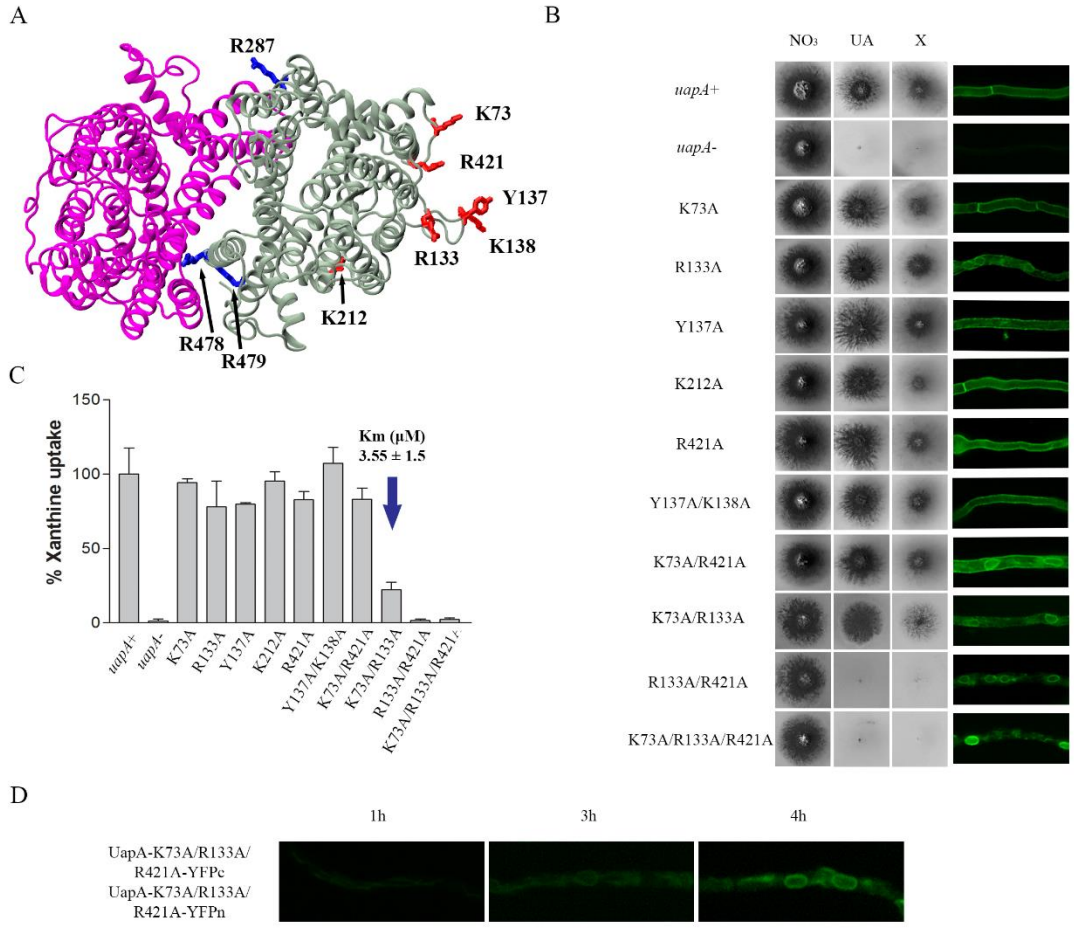


Figure 3.20 Arg^{133} and Arg^{421} are essential for ER exit, function, and lipid-dependent stability. (A) Topology of the predicted lipid-binding sites near the dimer interface (blue) and those in the membrane-facing regions (red) of the UapA dimer. (B) Growth tests of *A. nidulans* strains in minimal media supplemented with nitrate (NO_3), uric acid (UA), or xanthine (X) as a nitrogen source at 37°C (left panel). Control strains and concentrations of supplements are as in Figure 3.17. All UapA mutant strains are isogenic to the negative and positive control strains, expressing *uapA* alleles tagged with gfp. Inverted fluorescence microscopy images show localization of the GFP-tagged UapA constructs (right panel). (C) Relative ^{3}H-xanthine transport rates of UapA mutants expressed as percentages of initial uptake rates (V) compared to the wild-type (*uapA+)* rate. 3H-xanthine uptakes were performed at 37°C. The K_m value (μM) for xanthine for mutant K73A/R133A is indicated by a blue arrow. The results are averages of three measurements for each concentration point. SD was 20%. (D) Bimolecular complementation analysis of the K73A/R133A/R421A UapA mutant, performed as previously described in Figure 3.16 and Figure 3.18. Notice that the K73A/R133A/R421A mutant retains the ability to reconstitute a fluorescent signal in the ER, but not in the plasma membrane.

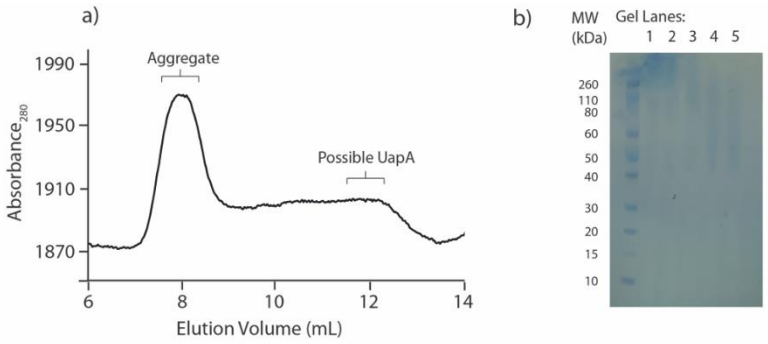


Figure 3.21. Purification of UapA K73A/R133A/R421A (a) SEC profile of the mutant. Fractions used for SDS-PAGE gel lanes are indicated. (b) SDS-PAGE gel from SEC. Gel lanes 1-2: Aggregate; 3-5: Possible UapA. Note the smear at the top of lanes 1 and 2 due to aggregated protein.

3.3.6 Substitution T401P partially restores K73A/R133A/R421A sorting to the PM and function

In order to expand our understanding on the possible core-lipid interactions, we performed UV mutagenesis, as described before, to obtain suppressors that restore growth of the triple K73A/R133A/R421A mutant on uric acid. We obtained 10 suppressors that all contained the same mutation, namely T401P (Figure 3.22a), which was isolated among the suppressors of the R287A/R478A/R479A mutant. Based on these findings and in order to explore the role of T401 in UapA function, we also constructed by directed mutagenesis the T401P mutation, in a wild-type uapA genetic context. Figure 3.22b shows that both T401P and K73A/R133A/R421A/T401P were able to grow on uric acid and xanthine, both at 25°C (not shown) and 37°C. Additionally, none of the mutants conferred growth on adenine or hypoxanthine. Fluorescence microscopy revealed that T401P suppressed partially K73A/R133A/R421A sorting to the PM and transport activity as single copy transformants of K73A/R133A/R421A/T401P still showed a major degree of ER-retention (Figure 3.22b right panel) and significantly reduced transport rates (Figure 3.22c). However high copy transformants confirmed that trafficking to the PM and transport were significantly restored. T401P by itself did not affect UapA localization to the PM and behaved nearly as well as a wild-type UapA (~70% transport rates).

As the K73A/R133A/R421A protein was unstable upon purification, we decided to see if we were able to isolate the K73A/R133A/R421A/T401P protein. Indeed, K73A/R133A/R421A/T401P protein was fairly stable but native MS showed that was mostly monomeric with only a small amount of the protein forming dimers (Figure 3.23). This is inconsistent with the fact that the K73A/R133A/R421A mutant was able to dimerize based on the results of the BiFC assay, as this would mean that T401P in the context of K73A/R133A/R421A also results in dimer unfolding. However, at the same time T401P

restored R287A/R478A/R479A dimerization, thus the definite effect of the T401P mutation in respect to K73A/R133A/R421A dimerization remains unclear.

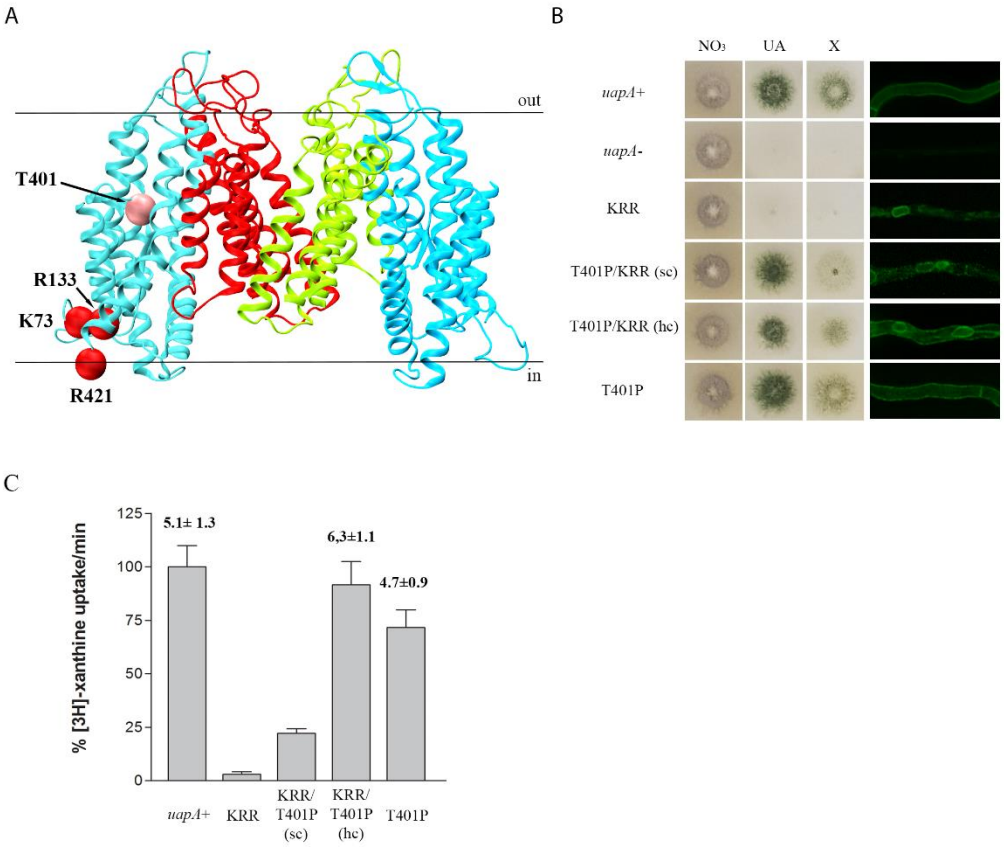


Figure 3.22 Substitution T401P partially restores the lipid-dependent functional defects of K73A/R133A/R421A. (A) Topologies of amino acids modified in K73A/R133A/R421A suppressors. Core domains are colored light blue, and dimerization domains red and green. Mutated amino acids in the original strain are shown with red spheres and in the suppressor with pink spheres. (B) Growth tests of K73A/R133A/R421A suppressors in minimal media supplemented with nitrate (NO_3), uric acid (UA), and xanthine (X) as nitrogen sources at 37°C (left panel). Control strains and supplement concentrations are as in previous Figures. All mutants are isogenic to the negative and positive control strains, expressing uapA alleles tagged with GFP. Inverted fluorescence microscopy images show localization of the GFP-tagged UapA constructs (right panel). KRR depicts K73A/R133A/R421A, sc depicts single copy transformants and hc depicts high copy transformants. (C) Relative ^{3}H-xanthine transport rates of UapA mutant versions expressed as percentages of initial uptake rates (V) compared to the wild-type (uapA+) rate. ^{3}H-xanthine uptakes were performed at 37°C. K_m values (μM) for xanthine are shown. Results are averages of three measurements for each concentration point. SD was 20%.

3.3.7 Conclusions

It is becoming clear that the composition of lipid bilayers affects many aspects of transporter life such as folding, oligomerization state, turnover and function [106,115–117,165–168]. Previous studies on the eukaryotic transporter UapA revealed that removal of tightly bound lipids causes dissociation of the dimer into a monomer. This result was reversed by addition

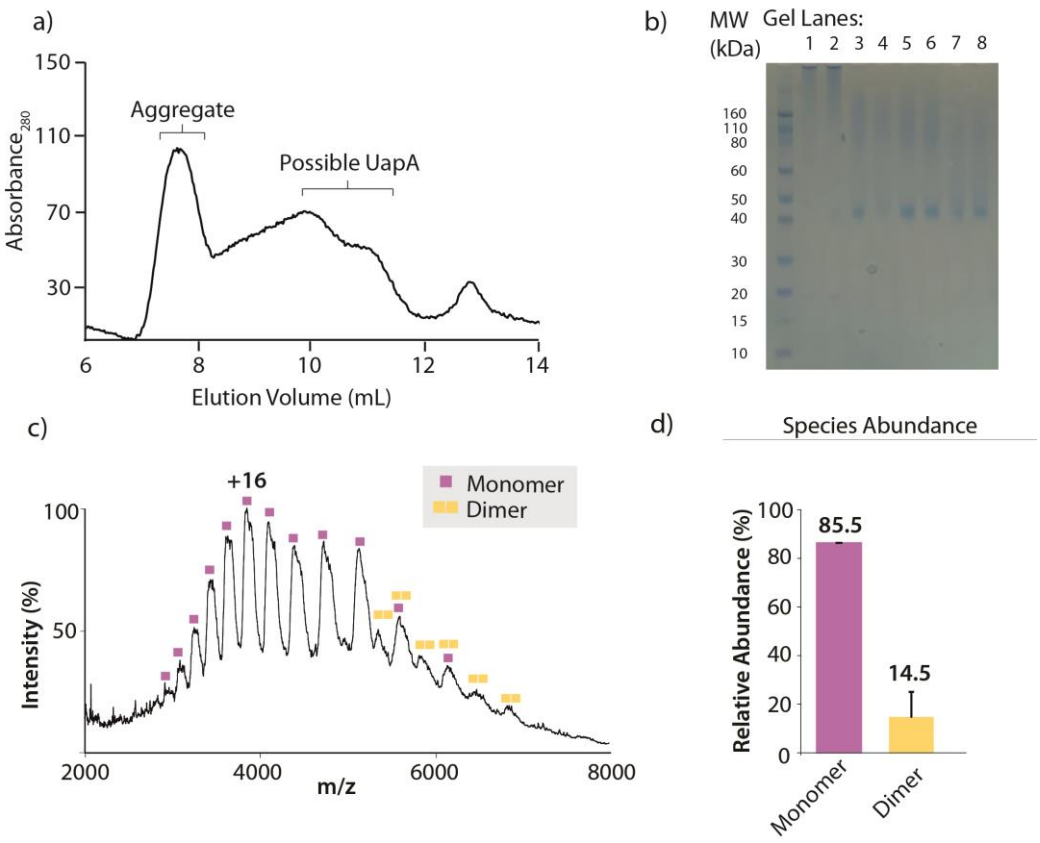


Figure 3.23. Purification and native MS of UapA K73A/R133A/R421A/T401P (a) SEC profile of the mutant. Fractions used for SDS-PAGE gel lanes are indicated with labels. (b) SDS-PAGE gel from SEC. Gel lanes 1-2: Aggregate; 3-8: Possible UapA. The peak at 13 mL is due to DDM eluting from the column. (c) Native mass spectrum of UapA K73A/R133A/R421A/T401P with monomer and dimer peaks indicated. (d) Relative abundance of monomer and dimer was calculated using UniDec [126]. Average abundance was calculated from three repeats carried out under identical conditions in the mass spectrometer.

of PIs or PEs. MD simulations predicted lipid binding sites that located in different regions of the protein. Mutagenesis of the predicted lipid-binding site at the dimer interface that is formed by the residues Arg^{287}, Arg^{478}, and Arg^{479}, abolished UapA dimerization and function. Here we decided to understand more about the role of the residues Arg^{287}, Arg^{478}, and Arg^{479} in UapA dimer stability and function and explore the potential role of other predicted lipid-binding residues located in the opposite side of the core domain. Genetic suppressors of the triple R287A/R478A/R479A mutant that restored UapA-mediated transport activity mapped both in the core and the dimerization domain. MD proposed that suppressors in the core domain enhanced hydrophobic interactions increasing stability while suppressors in the dimerization domain could either increase the strength of interactions between the two domains or replace interactions with lipids of the mutated Arg^{287}. In parallel, Lys^{73}, Arg^{133}, and Arg^{421}, exposed proved to be essential for ER exit and sorting to the PM, but not for the initial formation of dimers at the ER. Problematic sorting of the K73A/R133A/R421A mutant was

rescued only by the mutation T401P that surprisingly restored both ER exit of the K73A/R133A/R421A mutant and dimerization of the R287A/R478A/R479A mutant. It still remains quite unclear how T401P suppress problematic ER exit, as the molecular basis of the trafficking defect is not understood at the moment. However we can hypothesize that loss of core-lipid interactions results in structural destabilization that is partially restored by enhancing the network of the hydrophobic interactions that stabilize the two transmembrane layers of the core domain.

3.4 Expression of mammalian NAT homologues in *A. nidulans*

Includes results from Kourkoulou, A., Grevias, P., Lambrinidis, G., Pyle, E., Dionysopoulou, M., Politis, A., Mikros, E., Byrne, B. & Diallinas, G. ***Genetics.*** *213, 1357-1372 (2019).*

3.4.1 Rationale of the study

Knowledge on structure function relationships of mammalian transporters is limited despite their biological and medical importance. This is mainly due to the technical complications that accompany their study. Fungi offer many advantages as heterologous expression hosts for mammalian membrane proteins as they are much easier to handle. Both *A. nidulans* and *S. cerevisiae* have been successfully used to express plant solute transporters [169–172]. However long-standing efforts of our group and many others have failed to functionally express metazoan solute transporters in model fungi, as they are retained at the ER probably because of the differences in the membrane environment or recognition inability by the trafficking mechanism. Here we made use of the information we obtained about transporter-lipid interactions described in the previous chapter and try to achieve the functional expression of metazoan NAT transporters in *A. nidulans*.

3.4.2 Manipulation of a residue topologically equivalent to T401P leads to functional expression of a mammalian NAT homolog in *A. nidulans*

In order to understand more about the evolution of substrate specificity in the NAT family, we tried to functionally express characterized NAT homologues in *A. nidulans* and study structure-function relationships. At first, we decided to express the well characterized NAT homolog from rat, rSNBT1 that is a rather promiscuous Na^+ symporter specific for pyrimidines (uracil

and thymine) and purines (hypoxanthine, guanine, xanthine, and uric acid) [4,173]. Previous attempts to express rSNBT1 or rSNBT1/UapA chimeric transporters in *A. nidulans* or *S. cerevisiae* have failed, always due to total ER retention (C. Gournas, S. Amillis, B. Byrne, and G. Diallinas, unpublished data). Furthermore, attempts, via isolation of genetic suppressors, to restore rSNBT1 ER-exit and translocation into the PM were unsuccessful. Based on the fact that Thr^{401} (T401P) proved to be a key residue in restoring defects in UapA subcellular sorting and function that may be related to loss of lipid interactions, we thought that by manipulating the topologically equivalent to Thr^{401} residue in rSNBT1 we might achieve its functional expression. For this reason we built a structural model of rSNBT1 by homology threading using the available UapA crystal structure and found that the equivalent residue to Thr^{401} of UapA is Asn^{390} (Figure 3.24a). Notably, we found that this residue is conserved in all metazoan NAT members irrespective of their specificity and what clearly distinguishes metazoans from fungi in this region is the replacement of the Asn^{390} with a Thr and the Gly^{391} with a Pro.

Based on the above we transformed Δ7 *A. nidulans* strains with rSNBT1 versions containing substitutions N390T, N390P and G391P expressed via the *gpdA* promoter. Figure 3.24b shows that similarly to the negative control (Δ7), strains expressing wild-type rSNBT1 or the N390P and G391P mutants did not grow on purines and were resistant to 5FU but the strain expressing the N390T mutant was obviously sensitive to 5FU. Notably, 5FU sensitivity of this strain was Na^+ dependent, compatible with the physiological mechanism of rSNBT1 functioning, as sensitivity was reversed when we significantly reduce Na^+ in the medium (Figure 3.24c). Notice that in growth tests presented in Figure 3.24c we used proline as a non-repressing N source in order to avoid addition of Na^+ present in $NaNO_3$ that was used in all other experiments. Moreover, sensitivity was significantly inhibited by the presence of excess hypoxanthine, uracil and thymine that are considered substrates of rSNBT1 but not adenine (Figure 3.24b). Contrary to the above observations, rSNBT1-N390T did not confer growth on any purine tested (see later). Uptake assays using radiolabeled uracil further confirmed the previous results as N390T increased uracil accumulation compared with wild-type rSNBT1 although transport capacity was low (Figure 3.24d). Furthermore, uracil uptake was Na^+ dependent and was inhibited by excess uracil, hypoxanthine, guanine and thymine but not xanthine and adenine while inhibition by uric acid was moderate. Xanthine failed to inhibit 5FU sensitivity and uracil accumulation even though is also considered as an rSNBT1 substrate. Considering all the previous results this can be due to the different membrane environment (i.e. heterologous expression) which can affect structure and kinetics.

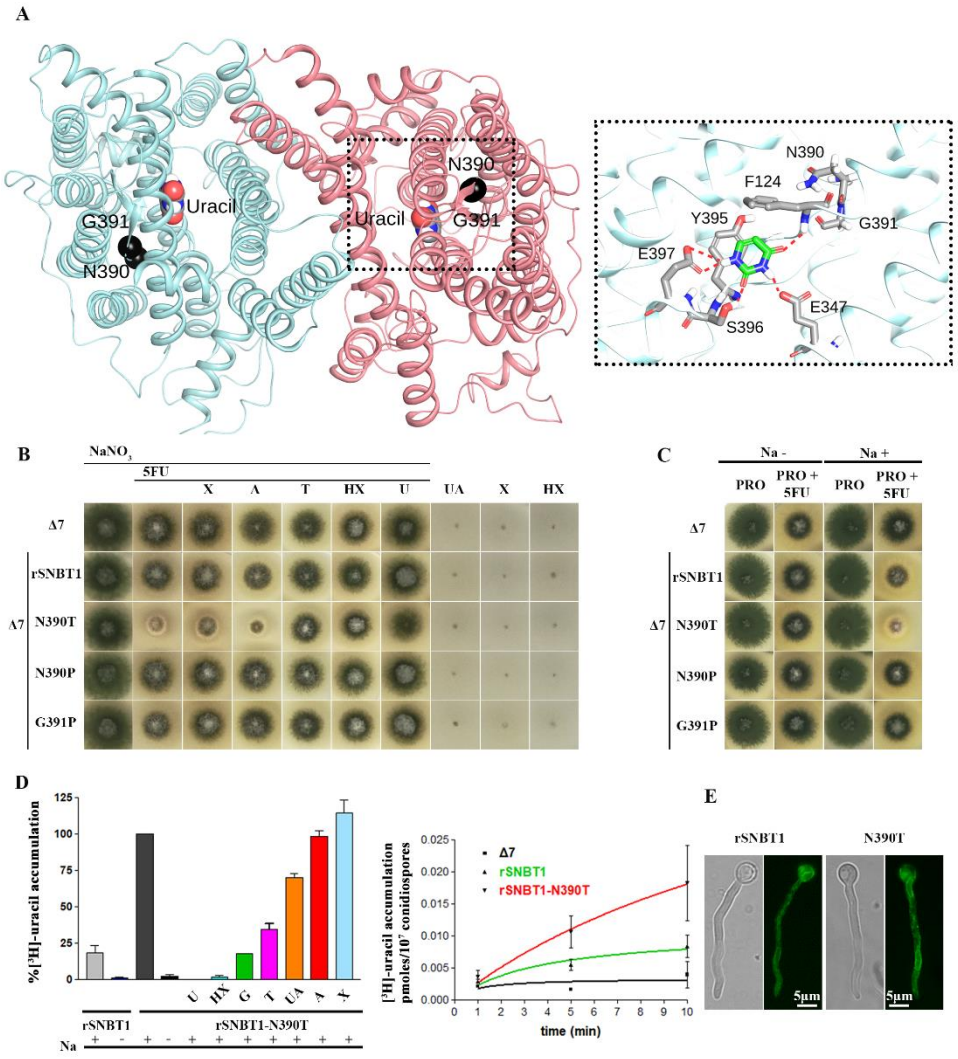


Figure 3.24. Manipulation of a residue topologically equivalent to T401P leads to functional expression of a mammalian NAT homolog in *A. nidulans*. (A) Homology modeling of the topology of rSNBT1, constructed using, as described in the Materials and Methods, the inward-facing conformation of the crystal structure of the UapA dimer. The two mutated and functionally analyzed residues, Asn^{390} and Gly^{391}, are shown as black spheres. The location of uracil, the major substrate of rSNBT1, is also depicted, as determined by dynamic docking (left panel). In the right panel, a zoomed-out picture of the substrate-binding site depicting the major interactions of uracil with specific residues. (B) Growth tests of isogenic *A. nidulans* strains expressing a single copy wild-type rSNBT1 or its mutated versions rSNBT1-N390T, rSNBT1-N390P, and rSNBT1-G391P. A negative control strain (i.e., the recipient Δ7 strain that has null activity for nucleobase transport; see text) is included for comparison. Growth tests were performed at 37°C on minimal media supplemented with Na^+ (100 mM NaCl); 10 mM $NaNO_3$ was used as a control nitrogen source unrelated to purine transport activities in all tests scoring resistance/sensitivity to 5FU (rows 1–7). Rows 3–7 represent *in vivo* competition assays scoring the ability of excess purines (2 mM) to compete with the uptake of 5FU (100 μM), and thus revert 5FU sensitivity. X is xanthine, A is adenine, T is Thymine, HX is hypoxanthine, and U is uracil. Notice that T, HX and U competed with 5FU uptake and suppressed sensitivity. Growth was also scored on minimal media containing UA (uric acid), X, or HX as sole nitrogen sources, none of which supported growth of the strains tested (three last rows). (C) Growth tests of *A. nidulans* of the same strains as in (A), on minimal media plus proline as a sole nitrogen source, supplemented or not with 100 mM NaCl. Notice that rSNBT1-N390T-mediated 5FU sensitivity is dependent on the presence of Na^+ supplementation. Notice also that in the presence of proline as a nitrogen source, the wild-type rSNBT1 allele confers very

moderate sensitivity to 5FU. (D) Left panel: ^{3}H-uracil (0.1 µM) accumulation in strains expressing rSNBT1 and rSNBT1-N390T, performed in the presence or absence of 100 mM Na^{+}, and in the presence or absence of excess (2 mM) unlabeled nucleobases, after a period of 10 min incubation with radiolabeled substrate. ^{3}H-uracil accumulation in rSNBT1-N390T in the presence of 100 mM Na^{+} and absence of unlabeled nucleobase is arbitrarily taken as 100%. Right panel: relative ^{3}H-uracil transport accumulation in Δ7 (negative control), rSNBT1, or rSNBT1-N390T strains as a time course. Uptake results are averages of three measurements for each concentration point. SD was 20%. (E) Inverted fluorescence microscopy images showing the subcellular localization of the GFP-tagged rSNBT1 and rSNBT1-N390T constructs. Notice that the strains used for microscopy are identical to those used in growth tests and uptake assays, as in all cases rSNBT1 sequences were tagged C-terminally with GFP (see Materials and Methods).

Epifluorescence microscopy showed that N390T restored rSNBT1 sorting partially, as it still showed a major degree of ER-retention (Figure 3.24e).

To further confirm that the phenotype observed in the corresponding transformants of rSNBT1-N390T is due to the transporter protein, we analyzed the meiotic progeny of an rSNBT1-N390T transformant. *A. nidulans* undergoing meiosis during a process called "selfing" [84] are prone to high recombination rates that often lead to the loss of sequences introduced by transformation. Analysis of 28 meiotic progenital colonies of an original rSNBT1-N390T transformant showed that all 21 colonies that conserved the original sensitivity to 5FU exhibited a GFP signal (not shown) from rSNBT1-N390T (Figure 3.25). On the contrary, the other 7 colonies that appeared to be 5FU-resistant lost the fluorescent signal confirming that 5FU toxicity is rSNBT1 dependent.

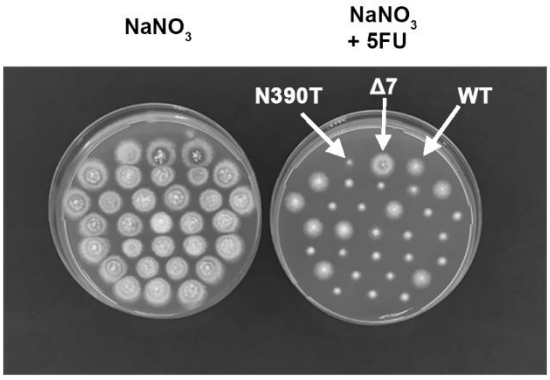


Figure 3.25. Co-segregation of 5-FU sensitivity with genetic loss of rSNBT1-N390T activity. The figure shows a growth test on minimal media, supplemented with 100 $NaNO_3$ mM or 100 mM $NaNO_3$ plus 100 µM 5-fluorouracil (5-FU) of 29 meiotic progeny colonies originating from individual ascospores from a 'selfed' cleistothecium of a strain that contained a genome-integrated single plasmid copy of rSNBT1-N390T. Isogenic control strains are a strain with total deletions in all major purine transporters (Δ7 negative control), a Δ7 transformant expressing rSNBT1-N390T-gfp and a standard wild-type strain expressing all known purine transporters. Notice that 8 colonies are resistant to 5-FU, contrasting the sensitivity of the parental strain. These colonies were subsequently shown to have lost the rSNBT1-N390T sequences, apparently due to genetic excision of integrated plasmid sequences, a frequent meiotic phenomenon in *A. nidulans*.

We also tried to functionally express the ascorbate-specific NAT from human hSVCT1 carrying similar mutations with the ones that allowed successful functional expression rSNBT1 in *A. nidulans*. The topologically equivalent critical residue (i.e the one that corresponds to Thr^{401} in UapA) in hSVCT1 is Asn^{377}. We thus constructed and analyzed Δ7 *A. nidulans* strains expressing wild-type hSVCT1 or mutated versions with substitutions N377T and N377P expressed via the *gpdA* promoter. Figure 3.26 shows that in this case the Asn^{377} substitutions did not confer growth on purines or sensitivity to 5FU. Additionally, they did not confer any growth phenotype on media containing L-ascorbate (not shown). *In vivo* epifluorescence microscopy showed that hSVCT1 remains trapped in the ER network.

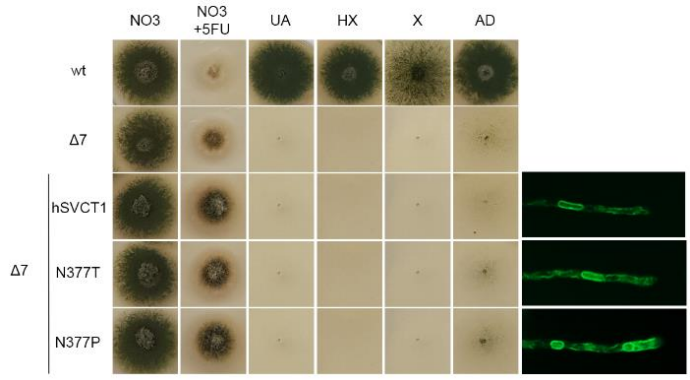


Figure 3.26. Manipulation of the residue topologically equivalent to T401P in hSVCT1. Details are as described in Figure 3.25.

3.4.3 Conclusions

Eukaryotic systems such as model fungi are preferred for heterologous expression of mammalian proteins due to the similarities in transcription and translation mechanisms and basic cellular composition compared to prokaryotic cells. However, functional characterization via heterologous expression is often hampered by mislocalization or degradation resulting in no or low expression levels. Many strategies have been implemented in order to achieve functional expression of mammalian membrane proteins in yeast, most focusing in modifying the lipid composition of yeast cells towards the synthesis of cholesterol instead of ergosterol. In UapA, mutation T401P restored trafficking of a UapA mutant that was predicted to have defective interactions with lipids. Using this information we achieved, for the first time, the functional expression of a rat homolog of UapA, rSNBT1 in *A. nidulans*. This supports the idea that the membrane lipid composition of the expression system is in principle the reason for non-proper folding and ER-retention. Furthermore, this achievement opens the way for further manipulations regarding residues involved in lipid interactions or residues that by-pass the need for lipid interactions like Thr^{401} in UapA.

Discussion

4.1 Evolution of L-ascorbate transporters

A major goal of this study was to reveal the changes needed in order for members of the same transporter family (NAT) to be specific for substrates with such different structural and chemical formulas. Nucleobases are planar molecules with relatively low solubility, whereas L-ascorbate is non-planar and extremely soluble. Thus, the corresponding transporters should have major differences in the architecture of the substrate binding site and the substrate translocation pathway. We decided to address this issue by making a mutational analysis on UapA, that is nucleobase specific, introducing residues found in L-ascorbate specific transporters. Study of L-ascorbate specificity is difficult in *A. nidulans* as it cannot be used as nitrogen or carbon source in order to design genetic screens selecting L-ascorbate specific mutants. Additionally, radiolabeled L-ascorbic acid is expensive and unstable, not permitting performance of many experiments. Therefore, we made use of indirect ways to explore L-ascorbate binding and transport which indicated at least better recognition by the mutants. In respect to purine recognition our analysis showed that UapA-TSSP, lost any detectable transport activity for xanthine and uric acid or purines in general probably due to replacement of residues critical for purine recognition (Phe^{406} and Gln^{408}). Additionally, UapA-SSSQ also showed very little purine transport activity, despite conserving $Gln^{408.}$ Thus the introduction of a series of polar Ser residues, as seen in L-ascorbate specific members, upstream of Gln^{408} is detrimental for purine transport, at least within the context of an otherwise wild-type UapA sequence.

In addition, our present analysis supports that NAT neo- functionalization took place via progressive improvement of a sub-function of ancestral NAT transporters (i.e. very low affinity binding of L-ascorbate by UapA). A duplication event of a NAT transporter in a vertebrate ancestor and further diversification of the one copy towards L-ascorbate seems to have resulted in generation of L-ascorbate transporters. This idea leads to many questions as for example if there are present or extinct transporters with double specificity that could represent intermediate versions. These versions may have specificities other than nucleobases or L-ascorbate or no function at all? We still don't know the specificity of the SVCT3 clade. Based on the crucial role of the NAT signature motif in substrate binding and specificity, its diversification in respect to specificity and the results of this study, we believe

that NATs containing transitory variations of this motif might reflect these intermediate steps in the evolution of L-ascorbate transporters.

The selective pressures that contributed to this event may include evolution of L-ascorbate biosynthesis, which seems that in the group of metazoa 'appeared' in vertebrates. Interestingly teleost fishes (along with guinea pigs, most bats, some bird species, as well as humans and many other primates) have lost the ability to synthesize vitamin C while all cartilaginous and non-teleost bony fishes maintain it [174]. Animals use one of the three known pathways for L-ascorbate synthesis, while photosynthetic eukaryotes (plants, algae and some single-celled organisms) employ the other two. These observations give rise to more questions. For example, whether transporters can acquire the ability to recognize a new substrate before or after it first "appears", and if this can be related to the evolution of biosynthetic pathways. A recent study showed that transporters of the NPF family that transport both cyanogenic glucosides and glucosinolates existed before plants evolved the ability to synthesize glucosinolates and subsequently, these transporters became glucosinolate-specific [175]. A similar case was also reported in vertebrate corticoid receptors where corticoid and glucocorticoid receptors evolved from a dual specificity receptor recognizing aldosterone and cortisol before aldosterone biosynthesis, possibly as a by-product of affinity towards chemically similar ligands [176,177]. These findings support the idea that transporters or receptors might evolve together with their substrate biosynthetic pathways.

NAT evolution could be also correlated with purine catabolism and more specifically to uric acid oxidase. The loss of uric acid oxidase which oxidizes uric acid to allantoin (that is excreted from the body) in higher primates parallels the similar loss of L-ascorbate biosynthesis [178]. This suggests that accumulation of uric acid may partially substitute L-ascorbate loss as these molecules have similar but not identical properties i.e. both are strong reducing agents and potent antioxidants [179]. It has been suggested that the elevation of uric acid levels in blood plasma due to this gene loss contributed to the development of the brain in primates and human intelligence as a result of its protective role against oxidation related to neurotoxic activity [180]. Additionally primates have also lost nucleobase-specific NATs. A recent study supported the co-evolution of the uric acid transporter URAT1 (distinct from NAT transporters) and uric acid oxidase during primate evolution [181]. Understanding the evolution of substrate specificity in the NAT family but also of other families will shed light into these questions and theories.

4.2 Context-dependent cryptic roles of residues in UapA substrate specificity

In order to better understand the process of evolution of substrate specificity in the NAT family we decided to extend our mutational analysis of the NAT signature motif. Additionally, we decided to re-address the role of residues affecting specificity that map outside the substrate binding site. An important novelty of this study was the use of the Δ7 *A. nidulans* strain for the study of NAT transporters. This strain genetically lacks not only a wild-type UapA endogenous copy, but also all other major transporters related to purine, pyrimidine, nucleoside or allantoin transport. Consequently, Δ7 has no transport activity for purine-related substrates, so that any UapA mutant introduced by transformation could be analyzed for its ability to transport any purine substrates in a 'clean' background.

The analysis of the mutants revealed context dependent cryptic roles of located in the first part of the NAT signature motif. Importantly, introducing the SYSE sequence, which is found in many metazoan NATs of the SVCT4/SLC23A4 clade, in UapA results in a transporter with no transport capacity for its physiological substrates or any purine tested, but high capacity for 5FU accumulation. Notably, bacterial uracil transporters conserve a Tyr and Glu residues in these positions. MDs confirmed binding of uracil in UapA-SYSE which showed remarkably similar binding structure with that in its natural transporter UraA. This binding architecture is not seen in wild-type UapA that does not transport uracil or 5FU. The loss of transport activity towards purines was due to the combination of Tyr^{406} and Glu^{408} suggesting a negative epistatic interactions between those two residues. The important context-dependent role of Glu^{408} in UapA specificity is corroborated by comparing UapA-SYSE and UapA-SYSQ, the former being rather specific for 5FU accumulation and the latter promiscuous (transport of purines-uracil-5FU).

MDs also revealed that the presence of the F406Y mutation affects the calculated pKa of acidic residues such as the invariable Glu^{356} and Asp^{360} in TMS8 and Glu^{408} in TMS10. More specifically in UapA-SYSE Glu^{356} is considered protonated, while Asp^{360} is thought to be negatively charged. Interestingly, MDs performed in UraA in order to investigate the role of Glu^{241} and His^{245} (analogous to Glu^{356} and Asp^{360}) in H^+ translocation, revealed that protonation of Glu^{241} resulted in disruption of the hydrogen bond network keeping uracil bound [3]. This could explain partially why UapA-SYSE scores as a loss of function mutant in respect to purines as alterations, by Tyr^{406}, in the protonation status of residues implicated in substrate binding in combination with the presence of Glu^{408} may disrupt binding of specific substrates. Finally,

our MD analysis also revealed possible interactions that rationalize why UapA-TFAE binds, but does not transport hypoxanthine.

UapA-SYSE loss of transport capacity for purines was reversed by mutations in residues Phe^{528}, Ala^{519} and Leu^{234} that led to a transporter able to translocate uracil or 5FU, in addition to purines. This suggests that these residues interact functionally with the NAT motif, a conclusion made also before at least in the case of Phe^{528} (its role will be further discussed in the next chapter). This residue firstly 'appeared' in a genetic screen searching for mutations restoring UapA-TFAE (Q408E) cryosensitivity [95]. Except from Phe^{528}, UapA-TFAE cryosensitivity was also reversed by mutations in residues Gln^{113} and Thr^{526} [10] but, as shown in this study, Thr^{526} failed to reverse UapA-SYSE loss of transport activity for purines. Interestingly, one of the mutations in Phe^{528} that 'rescued' UapA-SYSE, namely F528M, was a mutation rationally designed based on the conservation of a Met residue at this position in metazoan NATs. In conclusion, these observations clearly show that UapA specificity is determined by complex interactions of residues both inside and outside of the substrate binding site. Thus probably the key for engineering L-ascorbate transport capacity in UapA may be the combination of mutations in the substrate binding site towards L-ascorbate with mutations in distant residues known already to enlarge specificity. Given the limitations of our system for detecting low-affinity L-ascorbate transport, we cannot exclude the possibility that we already have mutants (e.g. UapA-SSSP) with an ameliorated capacity for recognizing this metabolite.

4.3 The role of Phe^{528} and other specificity residues- 'locks' or 'brakes' of transport dynamics

Mutations in residue Phe^{528} were isolated several times in this work but also previously. As mentioned before a mutation in this residue (F528S) firstly "appeared" in 2001 as a second-site suppressor of UapA-TFAE cryosensitivity [95]. F528S did not affect binding affinity for UapA physiological substrates but converted UapA into a promiscuous, low-affinity (regarding other purines), high-capacity purine transporter. In 2006 it was further shown that also distinct Phe^{528} substitutions (e.g. Ala, Ser, Met Asn) can confer a similar 'promiscuity' effect, with Tyr^{528} being the only exception leading to no change in UapA specificity, suggesting that removal of the aromatic ring is responsible for enlargement of specificity [154]. In addition, several Phe^{528} substitutions led to a moderate increase in UapA-mediated initial uptake rate

of transport of xanthine. None of these Phe^{528} mutations affected the H^+-coupled mechanism of transport or expression in the plasma membrane. Based on these results it was proposed that Phe^{528} is not an element of the *bona fidae* purine-binding site, but rather acts as a *selectivity filter* that somehow controls substrate translocation by 'selecting' which substrates reach the substrate binding site.

In 2008 two more residues that enlarged UapA specificity mapping outside the substrate binding site were selected [10]. These were Gln^{113} and Thr^{526}, the latter being very close to Phe^{528}, and again they were not associated with significant changes in UapA substrate binding affinities, but only moderately increased transport rates. In 2010 four more residues with similar results were selected (Ala^{441}, Val^{463}, Ala^{469} and Arg^{481}) [9]. One of them, Arg^{481}, was predicted to be a cytoplasm-facing residue. The UapA crystal structure showed later that this residue is close to the substrate binding site of the opposite protomer [2]. Combinations of Arg^{481} substitutions with substitutions in residues located towards the extracellular face of the transporter, such as those concerning Thr^{526} and Phe^{528}, were additive (i.e. more efficient transport of non-native substrates). The topology of the residues together with their additive nature suggested that these amino acids may actually be *selective gates* or *gating elements* that act along the substrate translocation trajectory. The term *gate* is usually associated with ion channels, which have a rather rigid and continuous transmembrane pore rather a specific binding site and thus the function of gates (*gating*) is needed to allow or restrict the passage of ions. However, after 2008, several crystal structures were published in 'occluded' or 'open' outward- and inward-facing conformations, suggesting the existence of domains or loops that do act as dynamic gates controlling the access and release of substrates form transporter binding sites.

The present work provides novel findings concerning how Phe^{528} might function as a key amino acid in determining UapA specificity. The modelling of the outward topology of UapA revealed that Phe^{528} approaches very close and probably interacts via pi-pi stacking with Glu^{356}, which is a major substrate-binding residue. This interaction is apparently dynamic, as it is not present in the inward-facing UapA crystal. This observation led us to form the following hypothesis. In the apo state (i.e. no substrate bound) a Phe^{528}-Glu^{356} interaction may 'lock' UapA in its outward topology. When the substrate (and probably H^+) is available it interacts with Glu^{356} and Gln^{408} resulting in loss of the Phe^{528}-Glu^{356} interaction and this shift allows UapA to shift to its inward topology via sliding of the elevator domain. Thus, replacement of the aromatic residue in the Phe^{528} position can genetically 'unlock' the sliding mechanism in a way that becomes less dependent on substrate/H^+ binding, and thus convert

UapA into a more promiscuous transporter for low-affinity substrates. This means that the other specificity mutations shown previously, but also here (e.g. Leu^{234} and Ala^{519}), might similarly modify the sliding mechanism converting UapA into a more promiscuous transporter for weakly binding substrates. More structural data and relative MDs in several distinct topologies are needed for understanding how specificity of UapA, and probably of other elevator-type transporters, is determined.

This hypothesis for the function of Phe^{528} opens new insights into the substrate translocation mechanism of elevator-type transporters and how UapA specificity is modified. Based on this scenario, strong binding of UapA physiological substrates seems to be the key for unlocking the sliding of the elevator. In the absence of substrates or when low-affinity substrates enter the binding pocket UapA remains in the outward topology as their binding is not "strong" enough to break Phe^{528}-Glu^{356} and maybe other interactions. Thus, Phe^{528} and probably other residues that map in the interface between the two domains and are involved in specificity such as Val^{463}, Ala^{469}, Arg^{481} and Thr^{526} may work as "*locks*" or "*brakes*" that prevent sliding of the elevator in the absence of tightly bound native substrates. Therefore, mutations in these residues might "loosen" the sliding mechanism in a way that becomes independent of tight substrate binding. This scenario is further supported by both the increase in xanthine transport rates seen in mutants containing substitutions of these residues and the additive phenotype of their combinations.

4.4 Lipid interactions are critical for UapA dimerization and ER exit

The study of protein-lipid interactions and especially transporter-lipid interactions is an emerging theme. The recent study on UapA-lipid interactions was among the first focusing on a eukaryotic transporter. In brief, it has been shown that UapA, which primarily exists as a dimer, dissociates into monomers after removal of lipids that co-purify with it, and that the dimer can be recovered by the addition of PI or PE and more efficiently of both. Mutagenesis of a tentative lipid-binding site at the dimer interface formed by three Arg residues 287, 478, and 479, predicted by MDs, abolished lipid binding and function [12]. Thus, it was proposed that binding of PI and PE at specific sites in the dimer interface stabilizes the dimeric functional form of UapA. Here, we decided to study further the role of this specific lipid binding site and in parallel investigate the potential role of residues at the periphery of UapA, predicted also by MDs to bind lipids.

At first we found that Arg^{287}, Arg^{478}, and Arg^{479} are essential for early *de novo* formation in the ER membrane and this is not essential for further trafficking to the plasma membrane (PM). In parallel, predicted lipid binding peripheral residues that map in the core domain of UapA (Lys^{73}, Arg^{133}, and Arg^{421}), are essential for stability, ER exit and sorting to the PM, but apparently not essential for the initial formation of dimers in the ER. Thus, the two sets of positively charged residues that define two distinct lipid binding sites are both essential for function, albeit due to different reasons. However, none of the mutants studied elicited an unfolded protein response (results not shown), suggesting that they probably do not lead to significant misfolding. This observation may also indicate that lipid interactions at the periphery of UapA are needed for the trafficking mechanism.

The most original findings of this work stems from the isolation of genetic suppressors restoring defects of the two loss of function mutants, R287A/R478A/R479A and K73A/R133A/R421A. In both cases all suppressors concerned intragenic mutations maybe due to fact that the only known factors affecting UapA sorting to the PM, COPII components, actin and clathrin heavy chain that are essential for cell viability and thus mutations in their genes might affect the trafficking of other essential cargoes. In the first case suppressor mutations can be classified into three types. Type I, containing the majority of suppressors, map in the center of the core domain in TMS2, TMS3, TMS4, TMS9, TMS10, and TMS11 and all, except one, introduce residues with increased hydrophobicity and/or aromaticity (V150I, I157F, I157L, L192F, A396P, T401P, T401F, or L431F). Only V153T introduces a polar residue, while S119T replaces a polar residue with a bigger residue with similar properties. Type II includes L234M in TMS5 in the middle plane of the dimerization domain. Finally, Type III includes E286Q and E286K at the end of the cytoplasmic-facing part of TMS6 next to the originally mutated Arg^{287}. In contrast, suppressors of K73A/R133A/R421A all concerned a single mutation, T401P, isolated also as a suppressor mutation of R287A/R478A/R479A.

MDs shed some light into how these mutations can restore dimerization or/and trafficking to the PM. Type I suppressors seem to increase the strength of relative TMS interactions in the core domain. Thus, probably they increase compactness and stability of the core domain. The type II mutation concerns a residue that maps in the dimerization domain. Interestingly this change was also selected as a suppressor mutation of the SYSE mutant (analyzed previously). This mutation can increase the strength of interactions between the dimerization and the core domain. How such a mutation can affect both specificity and dimerization opens new issues on the role of dimerization in finely regulating specificity. The most logical scenario for the two type III suppressor mutations is that they replace directly the

interactions with lipids of the nearby mutated Arg^{287} (i.e., in R287A). This is also in line with the fact that these were the suppressors with the highest xanthine transport activity. Additionally, BiFC assays showed that at least three selected suppressors I157F, L234M and T401P restored R287A/R478A/R479A dimerization and thus do not lead to functional UapA monomers. Finally, regarding K73A/R133A/R421A we do not understand, at the moment, how T401P restores trafficking of this mutant, mostly because we still do not understand the molecular basis of the trafficking defect in the original mutant. Noticeably, in XanQ, a neighboring residue (Pro^{318} equivalent to Pro^{402} in UapA) has been shown to be important for insertion/stability of the protein in the membrane [92]. Thus probably this region of the protein is important for stability.

In conclusion, our findings and especially those concerning type I and II suppressors of the R287A/R478A/R479A mutant strongly suggest that by stabilizing the core, the dimer is also stabilized and thus function is restored. In line with this, the core domain is the motile part of the monomeric units that undergoes dynamic up-and down elevator-like sliding in continuous dynamic interactions with membrane lipids, and thus prone to external forces. Additionally, the inability of the K73A/R133A/R421A mutant to reach the plasma membrane suggests that the predicted lipid interactions at the membrane-facing regions of the UapA core domain are important for either stabilization or the trafficking mechanism. In line with our results, a previous study on the prokaryotic NhaA Na^+/H^+ exchanger showed that lipid binding around domain interfaces is also involved in stabilizing the core domain during the conformational transitions required for transport by the elevator mechanism [116]. Therefore, it has been proposed that elevator-type antiporters may use a subset of peripheral or annular lipids as structural support to facilitate large-scale conformational changes within the membrane.

4.5 The successful functional expression of rSNBT1 in *A. nidulans* leads to new routes for NAT study

An important result of the present work was the functional expression of a metazoan NAT transporter for the first time in *A. nidulans*. This was achieved by using information gained during the study of UapA-lipid interactions. T401P proved to be a key substitution in restoring trafficking impairment due to predicted loss of interactions with lipids. Based on this observation we substituted the equivalent residue in the rat homolog, rSNBT1 that as all

metazoan transporters expressed in *A. nidulans* until now was “trapped” at the ER network. The N390T mutation led to the functional expression of this homolog indicating that the bottleneck in expressing metazoan transporters in fungi is proper folding in an environment of heterologous membrane lipid composition. However, it should be mentioned that this mutation only slightly improved rSNBT1 sorting to the PM and thus probably other substitutions are needed in order for a larger amount of the protein to reach the PM. Mutations in the other residues affecting core stability may be the case.

Unfortunately, introduction of the same mutation in another NAT homolog from human, hSVCT1, did not have the same effect. However, based on the fact that the effect of this mutation in rSNBT1 localization was negligible based on *in vivo* epifluorescence microscopy and on the limitations of our system for detecting L-ascorbate transport, we cannot exclude the possibility that also hSVCT1 functional expression is improved. Functional expression of rSNBT1 opens the way for further manipulations, via rational design or unbiased genetic screens, of NAT transporters for achieving their functional expression and manipulation in *A. nidulans*. This in turn will lead to new routes for studying the evolution of substrate specificity in the NAT family. Overall, the idea that manipulating lipid binding residues results in functional heterologous expression of transporters opens the way for heterologous expression of transporters also from other families.

References

[1] C. Gournas, I. Papageorgiou, G. Diallinas, The nucleobase–ascorbate transporter (NAT) family: genomics, evolution, structure–function relationships and physiological role, Mol. Biosyst. 4 (2008) 404. https://doi.org/10.1039/b719777b.

[2] Y. Alguel, S. Amillis, J. Leung, G. Lambrinidis, S. Capaldi, N.J. Scull, G. Craven, S. Iwata, A. Armstrong, E. Mikros, G. Diallinas, A.D. Cameron, B. Byrne, Structure of eukaryotic purine/H+ symporter UapA suggests a role for homodimerization in transport activity, Nat. Commun. 7 (2016) 11336. https://doi.org/10.1038/ncomms11336.

[3] X. Yu, G. Yang, C. Yan, J.L. Baylon, J. Jiang, H. Fan, G. Lu, K. Hasegawa, H. Okumura, T. Wang, E. Tajkhorshid, S. Li, N. Yan, Dimeric structure of the uracil:proton symporter UraA provides mechanistic insights into the SLC4/23/26 transporters, Cell Res. 27 (2017) 1020–1033. https://doi.org/10.1038/cr.2017.83.

[4] S. Yamamoto, K. Inoue, T. Murata, S. Kamigaso, T. Yasujima, J.Y. Maeda, Y. Yoshida, K.Y. Ohta, H. Yuasa, 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 (2010) 6522–6531. https://doi.org/10.1074/jbc.M109.032961.

[5] M. Bürzle, Y. Suzuki, D. Ackermann, H. Miyazaki, N. Maeda, B. Clémençon, R. Burrier, M.A. Hediger, The sodium-dependent ascorbic acid transporter family SLC23, Mol. Aspects Med. 34 (2013) 436–454. https://doi.org/10.1016/j.mam.2012.12.002.

[6] F. Lu, S. Li, Y. Jiang, J. Jiang, H. Fan, G. Lu, D. Deng, S. Dang, X. Zhang, J. Wang, N. Yan, Structure and mechanism of the uracil transporter UraA, Nature. 472 (2011) 243–246. https://doi.org/10.1038/nature09885.

[7] G. Diallinas, Dissection of Transporter Function: From Genetics to Structure, Trends Genet. 32 (2016) 576–590. https://doi.org/10.1016/j.tig.2016.06.003.

[8] M. Koukaki, A. Vlanti, S. Goudela, A. Pantazopoulou, H. Gioule, S. Tournaviti, G. Diallinas, The Nucleobase-ascorbate transporter (NAT) signature motif in UapA defines the function of the purine translocation pathway, J. Mol. Biol. 350 (2005) 499–513. https://doi.org/10.1016/j.jmb.2005.04.076.

[9] V. Kosti, I. Papageorgiou, G. Diallinas, 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 (2010) 1132–1143. https://doi.org/10.1016/j.jmb.2010.02.037.

[10] I. Papageorgiou, C. Gournas, A. Vlanti, S. Amillis, A. Pantazopoulou, G. Diallinas, Specific Interdomain Synergy in the UapA Transporter Determines Its Unique Specificity for Uric Acid among NAT Carriers, J. Mol. Biol. 382 (2008) 1121–1135. https://doi.org/10.1016/j.jmb.2008.08.005.

[11] V. Kosti, G. Lambrinidis, V. Myrianthopoulos, G. Diallinas, E. Mikros, Identification of the substrate recognition and transport pathway in a eukaryotic member of the nucleobase-ascorbate transporter (NAT) family, PLoS One. 7 (2012). https://doi.org/10.1371/journal.pone.0041939.

[12] E. Pyle, A.C. Kalli, S. Amillis, Z. Hall, A.M. Lau, A.C. Hanyaloglu, G. Diallinas, B. Byrne, A. Politis, Structural Lipids Enable the Formation of Functional Oligomers of the Eukaryotic Purine Symporter UapA, Cell Chem. Biol. 25 (2018) 840-848.e4. https://doi.org/10.1016/j.chembiol.2018.03.011.

[13] E. Krypotou, V. Kosti, S. Amillis, V. Myrianthopoulos, E. Mikros, G. Diallinas, 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 (2012) 36792–36803. https://doi.org/10.1074/jbc.M112.400382.

[14] G. Diallinas, Understanding transporter specificity and the discrete appearance of channel-like gating domains in transporters, Front. Pharmacol. 5 (2014) 207. https://doi.org/10.3389/fphar.2014.00207.

[15] M.A. Hediger, B. Clémençon, R.E. Burrier, E.A. Bruford, The ABCs of membrane transporters in health and disease (SLC series): Introduction, Mol. Aspects Med. 34 (2013) 95–107. https://doi.org/10.1016/j.mam.2012.12.009.

[16] A. César-Razquin, B. Snijder, T. Frappier-Brinton, R. Isserlin, G. Gyimesi, X. Bai, R.A. Reithmeier, D. Hepworth, M.A. Hediger, A.M. Edwards, G. Superti-Furga, A Call for Systematic Research on Solute Carriers, Cell. 162 (2015) 478–487. https://doi.org/10.1016/j.cell.2015.07.022.

[17] J.C. DiFrancesco, D. DiFrancesco, Dysfunctional HCN ion channels in neurological

diseases, Front. Cell. Neurosci. 9 (2015). https://doi.org/10.3389/fncel.2015.00071.

[18] Y. Moran, M.G. Barzilai, B.J. Liebeskind, H.H. Zakon, Evolution of voltage-gated ion channels at the emergence of Metazoa, J. Exp. Biol. 218 (2015) 515–525. https://doi.org/10.1242/jeb.110270.

[19] G.M. Cooper, R.E. Hausman, The Cell: A Molecular Approach., Sinauer Associates, Sunderland, MA, 2009.

[20] Y. Shi, Common folds and transport mechanisms of secondary active transporters, Annu. Rev. Biophys. 42 (2013) 51–72. https://doi.org/10.1146/annurev-biophys-083012-130429.

[21] M.H. Saier, V.S. Reddy, B. V. Tsu, M.S. Ahmed, C. Li, G. Moreno-Hagelsieb, The Transporter Classification Database (TCDB): Recent advances, Nucleic Acids Res. 44 (2016) D372–D379. https://doi.org/10.1093/nar/gkv1103.

[22] S. Ohno, Evolution by gene duplication, Springer-Verlag, Berlin Heidelberg, 1970.

[23] J.S. Taylor, J. Raes, Duplication and divergence: The evolution of new genes and old ideas, Annu. Rev. Genet. 38 (2004) 615–643. https://doi.org/10.1146/annurev.genet.38.072902.092831.

[24] C.T. Hittinger, S.B. Carroll, Gene duplication and the adaptive evolution of a classic genetic switch, Nature. 449 (2007) 677–681. https://doi.org/10.1038/nature06151.

[25] F.A. Kondrashov, A.S. Kondrashov, Role of selection in fixation of gene duplications, J. Theor. Biol. 239 (2006) 141–151. https://doi.org/10.1016/j.jtbi.2005.08.033.

[26] S.C. Wang, P. Davejan, K.J. Hendargo, I. Javadi-Razaz, A. Chou, D.C. Yee, F. Ghazi, K.J.K. Lam, A.M. Conn, A. Madrigal, A. Medrano-Soto, M.H. Saier, Expansion of the Major Facilitator Superfamily (MFS) to include novel transporters as well as transmembrane-acting enzymes, Biochim. Biophys. Acta - Biomembr. 1862 (2020) 183277. https://doi.org/10.1016/j.bbamem.2020.183277.

[27] E.M. Quistgaard, C. Löw, F. Guettou, P. Nordlund, Understanding transport by the major facilitator superfamily (MFS): Structures pave the way, Nat. Rev. Mol. Cell Biol. 17 (2016) 123–132. https://doi.org/10.1038/nrm.2015.25.

[28] Y. Huang, M.J. Lemieux, J. Song, M. Auer, D.N. Wang, Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli, Science (80-.). 301 (2003)

616–620. https://doi.org/10.1126/science.1087619.

[29] J. Abramson, I. Smirnova, V. Kasho, G. Verner, H.R. Kaback, S. Iwata, Structure and mechanism of the lactose permease of Escherichia coli, Science (80-.). 301 (2003) 610–615. https://doi.org/10.1126/science.1088196.

[30] H. Kumar, A.B. Waight, A.J. Risenmay, Z. Roe-, B.H. Chau, A. Schlessinger, M. Bonomi, W. Harries, A. Sali, A.K. Johri, R.M. Stroud, S. Francisco, T. Sciences, Crystal structure of a eukaryotic phosphate transporter, 496 (2013) 533–536. https://doi.org/10.1038/nature12042.Crystal.

[31] J. Sun, J.R. Bankston, J. Payandeh, T.R. Hinds, W.N. Zagotta, N. Zheng, Crystal structure of the plant dual-affinity nitrate transporter NRT1.1, Nature. 507 (2014) 73–77. https://doi.org/10.1038/nature13074.

[32] D. Deng, C. Xu, P. Sun, J. Wu, C. Yan, M. Hu, N. Yan, Crystal structure of the human glucose transporter GLUT1, Nature. 510 (2014) 121–125. https://doi.org/10.1038/nature13306.

[33] D. Drew, O. Boudker, Shared Molecular Mechanisms of Membrane Transporters, Annu. Rev. Biochem. 85 (2016) 543–572. https://doi.org/10.1146/annurev-biochem-060815-014520.

[34] V.S. Reddy, M.A. Shlykov, R. Castillo, E.I. Sun, M.H.S. Jr, The major facilitator superfamily (MFS) revisited, 279 (2022) 2022–2035. https://doi.org/10.1111/j.1742-4658.2012.08588.x.

[35] M. Niño-González, E. Novo-Uzal, D.N. Richardson, P.M. Barros, P. Duque, More Transporters, More Substrates: The Arabidopsis Major Facilitator Superfamily Revisited, Mol. Plant. 12 (2019) 1182–1202. https://doi.org/10.1016/j.molp.2019.07.003.

[36] J.L. Parker, S. Newstead, Molecular basis of nitrate uptake by the plant nitrate transporter NRT1.1, Nature. 507 (2014) 68–72. https://doi.org/10.1038/nature13116.

[37] F.H. Wong, J.S. Chen, V. Reddy, J.L. Day, M.A. Shlykov, S.T. Wakabayashi, M.H. Saier, The amino acid-polyamine-organocation superfamily, J. Mol. Microbiol. Biotechnol. 22 (2012) 105–113. https://doi.org/10.1159/000338542.

[38] A. Vastermark, S. Wollwage, M.E. Houle, R. Rio, M.H. Saier, Expansion of the APC superfamily of secondary carriers, Proteins Struct. Funct. Bioinforma. 82 (2014)

2797–2811. https://doi.org/10.1002/prot.24643.

[39] A. Yamashita, S.K. Singh, T. Kawate, Y. Jin, E. Gouaux, Crystal structure of a bacterial homologue of Na+/Cl --dependent neurotransmitter transporters, Nature. 437 (2005) 215–223. https://doi.org/10.1038/nature03978.

[40] J. Abramson, E.M. Wright, Structure and function of Na+-symporters with inverted repeats, Curr. Opin. Struct. Biol. 19 (2009). https://doi.org/10.1016/j.sbi.2009.06.002.

[41] Å. Västermark, M.H. Saier, Evolutionary relationship between 5+5 and 7+7 inverted repeat folds within the amino acid-polyamine-organocation superfamily, Proteins Struct. Funct. Bioinforma. 82 (2014) 336–346. https://doi.org/10.1002/prot.24401.

[42] H. De Koning, G. Diallinas, Nucleobase transporters (Review), Mol. Membr. Biol. 17 (2000) 75–94. https://doi.org/10.1080/09687680050117101.

[43] E. Krypotou, T. Evangelidis, J. Bobonis, A.A. Pittis, T. Gabaldón, C. Scazzocchio, E. Mikros, G. Diallinas, Origin, diversification and substrate specificity in the family of NCS1/FUR transporters, Mol. Microbiol. 96 (2015) 927–950. https://doi.org/10.1111/mmi.12982.

[44] K. Papakostas, M. Botou, S. Frillingos, 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 (2013) 36827–36840. https://doi.org/10.1074/jbc.M113.523340.

[45] T. Arakawa, T. Kobayashi-Yurugi, Y. Alguel, H. Iwanari, H. Hatae, M. Iwata, Y. Abe, T. Hino, C. Ikeda-Suno, H. Kuma, D. Kang, T. Murata, T. Hamakubo, A.D. Cameron, T. Kobayashi, N. Hamasaki, S. Iwata, Crystal structure of the anion exchanger domain of human erythrocyte band 3, Science (80-.). 350 (2015) 680–684. https://doi.org/10.1126/science.aaa4335.

[46] J. O., Simple allosteric model for membrane pumps., Nature. 211 (1966) 969–970.

[47] N.K. Karpowich, D.N. Wang, Structural biology: Symmetric transporters for asymmetric transport, Science (80-.). 321 (2008) 781–782. https://doi.org/10.1126/science.1161495.

[48] N. Yan, Structural advances for the major facilitator superfamily (MFS) transporters, Trends Biochem. Sci. 38 (2013) 151–159. https://doi.org/10.1016/j.tibs.2013.01.003.

[49] P. Mitchell, Osmochemistry of solute translocation, (n.d.).

[50] P. Mitchell, A general theory of membrane transport from studies of bacteria, Nature. 180 (1957) 134–136. https://doi.org/10.1038/180134a0.

[51] Y. Xu, Y. Tao, L.S. Cheung, C. Fan, L.Q. Chen, S. Xu, K. Perry, W.B. Frommer, L. Feng, Structures of bacterial homologues of SWEET transporters in two distinct conformations, Nature. 515 (2014) 448–452. https://doi.org/10.1038/nature13670.

[52] J. Wang, C. Yan, Y. Li, K. Hirata, M. Yamamoto, N. Yan, Q. Hu, Crystal structure of a bacterial homologue of SWEET transporters, Cell Res. 24 (2014) 1486–1489. https://doi.org/10.1038/cr.2014.144.

[53] Y. Lee, T. Nishizawa, K. Yamashita, R. Ishitani, O. Nureki, Structural basis for the facilitative diffusion mechanism by SemiSWEET transporter, Nat. Commun. 6 (2015). https://doi.org/10.1038/ncomms7112.

[54] N. Yan, Structural Biology of the Major Facilitator Superfamily Transporters, Annu. Rev. Biophys. 44 (2015) 257–283. https://doi.org/10.1146/annurev-biophys-060414-033901.

[55] P.W. Fowler, M. Orwick-Rydmark, S. Radestock, N. Solcan, P.M. Dijkman, J.A. Lyons, J. Kwok, M. Caffrey, A. Watts, L.R. Forrest, S. Newstead, Gating topology of the proton-coupled oligopeptide symporters, Structure. 23 (2015) 290–301. https://doi.org/10.1016/j.str.2014.12.012.

[56] L.S. Stelzl, P.W. Fowler, M.S.P. Sansom, O. Beckstein, Flexible gates generate occluded intermediates in the transport cycle of LacY, J. Mol. Biol. 426 (2014) 735–751. https://doi.org/10.1016/j.jmb.2013.10.024.

[57] M. Andersson, A.N. Bondar, J.A. Freites, D.J. Tobias, H.R. Kaback, S.H. White, Proton-coupled dynamics in lactose permease, Structure. 20 (2012) 1893–1904. https://doi.org/10.1016/j.str.2012.08.021.

[58] C.J. Law, J. Almqvist, A. Bernstein, R.M. Goetz, Y. Huang, C. Soudant, A. Laaksonen, S. Hovmöller, D.N. Wang, Salt-bridge Dynamics Control Substrate-induced Conformational Change in the Membrane Transporter GlpT, J. Mol. Biol. 378 (2008). https://doi.org/10.1016/j.jmb.2008.03.029.

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

215–223. https://doi.org/10.1038/nature03978.

[60] L.R. Forrest, G. Rudnick, The rocking bundle: A mechanism for ion-coupled solute flux by symmetrical transporters, Physiology. 24 (2009). https://doi.org/10.1152/physiol.00030.2009.

[61] A. V. Pedersen, T.F. Andreassen, C.J. Loland, A conserved salt bridge between transmembrane segments 1 and 10 constitutes an extracellular gate in the dopamine transporter, J. Biol. Chem. 289 (2014). https://doi.org/10.1074/jbc.M114.586982.

[62] H. Krishnamurthy, E. Gouaux, X-ray structures of LeuT in substrate-free outward-open and apo inward-open states, Nature. 481 (2012). https://doi.org/10.1038/nature10737.

[63] A.A. Garaeva, D.J. Slotboom, Elevator-type mechanisms of membrane transport, Biochem. Soc. Trans. 48 (2020) 1227–1241. https://doi.org/10.1042/BST20200290.

[64] A.A. Garaeva, A. Guskov, D.J. Slotboom, C. Paulino, A one-gate elevator mechanism for the human neutral amino acid transporter ASCT2, Nat. Commun. 10 (2019). https://doi.org/10.1038/s41467-019-11363-x.

[65] X. Yu, O. Plotnikova, P.D. Bonin, T.A. Subashi, T.J. McLellan, D. Dumlao, Y. Che, Y.Y. Dong, E.P. Carpenter, G.M. West, X. Qiu, J.S. Culp, S. Han, Cryo-EM structures of the human glutamine transporter SLC1a5 (ASCT2) in the outward-facing conformation, Elife. 8 (2019). https://doi.org/10.7554/eLife.48120.

[66] V. Arkhipova, A. Guskov, D.J. Slotboom, Structural ensemble of a glutamate transporter homologue in lipid nanodisc environment, Nat. Commun. 11 (2020). https://doi.org/10.1038/s41467-020-14834-8.

[67] J. Dechancie, I.H. Shrivastava, I. Bahar, The mechanism of substrate release by the aspartate transporter Glt Ph: Insights from simulations, Mol. Biosyst. 7 (2011). https://doi.org/10.1039/c0mb00175a.

[68] J.G. McCoy, Z. Ren, V. Stanevich, J. Lee, S. Mitra, E.J. Levin, S. Poget, M. Quick, W. Im, M. Zhou, The Structure of a Sugar Transporter of the Glucose EIIC Superfamily Provides Insight into the Elevator Mechanism of Membrane Transport, Structure. 24 (2016). https://doi.org/10.1016/j.str.2016.04.003.

[69] Y. Cao, X. Jin, E.J. Levin, H. Huang, Y. Zong, M. Quick, J. Weng, Y. Pan, J. Love, M. Punta, B. Rost, W.A. Hendrickson, J.A. Javitch, K.R. Rajashankar, M. Zhou, Crystal

structure of a phosphorylation-coupled saccharide transporter, Nature. 473 (2011). https://doi.org/10.1038/nature09939.

[70] Z. Ren, J. Lee, M.M. Moosa, Y. Nian, L. Hu, Z. Xu, J.G. McCoy, A.C.M. Ferreon, W. Im, M. Zhou, Structure of an EIIC sugar transporter trapped in an inward-facing conformation, Proc. Natl. Acad. Sci. U. S. A. 115 (2018). https://doi.org/10.1073/pnas.1800647115.

[71] M. Hirschi, Z.L. Johnson, S.Y. Lee, Visualizing multistep elevator-like transitions of a nucleoside transporter, Nature. 545 (2017). https://doi.org/10.1038/nature22057.

[72] A. Vergara-Jaque, C. Fenollar-Ferrer, C. Mulligan, J.A. Mindell, L.R. Forrest, Family resemblances: A common fold for some dimeric ion-coupled secondary transporters, J. Gen. Physiol. 146 (2015). https://doi.org/10.1085/jgp.201511481.

[73] N.J. Hu, S. Iwata, A.D. Cameron, D. Drew, Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT, Nature. 478 (2011) 408–411. https://doi.org/10.1038/nature10450.

[74] X. Zhou, E.J. Levin, Y. Pan, J.G. McCoy, R. Sharma, B. Kloss, R. Bruni, M. Quick, M. Zhou, Structural basis of the alternating-access mechanism in a bile acid transporter, Nature. 505 (2014). https://doi.org/10.1038/nature12811.

[75] M. Coincon, P. Uzdavinys, E. Nji, D.L. Dotson, I. Winkelmann, S. Abdul-Hussein, A.D. Cameron, O. Beckstein, D. Drew, Crystal structures reveal the molecular basis of ion translocation in sodium/proton antiporters, Nat. Struct. Mol. Biol. 23 (2016) 248–255. https://doi.org/10.1038/nsmb.3164.

[76] C. Lee, H.J. Kang, C. Von Ballmoos, S. Newstead, D.L. Dotson, S. Iwata, O. Beckstein, D. Alexander, D. Drew, A two-domain elevator mechanism for sodium / proton antiport, 501 (2014) 573–577. https://doi.org/10.1038/nature12484.A.

[77] M. Appel, D. Hizlan, K.R. Vinothkumar, C. Ziegler, W. Kühlbrandt, Conformations of NhaA, the Na/H Exchanger from Escherichia coli, in the pH-Activated and Ion-Translocating States, J. Mol. Biol. 386 (2009) 351–365. https://doi.org/10.1016/j.jmb.2008.12.042.

[78] C. Lee, S. Yashiro, D.L. Dotson, P. Uzdavinys, S. Iwata, M.S.P. Sansom, C. von Ballmoos, O. Beckstein, D. Drew, A.D. Cameron, Crystal structure of the sodium-proton antiporter NhaA dimer and new mechanistic insights, J. Gen. Physiol. 144

(2014) 529–544. https://doi.org/10.1085/jgp.201411219.

[79] J. Yin, W. Ren, X. Huang, J. Deng, T. Li, Y. Yin, Potential Mechanisms Connecting Purine Metabolism and Cancer Therapy, Front. Immunol. 9 (2018) 1697. https://doi.org/10.3389/fimmu.2018.01697.

[80] M. Kraupp, R. Marz, Membrane transport of nucleobases: Interaction with inhibitors, Gen. Pharmacol. 26 (1995) 1185–1190. https://doi.org/10.1016/0306-3623(95)00005-L.

[81] M. Kse, A.C. Schiedel, Nucleoside/nucleobase transporters: Drug targets of the future?, Future Med. Chem. 1 (2009) 303–326. https://doi.org/10.4155/fmc.09.29.

[82] P.M. Kirk, P.F. Cannon, D.W. Minter, J.A. Stalpers, Dictionary of the Fungi, 10th ed., Wallingford: CABI, 2008.

[83] F. Lutzoni, F. Kauff, C.J. Cox, D. McLaughlin, G. Celio, C. Dentinger, M. Padamsee, D. Hibbett, T.Y. James, E. Baloch, M. Grube, V. Reeb, V. Hofstetter, C. Schoch, A.E. Arnold, J. Miadlikowska, J. Spatafora, D. Johnson, S. Hambleton, M. Crockett, R. Shoemaker, G.H. Sung, R. Lücking, T. Lumbsch, K. O'Donnell, M. Binder, P. Diederich, D. Ertz, C. Gueidan, K. Hansen, R.C. Harris, K. Hosaka, Y.W. Lim, B. Matheny, H. Nishida, D. Pfister, J. Rogers, A. Rossman, I. Schmitt, H. Sipman, J. Stone, J. Sugiyama, R. Yahr, R. Vilgalys, Assembling the fungal tree of life: Progress, classification, and evolution of subcellular traits, Am. J. Bot. 91 (2004). https://doi.org/10.3732/ajb.91.10.1446.

[84] R.B. Todd, M.A. Davis, M.J. Hynes, Genetic manipulation of Aspergillus nidulans: meiotic progeny for genetic analysis and strain construction, Nat. Protoc. 2 (2007) 811–821. https://doi.org/10.1038/nprot.2007.112.

[85] J. Bruggeman, A.J.M. Debets, P.J. Wijngaarden, J.G.M. Arjan, R.F. Hoekstra, Sex Slows Down the Accumulation of Deleterious Mutations in the Homothallic Fungus Aspergillus nidulans, 485 (2003) 479–485.

[86] C. Scazzocchio, Aspergillus: A multifaceted genus, 2019. https://doi.org/10.1016/B978-0-12-809633-8.20803-0.

[87] E. Krypotou, G. Diallinas, Transport assays in filamentous fungi: Kinetic characterization of the UapC purine transporter of Aspergillus nidulans, Fungal Genet. Biol. 63 (2014) 1–8. https://doi.org/10.1016/j.fgb.2013.12.004.

[88] E. Karena, S. Frillingos, Role of intramembrane polar residues in the YgfO xanthine permease. HIS-31 and ASN-93 are crucial for affinity and specificity, and ASP-304 and GLU-272 are irreplaceable, J. Biol. Chem. 284 (2009) 24257–24268. https://doi.org/10.1074/jbc.M109.030734.

[89] A. Pantazopoulou, G. Diallinas, The first transmembrane segment (TMS1) of UapA contains determinants necessary for expression in the plasma membrane and purine transport, Mol. Membr. Biol. 23 (2006) 337–348. https://doi.org/10.1080/09687860600738239.

[90] S. Varma, C.E. Campbell, S.M. Kuo, Functional role of conserved transmembrane segment 1 residues in human sodium-dependent vitamin C transported, Biochemistry. 47 (2008) 2952–2960. https://doi.org/10.1021/bi701666q.

[91] G. Mermelekas, E. Georgopoulou, A. Kallis, M. Botou, V. Vlantos, S. Frillingos, 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 (2010) 35011–35020. https://doi.org/10.1074/jbc.M110.170415.

[92] P. Karatza, P. Panos, E. Georgopoulou, S. Frillingos, Cysteine-scanning analysis of the nucleobase-ascorbate transporter signature motif in YgfO permease of Escherichia coli: Gln-324 and Asn-325 are essential, and Ile-329 -Val-339 form an α-helix, J. Biol. Chem. 281 (2006) 39881–39890. https://doi.org/10.1074/jbc.M605748200.

[93] E. Georgopoulou, G. Mermelekas, E. Karena, S. Frillingos, 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 (2010) 19422–19433. https://doi.org/10.1074/jbc.M110.120543.

[94] S. Goudela, P. Karatza, M. Koukaki, S. Frillingos, G. Diallinas, Comparative substrate recognition by bacterial and fungal purine transporters of the NAT/NCS2 family, Mol. Membr. Biol. 22 (2005) 263–275. https://doi.org/10.1080/09687860500093016.

[95] S. Amillis, M. Koukaki, G. Diallinas, Substitution F569S converts UapA, a specific uric acid-xanthine transporter, into a broad specificity transporter for purine-related solutes, J. Mol. Biol. 313 (2001) 765–774. https://doi.org/10.1006/jmbi.2001.5087.

[96] K. Papakostas, E. Georgopoulou, S. Frillingos, Cysteine-scanning analysis of putative helix XII in the YgfO xanthine permease: ILE-432 and ASN-430 are important, J. Biol.

Chem. 283 (2008) 13666–13678. https://doi.org/10.1074/jbc.M800261200.

[97] S. Frillingos, 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 (2012) 250–272.

[98] B.H. Thurtle-Schmidt, R.M. Stroud, Structure of Bor1 supports an elevator transport mechanism for SLC4 anion exchangers, Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 10542–10546. https://doi.org/10.1073/pnas.1612603113.

[99] N. Coudray, S.L. Seyler, R. Lasala, Z. Zhang, K.M. Clark, M.E. Dumont, A. Rohou, O. Beckstein, D.L. Stokes, Structure of the SLC4 transporter Bor1p in an inward-facing conformation, 26 (2016) 130–145. https://doi.org/10.1002/pro.3061.

[100] K.W. Huynh, J. Jiang, N. Abuladze, K. Tsirulnikov, L. Kao, X. Shao, D. Newman, R. Azimov, A. Pushkin, Z.H. Zhou, I. Kurtz, CryoEM structure of the human SLC4A4 sodium-coupled acid-base transporter NBCe1, Nat. Commun. 9 (2018). https://doi.org/10.1038/s41467-018-03271-3.

[101] E.R. Geertsma, Y.N. Chang, F.R. Shaik, Y. Neldner, E. Pardon, J. Steyaert, R. Dutzler, Structure of a prokaryotic fumarate transporter reveals the architecture of the SLC26 family, Nat. Struct. Mol. Biol. 22 (2015) 803–808. https://doi.org/10.1038/nsmb.3091.

[102] C. Wang, B. Sun, X. Zhang, X. Huang, M. Zhang, H. Guo, X. Chen, F. Huang, T. Chen, H. Mi, F. Yu, L. Liu, P. Zhang, Structural mechanism of the active bicarbonate transporter from cyanobacteria, Nat. Plants. 5 (2019) 1184–1193. https://doi.org/10.1038/s41477-019-0538-1.

[103] J.D. Walter, M. Sawicka, R. Dutzler, Cryo-EM structures and functional characterization of murine Slc26a9 reveal mechanism of uncoupled chloride transport, Elife. 8 (2019) 1–31. https://doi.org/10.7554/eLife.46986.

[104] X. Chi, X. Jin, Y. Chen, X. Lu, X. Tu, X. Li, Y. Zhang, J. Lei, J. Huang, Z. Huang, Q. Zhou, X. Pan, Structural insights into the gating mechanism of human SLC26A9 mediated by its C-terminal sequence, Cell Discov. 6 (2020). https://doi.org/10.1038/s41421-020-00193-7.

[105] A.M. Velho, S.M. Jarvis, Topological studies of hSVCT1, the human sodium-dependent vitamin C transporter and the influence of N-glycosylation on its intracellular

targeting, Exp. Cell Res. 315 (2009) 2312–2321. https://doi.org/10.1016/j.yexcr.2009.04.007.

[106] C. Koshy, E.S. Schweikhard, R.M. Gärtner, C. Perez, Ö. Yildiz, C. Ziegler, Structural evidence for functional lipid interactions in the betaine transporter BetP, EMBO J. 32 (2013) 3096–3105. https://doi.org/10.1038/emboj.2013.226.

[107] J.R. Bolla, M.T. Agasid, S. Mehmood, C. V. Robinson, Membrane protein-lipid interactions probed using mass spectrometry, Annu. Rev. Biochem. 88 (2019) 85–111. https://doi.org/10.1146/annurev-biochem-013118-111508.

[108] D.L. Nelson, M.M. Cox, Lehninger Principles of Biochemistry, 6th ed., n.d.

[109] M. Pateiro, R. Domínguez, P.E.S. Munekata, F.J. Barba, J.M. Lorenzo, Innovative Thermal and Non-Thermal Processing, Bioaccessibility and Bioavailability of Nutrients and Bioactive Compounds, Woodhead Publishing Series in Food Science, Technology and Nutrition, 2019.

[110] M. Baginski, H. Resat, E. Borowski, Comparative molecular dynamics simulations of amphotericin B-cholesterol/ergosterol membrane channels, Biochim. Biophys. Acta - Biomembr. 1567 (2002) 63–78. https://doi.org/10.1016/S0005-2736(02)00581-3.

[111] A.G. Lee, Lipid-protein interactions, in: Biochem. Soc. Trans., 2011: pp. 761–766. https://doi.org/10.1042/BST0390761.

[112] H. Palsdottir, C. Hunte, Lipids in membrane protein structures, Biochim. Biophys. Acta - Biomembr. 1666 (2004). https://doi.org/10.1016/j.bbamem.2004.06.012.

[113] M. Frick, C. Schmidt, Mass spectrometry—A versatile tool for characterising the lipid environment of membrane protein assemblies, Chem. Phys. Lipids. 221 (2019) 145–157. https://doi.org/10.1016/j.chemphyslip.2019.04.001.

[114] J.M. East, D. Melville, A.G. Lee, Exchange Rates and Numbers of Annular Lipids for the Calcium and Magnesium Ion Dependent Adenosinetriphosphatase, Biochemistry. 24 (1985) 2615–2623. https://doi.org/10.1021/bi00332a005.

[115] A. Laganowsky, E. Reading, T.M. Allison, M.B. Ulmschneider, M.T. Degiacomi, A.J. Baldwin, C. V. Robinson, Membrane proteins bind lipids selectively to modulate their structure and function, Nature. 510 (2014) 172–175. https://doi.org/10.1038/nature13419.

[116] M. Landreh, E.G. Marklund, P. Uzdavinys, M.T. Degiacomi, M. Coincon, J. Gault, K. Gupta, I. Liko, J.L.P. Benesch, D. Drew, C. V. Robinson, Integrating mass spectrometry with MD simulations reveals the role of lipids in Na+/H+ antiporters, Nat. Commun. 8 (2017) 1–9. https://doi.org/10.1038/ncomms13993.

[117] K. Gupta, J.A.C. Donlan, J.T.S. Hopper, P. Uzdavinys, M. Landreh, W.B. Struwe, D. Drew, A.J. Baldwin, P.J. Stansfeld, C. V. Robinson, The role of interfacial lipids in stabilizing membrane protein oligomers, Nature. 541 (2017) 421–424. https://doi.org/10.1038/nature20820.

[118] E. Reading, C. V. Robinson, M.T. Degiacomi, T.M. Allison, A.J. Baldwin, M.B. Ulmschneider, A. Laganowsky, Membrane proteins bind lipids selectively to modulate their structure and function, Nature. 510 (2014) 172–175. https://doi.org/10.1038/nature13419.

[119] E. Käfer, Origins of translocations in Aspergillus nidulans., Genetics. 52 (1965). https://doi.org/10.1093/genetics/52.1.217.

[120] S. J, R. D, Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Laboratory Press, NY, 2001.

[121] A. Pantazopoulou, N.D. Lemuh, D.G. Hatzinikolaou, C. Drevet, G. Cecchetto, C. Scazzocchio, G. Diallinas, Differential physiological and developmental expression of the UapA and AzgA purine transporters in Aspergillus nidulans, Fungal Genet. Biol. 44 (2007) 627–640. https://doi.org/10.1016/j.fgb.2006.10.003.

[122] G. Diallinas, J. Valdez, V. Sophianopoulou, A. Rosa, C. Scazzocchio, Chimeric purine transporters of Aspergillus nidulans define a domain critical for function and specificity conserved in bacterial, plant and metazoan homologues, EMBO J. 17 (1998) 3827–3837. https://doi.org/10.1093/emboj/17.14.3827.

[123] N. Takeshita, Y. Higashitsuji, S. Konzack, R. Fischer, Apical Sterol-rich Membranes Are Essential for Localizing Cell End Markers That Determine Growth Directionality in the Filamentous Fungus Aspergillus nidulans, Mol. Biol. Cell. 19 (2008) 339–351. https://doi.org/10.1091/mbc.E07.

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

https://doi.org/10.1016/j.jmb.2015.05.021.

[125] M. Koukaki, E. Giannoutsou, A. Karagouni, G. Diallinas, A novel improved method for Aspergillus nidulans transformation, J. Microbiol. Methods. 55 (2003) 687–695. https://doi.org/10.1016/S0167-7012(03)00208-2.

[126] M.T. Marty, A.J. Baldwin, E.G. Marklund, G.K.A. Hochberg, J.L.P. Benesch, C. V. Robinson, Bayesian deconvolution of mass and ion mobility spectra: From binary interactions to polydisperse ensembles, Anal. Chem. 87 (2015) 4370–4376. https://doi.org/10.1021/acs.analchem.5b00140.

[127] S. Capella-gutiérrez, J.M. Silla-martínez, T. Gabaldón, trimAl : a tool for automated alignment trimming in large-scale phylogenetic analyses, 25 (2009) 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

[128] D. Darriba, G.L. Taboada, R. Doallo, D. Posada, ProtTest 3 : fast selection of best-fit models of protein evolution, 27 (2011) 1164–1165. https://doi.org/10.1093/bioinformatics/btr088.

[129] E. Wagenmakers, S. Farrell, AIC model selection using Akaike weights, 11 (2004) 192–196.

[130] A. Stamatakis, Using RAxML to Infer Phylogenies, Curr. Protoc. Bioinforma. 51 (2015) 6.14.1-6.14.14. https://doi.org/10.1002/0471250953.bi0614s51.

[131] J. Huerta-Cepas, F. Serra, P. Bork, ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data, Mol. Biol. Evol. 33 (2016) 1635–1638. https://doi.org/10.1093/molbev/msw046.

[132] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J.Y. Xiang, L. Wang, D. Lupyan, M.K. Dahlgren, J.L. Knight, J.W. Kaus, D.S. Cerutti, G. Krilov, W.L. Jorgensen, R. Abel, R.A. Friesner, OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins, J. Chem. Theory Comput. 12 (2016) 281–296. https://doi.org/10.1021/acs.jctc.5b00864.

[133] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. De Groot, H. Grubmüller, A.D. MacKerell, CHARMM36m: An improved force field for folded and intrinsically disordered proteins, Nat. Methods. 14 (2016) 71–73. https://doi.org/10.1038/nmeth.4067.

[134] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type

perception in molecular mechanical calculations, 25 (2006) 247–260. https://doi.org/10.1016/j.jmgm.2005.12.005.

[135] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and Testing of a General Amber Force Field, J. Comput. Chem. 56531 (2004) 1157–1174.

[136] M.A. Lomize, I.D. Pogozheva, H. Joo, H.I. Mosberg, A.L. Lomize, OPM database and PPM web server : resources for positioning of proteins in membranes, 40 (2012) 370–376. https://doi.org/10.1093/nar/gkr703.

[137] J.M. Abraham, T. Murtola, R. Schulz, S. Pall, J.C. Smith, B. Hess, E. Lindahl, GROMACS : High performance molecular simulations through multi-level parallelism from laptops to supercomputers, 2 (2015) 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

[138] D.J. Evans, B.L. Holian, The Nose–Hoover thermostat, 52 (1985) 4069–4074. https://doi.org/10.1063/1.449071.

[139] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190. https://doi.org/10.1063/1.328693.

[140] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092. https://doi.org/10.1063/1.464397.

[141] E.L. Wu, X. Cheng, S. Jo, H. Rui, K.C. Song, E.M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R.M. Venable, J.B. Klauda, W. Im, CHARMM-GUI membrane builder toward realistic biological membrane simulations, J. Comput. Chem. 35 (2014) 1997–2004. https://doi.org/10.1002/jcc.23702.

[142] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935. https://doi.org/10.1063/1.445869.

[143] J. Huang, A.D. Mackerell, CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data, J. Comput. Chem. 34 (2013) 2135–2145. https://doi.org/10.1002/jcc.23354.

[144] J. Lee, X. Cheng, J.M. Swails, M.S. Yeom, P.K. Eastman, J.A. Lemkul, S. Wei, J. Buckner, J.C. Jeong, Y. Qi, S. Jo, V.S. Pande, D.A. Case, C.L. Brooks, A.D. MacKerell, J.B. Klauda, W. Im, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and

CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field, J. Chem. Theory Comput. 12 (2016) 405–413. https://doi.org/10.1021/acs.jctc.5b00935.

[145] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

[146] M.P. Jacobson, D.L. Pincus, C.S. Rapp, T.J.F. Day, B. Honig, D.E. Shaw, R.A. Friesner, A Hierarchical Approach to All-Atom Protein Loop Prediction, Proteins Struct. Funct. Genet. 55 (2004) 351–367. https://doi.org/10.1002/prot.10613.

[147] J.L. Banks, H.S. Beard, Y. Cao, A.E. Cho, W. Damm, R. Farid, A.K. Felts, T.A. Halgren, D.T. Mainz, J.R. Maple, R. Murphy, D.M. Philipp, M.P. Repasky, L.Y. Zhang, B.J. Berne, R.A. Friesner, E. Gallicchio, R.M. Levy, Integrated Modeling Program, Applied Chemical Theory (IMPACT), J. Comput. Chem. 26 (2005) 1752–1780. https://doi.org/10.1002/jcc.20292.

[148] W. Sherman, T. Day, M.P. Jacobson, R.A. Friesner, R. Farid, Novel procedure for modeling ligand/receptor induced fit effects, J. Med. Chem. 49 (2006) 534–553. https://doi.org/10.1021/jm050540c.

[149] G. Diallinas, C. Gournas, Structure-function relationships in the nucleobase-ascorbate transporter (NAT) family: Lessons from model microbial genetic systems, Channels. 2 (2008) 363–372. https://doi.org/10.4161/chan.2.5.6902.

[150] S.R. Eddy, Accelerated profile HMM searches, PLoS Comput. Biol. 7 (2011). https://doi.org/10.1371/journal.pcbi.1002195.

[151] S.Q. Le, O. Gascuel, An improved general amino acid replacement matrix, Mol. Biol. Evol. 25 (2008) 1307–1320. https://doi.org/10.1093/molbev/msn067.

[152] E. Krypotou, G. Lambrinidis, T. Evangelidis, E. Mikros, G. Diallinas, Modelling, substrate docking and mutational analysis identify residues essential for function and specificity of the major fungal purine transporter AzgA, Mol. Microbiol. 93 (2014) 129–145. https://doi.org/10.1111/mmi.12646.

[153] C. Gournas, S. Amillis, A. Vlanti, G. Diallinas, Transport-dependent endocytosis and turnover of a uric acid-xanthine permease, Mol. Microbiol. 75 (2010) 246–260. https://doi.org/10.1111/j.1365-2958.2009.06997.x.

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

808–819. https://doi.org/10.1016/j.jmb.2005.12.070.

[155] G.F. Papadaki, S. Amillis, G. Diallinas, Substrate Speci fi city of the FurE Transporter Is Determined by Cytoplasmic Terminal Domain Interactions, 207 (2017) 1387–1400. https://doi.org/10.1534/genetics.117.300327/-/DC1.1.

[156] G. Sioupouli, G. Lambrinidis, E. Mikros, S. Amillis, G. Diallinas, Cryptic purine transporters in Aspergillus nidulans reveal the role of specific residues in the evolution of specificity in the NCS1 family, Mol. Microbiol. 103 (2017) 319–332. https://doi.org/10.1111/mmi.13559.

[157] S. Amillis, Z. Hamari, K. Roumelioti, C. Scazzocchio, G. Diallinas, Regulation of expression and kinetic modeling of substrate interactions of a uracil transporter in Aspergillus nidulans, Mol. Membr. Biol. 24 (2007) 206–214. https://doi.org/10.1080/09687860601070806.

[158] Z. Hamari, S. Amillis, C. Drevet, A. Apostolaki, C. Vágvölgyi, G. Diallinas, C. Scazzocchio, Convergent evolution and orphan genes in the Fur4p-like family and characterization of a general nucleoside transporter in Aspergillus nidulans, Mol. Microbiol. 73 (2009) 43–57. https://doi.org/10.1111/j.1365-2958.2009.06738.x.

[159] M. Sahin-Tóth, K.M. Akhoon, J. Runner, H.R. Kaback, Ligand recognition by the lactose permease of Escherichia coli: Specificity and affinity are defined by distinct structural elements of galactopyranosides, Biochemistry. 39 (2000) 5097–5103. https://doi.org/10.1021/bi0000263.

[160] M. Sahin-Tóth, P. Gunawan, M.C. Lawrence, T. Toyokuni, H.R. Kaback, Binding of hydrophobic D-galactopyranosides to the lactose permease of Escherichia coli, Biochemistry. 41 (2002) 13039–13045. https://doi.org/10.1021/bi0203076.

[161] K. Papakostas, S. Frillingos, Substrate selectivity of YgfU, a uric acid transporter from Escherichia coli, J. Biol. Chem. 287 (2012) 15684–15695. https://doi.org/10.1074/jbc.M112.355818.

[162] M. Botou, V. Yalelis, P. Lazou, I. Zantza, K. Papakostas, V. Charalambous, E. Mikros, E. Flemetakis, S. Frillingos, Specificity profile of NAT/NCS2 purine transporters in Sinorhizobium (Ensifer) meliloti, Mol. Microbiol. 114 (2020) 151–171. https://doi.org/10.1111/mmi.14503.

[163] R.A.F. Reithmeier, J.R. Casey, A.C. Kalli, M.S.P. Sansom, Y. Alguel, S. Iwata, Band 3, the

human red cell chloride/bicarbonate anion exchanger (AE1, SLC4A1), in a structural context, Biochim. Biophys. Acta - Biomembr. 1858 (2016) 1507–1532. https://doi.org/10.1016/j.bbamem.2016.03.030.

[164] S. Dimou, O. Martzoukou, M. Dionysopoulou, V. Bouris, S. Amillis, G. Diallinas, Translocation of nutrient transporters to cell membrane via Golgi bypass in Aspergillus nidulans , EMBO Rep. 21 (2020) 15–17. https://doi.org/10.15252/embr.201949929.

[165] E. Van Den Brink-Van Der Laan, J. Antoinette Killian, B. De Kruijff, Nonbilayer lipids affect peripheral and integral membrane proteins via changes in the lateral pressure profile, Biochim. Biophys. Acta - Biomembr. 1666 (2004) 275–288. https://doi.org/10.1016/j.bbamem.2004.06.010.

[166] C. Koshy, C. Ziegler, Structural insights into functional lipid-protein interactions in secondary transporters, Biochim. Biophys. Acta - Gen. Subj. 1850 (2015) 476–487. https://doi.org/10.1016/j.bbagen.2014.05.010.

[167] C. Martens, R.A. Stein, M. Masureel, A. Roth, S. Mishra, R. Dawaliby, A. Konijnenberg, F. Sobott, C. Govaerts, H.S. McHaourab, Lipids modulate the conformational dynamics of a secondary multidrug transporter, Nat. Struct. Mol. Biol. 23 (2016) 744–751. https://doi.org/10.1038/nsmb.3262.

[168] E. Henrich, O. Peetz, C. Hein, A. Laguerre, B. Hoffmann, J. Hoffmann, V. Dötsch, F. Bernhard, N. Morgner, Analyzing native membrane protein assembly in nanodiscs by combined non-covalent mass spectrometry and synthetic biology, Elife. 6 (2017) 1–19. https://doi.org/10.7554/eLife.20954.

[169] D.P. Schachtman, R. Kumar, J.I. Schroeder, E.L. Marsh, Molecular and functional characterization of a novel low-affinity cation transporter (LCT1) in higher plants, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 11079–11084. https://doi.org/10.1073/pnas.94.20.11079.

[170] N. Von Wirén, F.R. Lauter, O. Ninnemann, B. Gillissen, P. Walch-Liu, C. Engels, W. Jost, W.B. Frommer, Differential regulation of three functional ammonium transporter genes by nitrogen in root hairs and by light in leaves of tomato, Plant J. 21 (2000) 167–175. https://doi.org/10.1046/j.1365-313X.2000.00665.x.

[171] W.N. Fischer, D.D.F. Loo, W. Koch, U. Ludewig, K.J. Boorer, M. Tegeder, D. Rentsch,

E.M. Wright, W.B. Frommer, Low and high affinity amino acid H+-cotransporters for cellular import of neutral and charged amino acids, Plant J. 29 (2002) 717–731. https://doi.org/10.1046/j.1365-313X.2002.01248.x.

[172] E. Argyrou, V. Sophianopoulou, N. Schultes, G. Diallinas, Functional Characterization of a Maize Purine Transporter by Expression in Aspergillus nidulans, Plant Cell. 13 (2007) 953. https://doi.org/10.2307/3871352.

[173] T. Yasujima, C. Murata, Y. Mimura, T. Murata, M. Ohkubo, K. Ohta, K. Inoue, H. Yuasa, Urate transport function of rat sodium-dependent nucleobase transporter 1, Physiol. Rep. 6 (2018) 1–12. https://doi.org/10.14814/phy2.13714.

[174] B. Ching, S.F. Chew, Y.K. Ip, Ascorbate synthesis in fishes: A review, IUBMB Life. 67 (2015) 69–76. https://doi.org/10.1002/iub.1360.

[175] M.E. Jørgensen, D. Xu, C. Crocoll, H.A. Ernst, D. Ramírez, M.S. Motawia, C.E. Olsen, O. Mirza, H.H. Nour-Eldin, B.A. Halkier, Origin and evolution of transporter substrate specificity within the NPF family, Elife. 6 (2017) 1–30. https://doi.org/10.7554/elife.19466.

[176] J.T. Bridgham, E.A. Ortlund, J.W. Thornton, An epistatic ratchet constrains the direction of glucocorticoid receptor evolution, Nature. 461 (2009) 515–519. https://doi.org/10.1038/nature08249.

[177] S.M. Carroll, J.T. Bridgham, J.W. Thornton, Evolution of hormone signaling in elasmobranchs by exploitation of promiscuous receptors, Mol. Biol. Evol. 25 (2008) 2643–2652. https://doi.org/10.1093/molbev/msn204.

[178] M. Oda, Y. Satta, O. Takenaka, N. Takahata, Loss of urate oxidase activity in hominoids and its evolutionary implications, Mol. Biol. Evol. 19 (2002) 640–653. https://doi.org/10.1093/oxfordjournals.molbev.a004123.

[179] B.N. AMES, R. CATHCART, E. SCHWIERS, P. HOCHSTEINt, Uric acid provides an antioxidant defense in humans against oxidant- and radical-caused aging and cancer: A hypothesis, Proc. Natl. Acad. Sci. USA. 78 (1981) 6858–6862. https://doi.org/10.1051/lhb/1972051.

[180] B. Álvarez-Lario, J. Macarrón-Vicente, Uric acid and evolution, Rheumatology. 49 (2010) 2010–2015. https://doi.org/10.1093/rheumatology/keq204.

[181] P.K. Tan, J.E. Farrar, E.A. Gaucher, J.N. Miner, Coevolution of URAT1 and Uricase

during Primate Evolution: Implications for Serum Urate Homeostasis and Gout, Mol. Biol. Evol. 33 (2016) 2193–2200. https://doi.org/10.1093/molbev/msw116.

Appendix

CV and reprints of original publications

Αθανασία Βασιλική (Ανεζία) Κούρκουλου, BSc

+30 6955456673 akourkoulou@hotmail.com

Άλδου Μανουτίου 9, Αμπελόκηποι, Αθήνα 11521

www.linkedin.com/in/anezia-kourkoulou-32a311196

Ημερομηνία γέννησης: 09/12/1995 Εθνικότητα: Ελληνική

Εκπαίδευση

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Βιολογίας

Υποψήφια διδάκτορας (PhD candidate) στην Μοριακή Μικροβιολογία 12/2017-03/2021

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Βιολογίας

Πτυχίο Βιολογίας (4ετές | 240 ECTS) 01/2014-10/2017
Βαθμός: 7,76/10

Διακρίσεις και βραβεία

Υποτροφία του Ιδρύματος Σταύρος Νιάρχος (ΙΣΝ) (3ετής υποτροφία PhD)

Πιστοποιήσεις

ICH Good Clinical Practice E6 (R2) e-learning course with a score of 100% (08/2020)

Ερευνητική εμπειρία

Διδακτορική διατριβή | ΕΚΠΑ, Τμήμα Βιολογίας 12/2017-03/2021
Τίτλος : «Σχέσεις δομής-λειτουργίας διαμεμβρανικών μεταφορέων»
Επιβλέπων: Καθ. Γ. Διαλλινάς

Πτυχιακή εργασία | ΕΚΠΑ, Τμήμα Βιολογίας 10/2016-10/2017
Τίτλος: "Μελέτη SVCT-like μεταλλαγών του μεταφορέα ουρικού οξέος-ξανθίνης και έκφραση των ανθρώπινων μεταφορέων ασκορβικού (SVCT1 & SVCT2) στον μύκητα *Aspergillus nidulans*."

Παρουσιάσεις σε συνέδρια:

•39th Conference of the Greek Society of Biological Sciences 2017 (EEBE): "In search of the molecular details underlying vitamin C transporter evolution", A. Kourkoulou, A. Pittis & G. Diallinas **(προφορική παρουσίαση και poster).**
•36th Small Meeting on Yeast Transport and Energetics 2018 (SMYTE): "Determinants of substrate specificity in the NAT family: tracing transporter evolution", A. Kourkoulou, I. Theologidis & G. Diallinas **(poster).**

•Gordon Research Seminar, Mechanisms of membrane transport 2019 (Dynamics and Allosteric Regulation of Membrane Transport): "Membrane lipids are critical for dimer formation and stability, trafficking, and function of a eukaryotic purine transporter", A. Kourkoulou, B. Byrne & G. Diallinas **(poster).**
•Gordon Research Conference, Mechanisms of membrane transport 2019 (Folding and Evolution in the Membrane Environment: New Ways of Understanding Transport Mechanisms): "Membrane lipids are critical for dimer formation and stability, trafficking, and function of a eukaryotic purine transporter", A. Kourkoulou, B. Byrne & G. Diallinas **(προφορική παρουσίαση and poster).**
•15th European Conference on Fungal Genetics 2020 (ECFG15): "UapA-membrane lipid interactions are crucial for ER-exit, dimerization, function and expression of mammalian transporters in *A. nidulans*" A. Kourkoulou, P. Grevias, G. Lamprinidis, E. Mikros, B. Byrne and G. Diallinas **(προφορική παρουσίαση and poster).**

Δημοσιεύσεις

1. **Kourkoulou A**., Scazzocchio C., Frillingos S., Mikros E., Byrne B., Diallinas G. (2018). Nucleobase-Ascorbate-Transporter (NAT) Family. 1-6. 10.1007/978-3-642-35943-9_10090-1.
2. **Kourkoulou A**., Pittis A.A., Diallinas G. (2018). Evolution of substrate specificity in the Nucleobase-Ascorbate Transporter (NAT) protein family. Microbial Cell, 5(6), 280-292.
3. Dimou S., **Kourkoulou A**., Amillis S., Percudani R., Diallinas G. (2019). The peroxisomal SspA protein is redundant for purine utilization but essential for peroxisome localization in septal pores in *Aspergillus nidulans*, Fungal Genetics and Biology, 132, 103259.
4. **Kourkoulou A**., Grevias P., Lamprinidis G., Pyle E., Dionysopoulou M., Politis A., Mikros E., Byrne B., Diallinas G. (2019) Specific Residues in a Purine Transporter Are Critical for Dimerization, ER-Exit and Function. Genetics, 213(4), 1357-1372.
5. **Kourkoulou A**., Zantza I., Foti K., Mikros E., Diallinas G. (2021) Context-dependent cryptic roles of specific residues in substrate selectivity of the UapA purine transporter. J Mol Biol. Epub ahead of print. PMID: 33497644.

Δεξιότητες

Computer/Data analysis skills
Microsoft Office Tools [Excel, PowerPoint, Word]
Adobe Photoshop
GraphPad Prism
Icy

Γλώσσες
Ελληνικά (μητρική)
Αγγλικά (Certificate of Proficiency in English, University of Michigan| Pearson Test of English General Level 5)
Ιταλικά (Κρατικό Πιστοποιητικό Γλωσσομάθειας, Επίπεδο Β1)

Άλλα

Δίπλωμα οδήγησης κατηγορία Β (αυτοκίνητο)

Εθελοντική εμπειρία| Μέλος της οργάνωσης των ακόλουθων ημερίδων:
•"Career Opportunities of Biologists in Greece", 2015
•"Biology of Today: Research and Innovation", 2016
•"Think Graduate", 2017

microbial cell

Research Report

www.microbialcell.com

Evolution of substrate specificity in the Nucleobase-Ascorbate Transporter (NAT) protein family

Anezia Kourkoulou[1,#], Alexandros A. Pittis[2,#] and George Diallinas[1,*]

[1] Department of Biology, National and Kapodistrian University of Athens, Panepistimioupolis, Athens 15784, Greece.
[2] Department of Botany, University of British Columbia, Vancouver, BC, Canada.
[#] Equal contribution.
* Corresponding Author:
George Diallinas, Department of Biology, National and Kapodistrian University of Athens, Panepistimioupolis, Athens 15784, Greece;
Tel. +302107274649; E-mail: diallina@biol.uoa.gr

ABSTRACT **L-ascorbic acid (vitamin C) is an essential metabolite in animals and plants due to its role as an enzyme co-factor and antioxidant activity. In most eukaryotic organisms, L-ascorbate is biosynthesized enzymatically, but in several major groups, including the primate suborder Haplorhini, this ability is lost due to gene truncations in the gene coding for L-gulonolactone oxidase. Specific ascorbate transporters (SVCTs) have been characterized only in mammals and shown to be essential for life. These belong to an extensively studied transporter family, called Nucleobase-Ascorbate Transporters (NAT). The prototypic member of this family, and one of the most extensively studied eukaryotic transporters, is UapA, a uric acid-xanthine/H^+ symporter in the fungus *Aspergillus nidulans*. Here, we investigate molecular aspects of NAT substrate specificity and address the evolution of ascorbate transporters apparently from ancestral nucleobase transporters. We present a phylogenetic analysis, identifying a distinct NAT clade that includes all known L-ascorbate transporters. This clade includes homologues only from vertebrates, and has no members in non-vertebrate or microbial eukaryotes, plants or prokaryotes. Additionally, we identify within the substrate-binding site of NATs a differentially conserved motif, which we propose is critical for nucleobase versus ascorbate recognition. This conclusion is supported by the amino acid composition of this motif in distinct phylogenetic clades and mutational analysis in the UapA transporter. Together with evidence obtained herein that UapA can recognize with extremely low affinity L-ascorbate, our results support that ascorbate-specific NATs evolved by optimization of a sub-function of ancestral nucleobase transporters.**

doi: *10.15698/mic2018.06.636*
Received *originally: 28.01.2018;*
In revised form: 14.03.2018,
Accepted *14.03.2018,*
Published *22.03.2018.*

Keywords*: transporters, L-ascorbate, nucleobase, Aspergillus nidulans, primates.*

Abbreviations:
APR – ancestral protein resurrection,
NAT – Nucleobase-Ascorbate Transporter.

INTRODUCTION

Gene duplication and divergence is a major mechanism of evolutionary novelty, providing protein families with an expanded repertoire of activities and novel specificities. The classic view from Susumu Ohno [1-4] is that duplication creates redundancy, and redundancy provides the material for novelty. Different evolutionary scenarios, such as dosage effects, neo-functionalization or sub-functionalization, might rationalize preservation of gene duplicates [5-10]. The long-term outcome of neo-functionalization is that one copy retains the original function of the gene, while the second copy, free from selective pressure, explores novel functions. The major criticism of this model is the high likelihood of loss of functionality of the gene product before acquiring a new function, as acquiring functional novelty often needs several intermediate evolutionary steps some of which might introduce deleterious mutations. Novelty might evolve via promiscuous side-activities or sub-functions of proteins, which are optimized in the course of evolution and under specific selective pressures. This model puts emphasis on the multi-functional pre-duplication state of the evolving genes and introduces the possibility that positive selection pressure might drive evolution after gene duplication event. A major issue in the evolution of proteins that remains little an-

swered is how prominent functional changes, which would require the accumulation of multiple serial mutations, take place over the course of time. If many steps in a defined order were necessary for a dramatic functional or specificity change, what would have been the function of 'intermediate' proteins? Additional questions remain as to how protein evolution proceeds after gene duplications. Does protein evolution proceed in small steps or in large steps? Is evolution reversible? How does complexity evolve? Is functional specificity determined solely by the substrate or effector binding-site? [7, 10, 11]. One way to approach these questions is to reproduce protein evolution in the laboratory [12]. This can be achieved by systematic rational or random mutagenesis and DNA shuffling of specific genes, given a functional assay is designed to follow functional modification, or more efficiently, by direct genetic selection or genetic screens using model microorganisms. What is emerging from these approaches of directed evolution is that functional evolution necessitates conservation of sufficient protein stability, as conformational epistasis might suppress several routes of neofunctionalization [13].

A "horizontal" or phylogenetic approach of molecular and functional comparison of homologous proteins is also a powerful approach to predict protein functions and how these might have evolved. However, this *in silico* approach, by itself, can be misleading as in several protein families, function and mostly specificity cannot be predicted due to overlapping divergent and convergent evolutionary events. Over the last decade, "ancestral protein resurrection" (APR) has developed as a novel strategy to reveal the mechanisms and dynamics of protein evolution [14]. APR gives access to a protein's lost properties that may have transiently arisen over evolutionary time, which in turn can be used to infer the potential selection pressures that resulted in present-day sequences. Thus, APR can identify the causative mutation that resulted in protein neofunctionalization after gene duplication. APR has confirmed that conformational epistasis limits the pathways of functional evolution as neutral mutations may act as ratchets to modification, and additionally supported that evolution seems to proceed from promiscuity to increased specificity, usually via loss-of-function mutations [15].

A still unexplored aspect in protein evolution concerns whether transmembrane proteins, especially those of eukaryotic cells, evolve at similar rates and under similar evolutionary constraints as soluble proteins. This question arises from a number of peculiarities of transmembrane proteins, as for example, their co-translational folding and translocation in the ER membrane, the very complex and highly regulated subcellular traffic, the highly controlled ubiquitination-dependent turnover routes, and their dynamic interaction with specific lipids [16-18]. These special aspects of transmembrane proteins should in principle impose additional constraints or alternative selective pressures for evolution.

Here, we ask how mammalian L-ascorbate (vitamin C) transporters might have evolved from, apparently, ancestral nucleobase transporters. This question is based on the fact that members of the so-called Nucleobase Ascorbate Transporter (NATs) family [19-23] are highly specific for either nucleobases or L-ascorbate (see also later). This shift in substrate specificity is a dramatic one as these solutes have very different chemical formulas and properties and probably necessitate prominent changes in the architecture of the substrate-binding site and substrate translocation trajectory in the relevant NATs. Our results support a scenario that ascorbate-specific NATs evolved by optimization of a sub-function of ancestral nucleobase transporters.

RESULTS AND DISCUSSION

NATs are functionally classified into two distinct specificity groups: nucleobase versus L-ascorbate transporters

More than 24 members of the NAT family have been functionally characterized via direct assessment of their transport activities or indirectly via genetic and physiological studies [19-23] (Table 1). The most extensively studied NATs come from *Escherichia coli* or the filamentous ascomycete *Aspergillus nidulans.* The rest are homologues from other bacteria (*Bacillus subtilis*, *Sinorhizobium meliloti*), ascomycetes (*Aspergillus fumigatus*, *Aspergillus brasiliensis, Candida albicans*) or plants (maize, *Arabidopsis thaliana*). All bacterial NATs are H^+ symporters highly specific for either uracil or xanthine or uric acid. The fungal and plant members are H^+ symporters specific either xanthine-uric acid, or for adenine-guanine-hypoxanthine-uracil (only in plants) [19-25]. All these NATs exhibit relatively high affinity for their physiological substrates (at the low μM range), but can also recognize with lower affinity several nucleobase analogues. Notably, the relatively specific xanthine-uric acid NATs of fungi can be genetically converted, by single or limited missense mutations, to become broad-specificity promiscuous nucleobase transporters [26-29]. However, no microbial or plant, wild-type or mutant NAT, is capable of transporting L-ascorbate.

In sharp contrast to the microbial and plant proteins, most functionally characterized mammalian NATs, coming from mouse, rat, wild boar or humans, called SVCT1 or SVCT2 (Sodium-dependent Vitamin C Transporters; SLC23 family members), are highly specific for L-ascorbate/Na^+ [23]. However, there is a case reported of a mammalian (rat) NAT transporter, called rSNBT1, which is specific for nucleobases, rather than L-ascorbate [31]. Interestingly, the genes encoding orthologues of rSNBT1 in most primates (called *SVCT4*) are truncated pseudogenes. A third distinct homologue of NATs (SVCT3) identified in mammals seems unrelated to either nucleobase or L-ascorbate transport [23]. Thus, the emerging view is that all NATs in microorganisms and plants are nucleobase-specific, whereas mammals have both nucleobase and L-ascorbate NATs, but primates have lost via gene truncation the nucleobase-specific ones. However, this picture is little supported by the limited number and biased selection of functionally characterized NATs.

Two members of the NAT family have been crystallized. These are the UraA uracil transporter [32] of *E. coli* and the UapA uric acid-xanthine transporter of *A. nidulans* [33]. Both proteins form tight dimers, the formation of which is essential for transport activity [33-35]. The UraA and UapA monomers are made of mostly α-helical segments (TMS) characterized by a 7+7-helix inverted repeat [30]. Interestingly, the same topology is also found in transporters with

Table 1. Characteristics of biochemically characterized NATs.

Transporter name	Origin	Physiological substrate	Other substrates	Cation symport	Known structure	Q-H motif (TMS1) Substrate transport	ExxGD motif (TMS8) Substrate bind & transport	NAT *signature* motif (TMS10) Substrate bind & transport & *specificity*
UraA	*E. coli*	**uracil**	5-FU	H^+	yes/dimer bound to uracil	yes	yes (ExxG**H**)	T-**Y**-G-**E**-NxGxxxxTG
XanQ (YgfO)	*E. coli*	**xanthine**	xanthine analogues[#]	H^+	*	yes (**T**-H)	yes	T-**F**-A-**Q**-NxGxxxxTG
XanP (YicE)	*E. coli*	**Xanthine** *low-capacity*	?	H^+	*	yes	yes	C-**F**-G-**Q**-NxGxxxxTG
UacT (YgfU)	*E. coli*	**uric acid** *low-affinity*	xanthine (very low capacity)	H^+	*	yes	yes/no ExxG**M**	S-**F**-S-**Q**-NxGxxxxTG
UapA	*A.nidulans*	**uric acid-xanthine**	uracil, xanthine-uric acid analogues oxypurinol, allopurinol	H^+	yes/ dimer bound to xanthine	yes	yes	T-**F**-A-**Q**-NxGxxxxTR
UapC	*A.nidulans* *A.fumigatus*	**xanthine – uric acid**	uracil, xanthine-uric acid analogues oxypurinol	H^+	*	yes	yes	V-**F**-A-**Q**-NxGxxxxTR/K
Xut1	*C. albicans*	**xanthine – uric acid**	uracil, xanthine-uric acid analogues oxypurinol	H^+	*	yes	yes	V-**F**-A-**Q**-NxGxxxxTK
Lpe1	*Zea mays*	**uric acid-xanthine**		H^+	*	yes	yes	A-**S**-V-**E**-NxGxxxxTG
Nat3	*A. thaliana*	**adenine, hypoxanthine guanine, uracil**	?	H^+	*	yes	yes	A-**S**-V-**E**-NxGxxxxTG
Nat12	*A. thaliana*	**adenine, hypoxanthine guanine, uracil**	?	H^+	*	yes	yes	T-**L**-T-**E**-NxGxxxxTG
rSNBT1 **(SVCT4)**	Rat, mouse *predominantly in small intestine* *pseudogene in primates*	**hypoxanthine, xanthine guanine, uracil, thymine**	5-FU, oxypurinol	Na^+	*	yes	yes	S-**Y**-S-**E**-NxGxxxxTR
SVCT1	Rat, mouse, human *Epithelia kidney, intestines, liver, lung, skin*	**L-ascorbate**	-	2 Na^+ (need of Ca2+ or Mg2+)	*	yes	yes	S-**S**-S-**P**-NxGxxxxTK
SVCT2	Rat, mouse, human *Widespread, including Brain*	**L-ascorbate**	-	2 Na^+ (need of Ca2+ or Mg2+)	*	yes	yes	S-**S**-S-**P**-NxGxxxxTK
SVCT3	Vertebrates *predominantly in kidney*	?	?	Na^+	*	yes	**no** (SxxGC)	S-**S**-F-**P**-NxGxxxxTG

All NATs shown are high-affinity transporters recognizing their physiological substrate at the low μM range, except UacT which is low affinity transporter. NATs marked with an (*) can be modeled, with 100% probability, on the crystal structure of UapA or UraA conforming to the 7+7 inverted repeat fold [30]. 5-FU is 5-fluorouracil. Consensus sequences of three motifs involved in function and specificity are shown. Notice that the most significant differences in the NAT *signature motif* of nucleobase *versus* L-ascorbate transporters concern substitutions of the aromatic Phe/Tyr at position 2 and of the polar Gln/Glu with a Pro. ?, not tested. [#]S. Frillingos pers. com.

very little primary amino acid sequence similarity with NATs, such as the AzgA-like purine transporters [36] the plant boron transporter Bor1 [37], the human Band3 anion exchanger [38], or members of SulP transporter family [39]. All these are homodimeric transporters which seem to function via the so-called "elevator mechanism" of transport [40].

L-ascorbate transporters belong to a phylogenetically distinct NAT clade

We performed an extensive phylogenetic analysis of the NAT family using different datasets and methodologies. NAT sequences used for the analysis are, with very few exceptions, made of 414-650 amino acid residues in length. Selected NAT representatives from all major taxa were also tested, using the HHpred modeling program (https://toolkit.tuebingen.mpg.de), and shown to conform to the 7+7-helix inverted repeat structure of UapA (analysis not shown). Initially we reconstructed a phylogenetic tree using identified NAT homologues in all domains of life, from archaea and bacteria, to fungi, algae, plants and animals (see later). Prominent evolutionary gene-losses of NATs are identified several pathogenic protozoa (e.g. *Trypanosoma, Plasmodium, Leishmania, Giardia*, etc.) and in mesozoa, in several basal fungi including Microsporidia, Cryptomycota and Blastocladiomycota, as well as, in several *Saccharomycetales*, after the whole genome duplication. Interestingly, NAT loss in *Saccharomycetales* is concomitant with gene loss of uric acid oxidase, formation of an allantoin utilization cluster [41] and the "generation" of xanthine dioxygenase (Claudio Scazzocchio, unpublished observations). Other NAT losses in the animal clade are detected in basal Metazoa, such as the Porifera (sponges), some Ctenophora (Eumetazoa), some jaw worms (Gnathostomulida and Xenacoelomorpha), and among early-diverging fish (jawless fish, lampreys or hagfish).

Figure 1 and Supplementary Figures S1 and S2 highlight our major findings in respect to the distribution of NAT sequences. The first clear observation is that there is a primary evolutionary split between microbial (homologues in archeae, bacteria, algae and fungi, collectively called UapA-like) and plant-animal NATs (Figure 1a). In the microbial/UapA-like group, there are distinct sub-clades for several bacterial groups, archaea, algae and fungi (Supplementary Figure S1). A handful of protists, such as *Dictyostelium discoideum*, are also clustered with specific bacterial clades, suggesting that these might be the products of horizontal gene transfers. A second basic observation is that in the plant-animal clade (SVCT clade), the plants form a clear monophyletic branch including dicots and monocots, as well as, bryophytes (e.g. *Physcomitrella patens*) and some green algae (e.g. *Klebsormidium nitens*) (Figure 1a and Supplementary Figure S2). The animal clade includes several distinct branches. These however can be grouped in four major branches, which herein we named SVCT1, SVCT2, SVCT3 and SVCT4, based on the terminology of the mammalian homologues (Figure 1 and Supplementary Figure S2). The amino acid sequence identity of members of the microbial, plant or animal clades is ~30-92%, whereas the identity between microbial-plant, microbial-animal or plant-animal NATs is ~24-33%.

To gain deeper insights into the evolution of the plant and animal – specific - SVCT clade, we focused on sequences from this clade and deepened our taxon sampling. We used HMMER [43] and the NAT Pfam domain (PF00860) to extract 728 NAT sequences from 171 plant and animal fully sequenced genomes (see methods). After reconstructing a phylogenetic tree using these sequences, we detected and extracted only the 674 sequences comprising the SVCT clade (Figure 1a and Supplementary Figure S2). In the animal clade, an "extended" SVCT4 group (> 45% identity among its members) is widely distributed across Metazoa, and includes members from most vertebrates, non-vertebrates and early diverging animals, as for example, *Trichoplax adhaerens* (the only extant representative of phylum Placozoa, which is a basal group of multicellular metazoa), *Nematostella vectensis* (sea anemone belonging to Cnidaria), *Ciona intestinalis* (sea squirt), *Strongylocentrotus purpuratus* (sea urchin), *Orbicella faveolata* (coral), *Macrostomum lignano* (flatworm), etc. (Supplementary Figure S2). Most SVCT4-positive species include a single SVCT4 homologue, but there is also significant genus-specific expansion of the group, apparently by gene duplications (e.g. 6-14 paralogues in *Caenorhabditis*, 8 in *Ciona intestinalis*, 5-6 in some fish; Figure 1b). Interestingly, all Haplorhini (anthropoid primates or simians and the tarsier *Carlito syrichta*) have lost functional SVCT4 homologues, due to gene truncations (i.e. generation of pseudogenes; [23]). "Patchy" SVCT4 loss is detected also in some non-primate species, but these correspond to genus-specific cases, rather than representing losses in major groups (Figure 1b). Interestingly, SVCT4 loss often coincides with the loss in the ability to synthesize L-ascorbate, although not all known cases that lack this ability lack SVCT4 proteins. Despite the fact that only a single member of the SVCT4 group is functionally characterized (i.e. the rat rSNBT1 nucleobase transporter [31]), the conservation of specific functional motifs present in bacterial, plant or fungal nucleobase transporters, allows us to hypothesize that members of the SVCT4 group are specific for nucleobase-related substrates (see later).

The SVCT1 and SVCT2 paralogous subfamilies form a separate clade (~50% identity with SVCT4 in the same species) that includes the well-characterized mammalian L-ascorbate transporters. The SVCT1/SVCT2 group appears exclusively in vertebrates and is probably the result of duplication and diversification of an ancestral animal transporter. Most species maintain one member of SVCT1 and SVCT2 (Figure 1b). Thus, a gene duplication event in the vertebrate ancestor seems to have generated the two subfamilies. As the SVCT1 and SVCT2 proteins show high similarity (>68% identity) and all functionally characterized homologues are L-ascorbate transporters, one can speculate that the specificity of L-ascorbate has evolved in the vertebrate ancestor and was retained in fish and tetrapods (see Fig.1). However, whether all extant members of the SVCT1/SVCT2 group are specific L-ascorbate transporters,

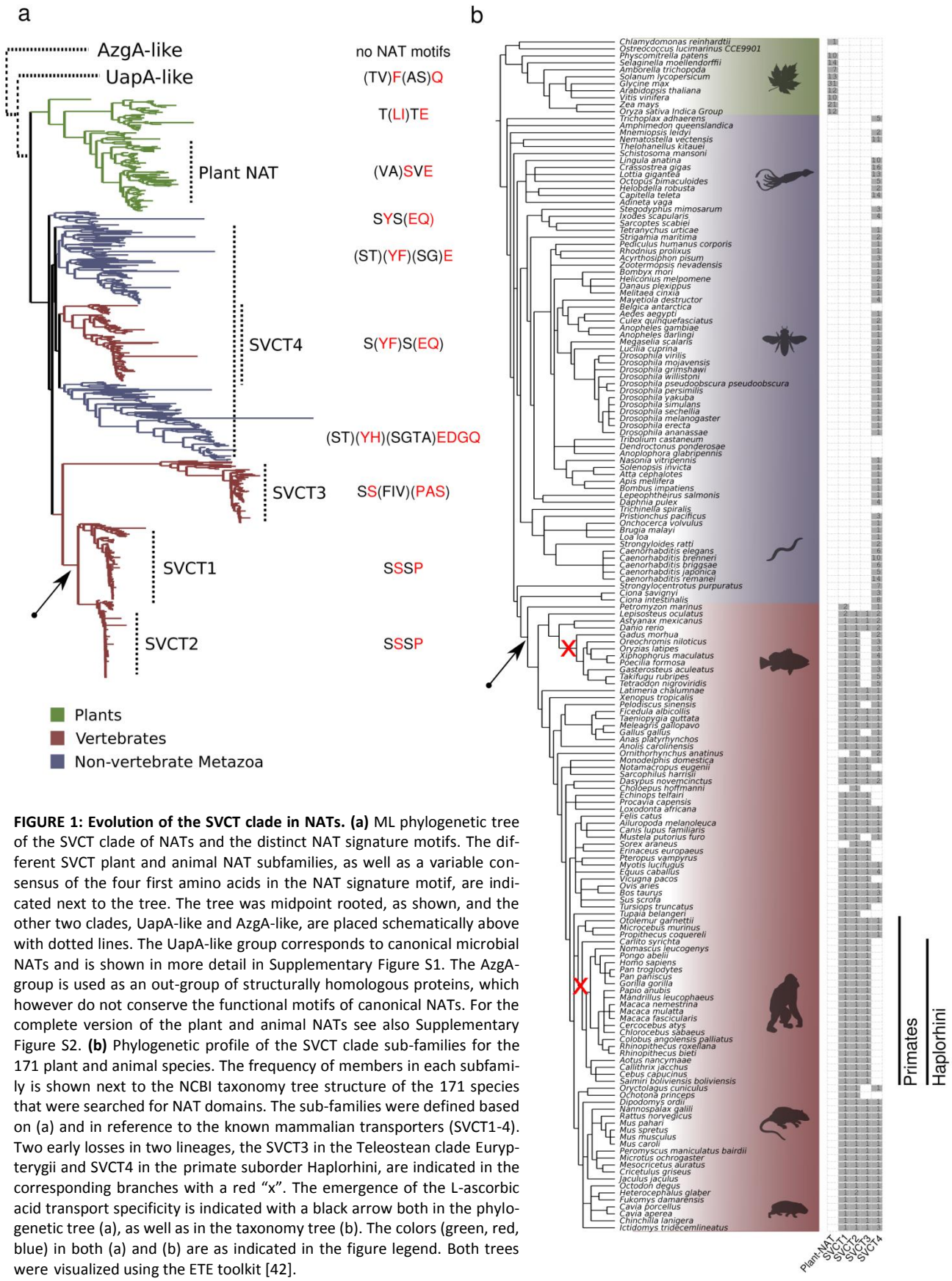


FIGURE 1: Evolution of the SVCT clade in NATs. (a) ML phylogenetic tree of the SVCT clade of NATs and the distinct NAT signature motifs. The different SVCT plant and animal NAT subfamilies, as well as a variable consensus of the four first amino acids in the NAT signature motif, are indicated next to the tree. The tree was midpoint rooted, as shown, and the other two clades, UapA-like and AzgA-like, are placed schematically above with dotted lines. The UapA-like group corresponds to canonical microbial NATs and is shown in more detail in Supplementary Figure S1. The AzgA-group is used as an out-group of structurally homologous proteins, which however do not conserve the functional motifs of canonical NATs. For the complete version of the plant and animal NATs see also Supplementary Figure S2. **(b)** Phylogenetic profile of the SVCT clade sub-families for the 171 plant and animal species. The frequency of members in each subfamily is shown next to the NCBI taxonomy tree structure of the 171 species that were searched for NAT domains. The sub-families were defined based on (a) and in reference to the known mammalian transporters (SVCT1-4). Two early losses in two lineages, the SVCT3 in the Teleostean clade Eurypterygii and SVCT4 in the primate suborder Haplorhini, are indicated in the corresponding branches with a red "x". The emergence of the L-ascorbic acid transport specificity is indicated with a black arrow both in the phylogenetic tree (a), as well as in the taxonomy tree (b). The colors (green, red, blue) in both (a) and (b) are as indicated in the figure legend. Both trees were visualized using the ETE toolkit [42].

or whether some SVCT1/SVCT2 homologues transport nucleobases or other substrates, is not formally known.

Finally, SVCT3-like sequences are found in nearly all tetrapod animals and several, but not all, fish (e.g. the *Lepisosteus oculatus, Astyanax mexicanus, Danio rerio, Rhincodon typus;* (Figure 1b). Noticeably, SVCT3 has been lost in some lineages including the Acanthomorpha, a very diverse clade of teleost fishes. The SVCT3 proteins are also the most divergent NAT group in the animal clade, as not only show the lowest overall amino acid sequence similarity with the other clades (< 34% identity with SVCT1/2 and SVCT4), but additionally do not fully conserve the specific amino acid motifs necessary for function and specificity in "canonical" NATs (see later). This might also explain the fact that its phylogenetic position in the tree was rather unstable across different phylogenetic algorithms and parameters (not shown). Thus, SVCT3 members are very probably transporters of substrates other than nucleobases or ascorbate. This is indeed supported by reports showing that the human hSVCT3 does not transport nucleobases or L-ascorbate [23]. An analogous "non-canonical" (i.e. lacking the motifs necessary for nucleobase or L-ascorbate transport) divergent NAT group, called UapD-like proteins, has recently been described in fungi [44], while other divergent NAT branches, all of unknown function, seem to exist also in prokaryotes (Stathis Frillingos pers. com. and observations not shown herein).

The NAT signature motif is differentially conserved in nucleobase and L-ascorbate transporter clades

Previous genetic, molecular, biochemical and structural studies have established that a small number of amino acid residues are essential or critical for NAT function. Most of these fall within three highly conserved motifs in transmembrane segments TMS1, TMS8 and TMS10. The two-amino acid motif **Gln**-**His** in TMS1 is nearly absolutely conserved in all NATs (only a small number of mostly microbial homologues have a Gln-Gln, Thr-His or Gln-*Φ* versions, where *Φ* is Met or Ile) and seems to be involved in dynamic interactions of TMS1-TMS3-TMS10 necessary for transport catalysis [19-22, 43, 44]. The TMS8 motif, conforming to the sequence **Glu**-X-X-**Gly**-**Asp**/His/Gly (in bold the most prominently conserved residues), is involved in interactions with substrates and transport [19-22, 46, 47]. Finally, a longer motif in TMS10, Gln/Glu/Pro-**Asn**-X-**Gly**-X-X-X-X-**Thr**-Gly/Arg/Lys, historically referred as the NAT signature motif [47], is essential for both substrate recognition and transport activity, as well as, substrate specificity [19-22,46-48]. The presence of these three motifs is taken, in the present analysis, as an additional criterion for classifying any uncharacterized transporter as a member of the "canonical" NAT family.

The phylogenetic analysis performed herein confirms the presence and exceptional sequence conservation of these motifs in nearly all NAT clades. More specifically, the motifs in TMS1 and TMS8 are highly conserved in all NATs, except from the highly divergent SVCT3 group, which conserves the TMS1 motif, but lacks the TMS8 motif (observations not shown). The motif in TMS10 exists in all NATs but appears in two major versions, especially if we consider three additional upstream amino acid residues (see Figure 1a and Supplementary Figure S2). More specifically, in functionally characterized nucleobase-specific NATs and their close homologues it conforms to the sequence Thr/Ser^{1}-Phe/Tyr^{2}-Ala/Ser^{3}-Gln/Glu/Pro^{4}-**Asn**^{5}-X-**Gly**^{7}-X-X-X-X-**Thr**$^{\mathbf{12}}$-Gly/Arg/Lys^{13} (numbers shown are arbitrary). In the L-ascorbate transporter clades (i.e. SVCT1 and SVCT2) this motif is Ser^{1}-Ser^{2}-Ser^{3}-Pro^{4}-**Asn**^{5}-X-**Gly**^{7}-X-X-X-X-**Thr**$^{\mathbf{12}}$-Arg^{13}. In other words, the most prominent changes, shown underlined, in the two versions of this extended TMS10 motif concern replacements of the aromatic Phe/Tyr^{2} and polar Gln/Glu^{4} (nucleobase transporters) by Ser^{2} and Pro^{4} residues (L-ascorbate transporters). Interestingly, the divergent SVCT3 group has an apparent TMS10 motif more similar to the one found in L-ascorbate transporters rather to that of nucleobase transporters, which also supports its close evolutionary relatedness to them (see Figure 1a).

Based on previous functional [19-22] and recent structural [32, 33, 35] studies performed in fungal (e.g. UapA) and bacterial (e.g. XanQ) homologues, Gln/Glu^{4} is one of the two major NAT residues interacting via a strong bipolar interaction with the nucleobases (the other being the absolutely conserved Glu in the TMS8 motif), while Phe/Tyr^{2} also interacts directly with the nucleobase ring via *pi-pi* stacking. Most importantly, Phe/Tyr^{2} and Gln/Glu^{4} have proved as key elements in determining the specificity of NATs in respect to different nucleobases. This is best exemplified by specific mutations in UapA, where substitution of Gln^{4} by Glu^{4} (e.g. Q408E) converts this relative specific uric acid-xanthine transporter into a promiscuous transporter capable of recognizing hypoxanthine, guanine or uracil [27], whereas replacement of Phe^{2} by Tyr^{2} (e.g. F406Y) confers a detectable capacity for hypoxanthine transport [29]. Thus, it becomes apparent that more dramatic substitutions of Phe/Tyr^{2} or Gln/Glu^{4}, introducing residues as the ones found in L-ascorbate specific NATs (i.e. Ser^{2} and Pro^{4}), are expected to result in even more prominent specificity modifications, and might in fact contribute to the "generation" of L-ascorbate transporters. All other conserved residues in the TMS10 signature motif of NATs, including the L-ascorbate transporters, are known to contribute to the kinetic parameters of transport (Thr/Ser^{1}, Ala/Ser^{3}, Asn^{5}, Gly^{7}, Thr^{12} and Gly/Arg/Lys^{13}), rather than playing a direct role in determining substrate specificity.

The differentially conserved residues in the NAT signature motif are essential for function and specificity

No mutation, up to date, in any nucleobase-specific NAT, has led to a transporter capable of recognizing L-ascorbate. However, it should be stressed that there has been no rational genetic design to do so, as most functional mutations were designed or genetically selected to change the specificity of NATs in respect to different nucleobases rather than to L-ascorbate. The only exception is the substitution of Gln^{4} by Pro^{4} in the NAT signature motif of the UapA [47] or XanQ transporters [48]. In this case, introducing a proline residue led to dramatic reduction (>100-fold) in nucleobase binding and thus no detectable transport

activity, while no evidence was obtained for acquisition of L-ascorbate recognition.

To further test whether specific variations of the NAT signature motif might be relevant to L-ascorbate recognition, we constructed and/or functionally analyzed a series of relevant UapA mutants. These mutants correspond to the following amino acid replacements in UapA: Q408P [46], A407S/Q408P, T405S/F406S/A407S and F406S/A407S/Q408P. In other words, we modified the first part of the NAT signature motif from TFAQ to TFA**P**, TF**SP**, **SSS**Q or T**SSP** (changes in bold underlined). These mutant versions of UapA were expressed into an appropriate *A. nidulans* strain genetically lacking all endogenous major nucleobase transporters (i.e. UapA, UapC and AzgA) and functionally analyzed in respect to their capacity to transport or bind nucleobases or L-ascorbate, and for proper plasma membrane sorting and stability. Growth tests on purines as sole nitrogen sources can be used as a tool to reveal the functionality of the relative purine transporters. Thus, a control strain expressing a functional UapA (i.e. *uapA*$^{+}$), but no other major purine transporter, grows on uric acid or xanthine, but not on other purines. Figure 2a shows that all UapA mutants tested have lost (T405S/F406S/A407S and F406S/A407S/Q408P) or have dramatically reduced (e.g. Q408P and A407S/Q408P) capacity for growth on uric acid or xanthine. This was confirmed by direct measurements of transport of radiolabeled xanthine (Figure 2b). In addition, none of the mutants tested have acquired any capacity for transport of any other purine, nucleoside or nucleobase-related metabolite (Figure 2a and results not shown). As all UapA mutants tested seem to correspond to abolished or reduced transport function, we wanted to exclude that this not due to misfolding, problematic sorting to the plasma membrane or protein instability. For this, we took advantage of the fact that all mutations analyzed are made in a fully functional UapA protein tagged with GFP. *In vivo* epifluorescence microscopy showed that all mutant versions of UapA analyzed are properly and stably localized in the plasma membrane (right panel in Figure 2a), which in turn confirms that all relevant mutations affect transport function per se.

We additionally tested whether the mutations introduced lead to any growth phenotype that could be related to UapA-dependent L-ascorbate accumulation. Previous results have shown wild-type *A. nidulans* cannot import L-ascorbate, at least when this is externally supplied to concentration up to 1 mM, while higher concentrations of L-ascorbate (5.6-56.8 mM or 0.1-1%) have a progressively moderate inhibitory effect on wild-type *A. nidulans* growth ([24], and unpublished observations). Figure 2c shows that none of the UapA mutations analyzed, or the total genetic absence of UapA (*uapA*$^{-}$), alters the growth phenotype of *A. nidulans* on L-ascorbate. We also tried to detect cytoplasmic accumulation of externally supplied L-ascorbate in the mutants or a wild-type strain, using HPLC or a biochemical assay, or by measuring the accumulation of radiolabeled L-ascorbate. In all cases, we did not obtain any evidence of L-ascorbate accumulation (results not shown).

Our results suggested that rational modification of the NAT signature motif is not sufficient, by itself, to confer detectable L-ascorbate accumulation. However, the stable and proper cellular expression of the mutants enabled us to use a sensitive *in vivo* cellular assay, based on substrate-elicited endocytosis, as an indirect measure of the ability of different UapA versions to "recognize", and not necessarily transport at a detectable level, L-ascorbate. This assay is based on the fact that UapA is highly sensitive to endocytic turnover triggered by substrate or ligand binding [49, 50]. Substrate-elicited transporter turnover is generally promoted by activity-dependent alteration of the conformation of a transporter from an outward-facing to a cytoplasm-facing topology, and can be followed by the appearance of fluorescing vacuoles and early endosome due to sorting of GFP-tagged transporters into these organelles [49, 50]. Based on this, we tested whether L-ascorbate could promote the endocytosis of wild-type or mutant versions of UapA-GFP, as well as, of a control transporter unrelated to NATs, namely FurA-GFP. FurA is a highly specific allantoin transporter that does not recognize either nucleobase or L-ascorbate [51].

Figure 2d shows that the presences of 0.1% L-ascorbate, for 15 min, promotes significant endocytosis of UapA mutants, but has no similar effect on wild-type UapA or FurA. After prolonged presence of L-ascorbate (1 h), endocytosis of the UapA mutants becomes more prominent, mostly evident in T405S/F406S/A407S and F406S/A407S/Q408P, but also becomes apparent, albeit at a much lower level, in wild-type UapA. Importantly, no endocytic turnover was detected in FurA, in all conditions tested. Additionally, prolonged incubation with L-ascorbate also elicited total loss of fluorescence, associated with a dramatic change in hyphae morphology, in a significant fraction of the UapA mutants. The L-ascorbate-dependent morphological effect, which was insignificant in strains expressing wild-type UapA or FurA, was reflected as a dramatic increase of the number and size of vacuoles and a reduction in the width of hyphae, two changes often seen under severe cell stress. The ascorbate-dependent effect was mostly observed in the triple mutants F406S/A407S/Q408P (in 55% hyphae) and T405S/F406S/A407S (in 68% hyphae) (Figure 2d and 2e), while a similar, albeit reduced, effect is seen in A407S/Q408P (40%) and Q408P (27%) (see Figure 2d and 2e). Given that the most prominent L-ascorbate associated effects are visible in F406S/A407S/Q408P and T405S/F406S/A407S, it seems that the presence of a tandem of Ser residues within the NAT substrate binding site, and the removal of the planar aromatic Phe reside, favor L-ascorbate recognition by the UapA mutants. Importantly, our results show that the L-ascorbate effect on hyphae morphology and the loss of GFP fluorescent are both UapA-specific, as we did not detect the same negative effects in a strain expressing FurA-GFP, which does not express UapA or any other nucleobase transporter (see strain list in Materials and methods).

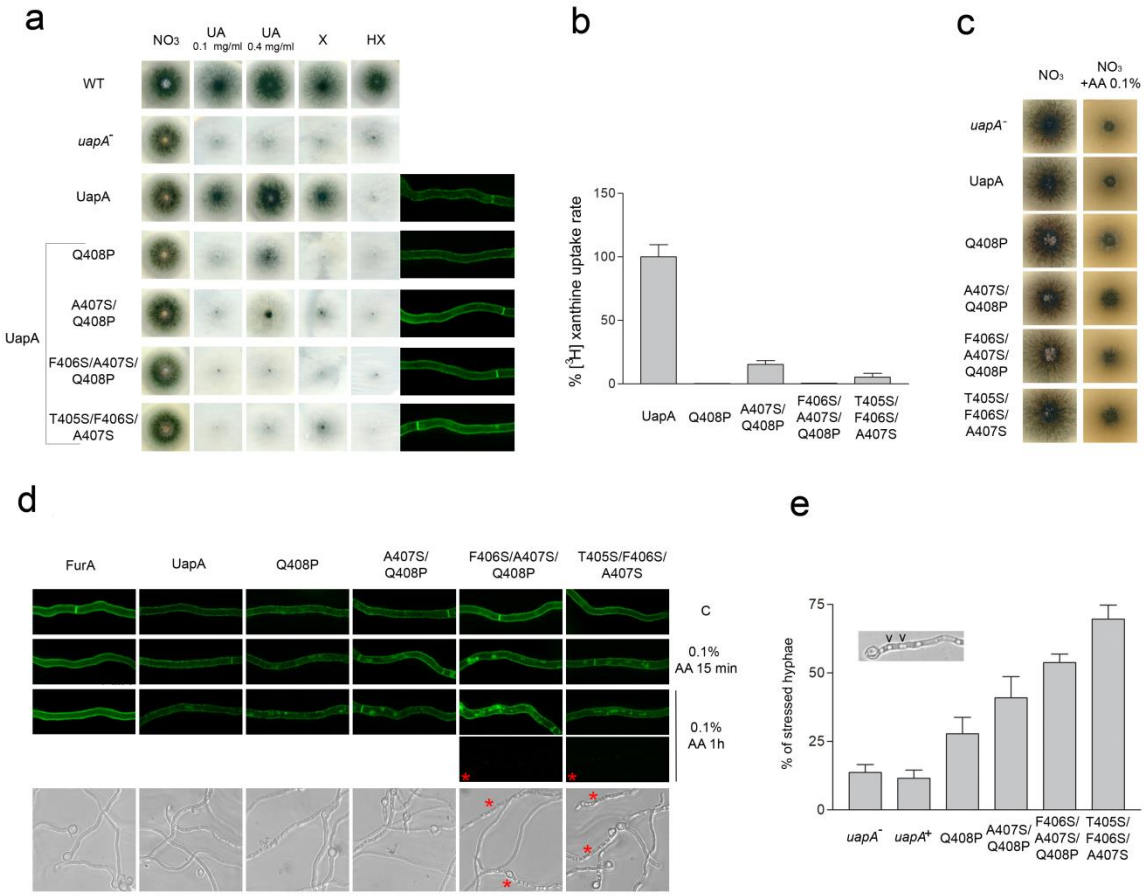


FIGURE 2: Specific residues within the NAT signature motif are critical for UapA function and specificity. (a) Growth tests, on standard MM, of isogenic strains expressing wild-type (UapA) or mutant UapA versions (Q408P, A407S/Q408P, T405S/F406S/A407S and F406S/A407S/Q408P), expressed in a genetic background genetically lacking other major purine transporters (i.e. *ΔuapC ΔazgA*). A strain lacking a functional UapA (*uapA*$^-$) and a standard wild-type *A. nidulans* strain (WT) are also included in the test as controls. The test depicts growth on 10mM sodium nitrate (NO_3) or 0.1 mg/ml uric acid (UA), xanthine (X), hypoxanthine (HX), after 48 h at 37°C. A 4-fold higher concentration of UA (0.4 mg/ml) was also used in order to distinguish UapA mutants with highly reduced affinity for substrates (Q408P, A407S/Q408P) from mutants with an apparent total loss of transport activity (T405S/F406S/A407S and F406S/A407S/Q408P). Subcellular localization of wild-type or mutant UapA-GFP versions, as revealed by *in vivo* epifluorescence microscopy, is also shown at the right side of the panel. Notice that in all cases UapA-GFP is stably localized in the plasma membrane and the septa of *A. nidulans* hyphae. For details of sample preparation see Material and methods. **(b)** Comparative UapA-mediated uptake rates of radiolabeled ^{3}H-xanthine in strains expressing wild-type (UapA) or mutant versions (Q408P, A407S/Q408P, T405S/F406S/A407S and F406S/A407S/Q408P) of UapA-GFP. Results are averages of three measurements for each concentration point. SD was < 20%. **(c)** Growth tests, at 37°C, on standard MM supplemented with 10 mM sodium nitrate (NO_3) as nitrogen source in the absence or presence of 0.1% L-ascorbate (AA) of isogenic strains lacking UapA (*uapA*$^-$), or expressing wild-type (UapA) or mutant UapA versions (Q408P, A407S/Q408P, T405S/F406S/A407S and F406S/A407S/Q408P). **(d)** *Upper panel*: *In vivo* epifluorescence microscopy following the effect of L-ascorbate (0.1% AA) on UapA-GFP or FurA-GFP subcellular localization. Notice that after 15 min addition of AA clear endocytic turnover of UapA-GFP (i.e. appearance of cytoplasmic structures corresponding to vacuoles; Gournas et al., 2010) is observed in all UapA mutants, but not in wild-type UapA or FurA. Vacuolar turnover is more evident after 1 h of AA in the triple mutants T405S/F406S/A407S and especially F406S/A407S/Q408P, where more and larger vacuoles appear. Notice that after 1 h of AA, moderate vacuolar turnover is also evident in the wild-type UapA, but not in FurA. Noticeably also, 1 h in AA triggers "loss" of GFP fluorescence in a significant fraction (> 50%) of hyphae in in the triple mutants T405S/F406S/A407S and F406S/A407S/Q408P (lowest black boxes marked with a red asterisk). The same effect, albeit significantly reduced, is also seen in A407S/Q408P (36%) and Q408P (22%), while in the wild-type UapA-GFP or FurA-GFP fluorescence loss is < 10% (not shown). *Lower panel*: L-ascorbate (1% AA for 1 h) triggers differential hypervacuolarization and reduction of hyphal width (i.e. cell stress) in UapA mutants and controls. Stressed hyphae are marked with a red asterisk. **(e)** Quantification of the effect of L-ascorbate on hyphal morphology, as recorded for 100 hyphae, of each strain analyzed. In all cases, hypervacuolarization and reduction of the width of hyphae was associated with loss of fluorescence (not shown). The results depicted in the graph confirm the apparent cytotoxicity of L-ascorbate in the triple mutants T405S/F406S/A407S and F406S/A407S/Q408P, where 50-70% of hyphae seem stressed, followed by progressively lower effects on A407S/Q408P (40%), Q408P (27%) and wt UapA (11%).

UapA can bind L-ascorbate with extremely low affinity: evidence for evolution of L-ascorbate transporters via optimization of a NAT sub-function

The observation that addition of L-ascorbate can also trigger a degree of vacuolarization and endocytic turnover of wild-type UapA-GFP, made us re-examine whether UapA can recognize, even with very low affinity, L-ascorbate. We performed standard competitive inhibition assays to test whether increasing concentrations of L-ascorbate (1-60 mM) inhibit the uptake of radiolabeled ^{3}H-xanthine (1 μM). Results shown in Figure 3 establish that wild-type UapA can indeed bind L-ascorbate with very low affinity, reflected in a K_i value of ~17 mM. We repeated the same experiment using a promiscuous UapA mutant capable of recognizing all nucleobases and obtained a moderately increased binding affinity also for L-ascorbate (Ki ~ 11 mM). This mutant (R481G/T526M) has an intact NAT signature motif, but includes two mutations that seem to loosen its specificity by modifying the functioning of putative outward and inward gates [29, 33]. Thus, UapA can in principle bind L-ascorbate with an affinity approx. 1000-fold lower than the one for its physiological substrates (xanthine or uric acid), or and 10- to 100-lower than its affinity for other nucleobase analogues [47]. This finding supports the hypothesis that other nucleobase-specific NATs might also be capable of very low-affinity recognition o L-ascorbate. This in turn supports a model where L-ascorbate NATs evolved via a progressive improvement of a sub-function of nucleobase-specific NATs.

CONCLUSIONS

Nucleobases are planar molecules with relatively low solubility, whereas L-ascorbate is non-planar and extremely soluble. Such a difference should necessitate prominent changes in the architecture of the substrate-binding site and of the substrate translocation trajectory in the relevant NATs. Our present analysis supports that NAT neofunctionalization took place via progressive improvement of a sub-function (i.e. very low affinity binding of L-ascorbate) in the course of evolution and probably under "new" selective pressures imposed by the evolution of L-ascorbate biosynthesis. This idea leads to many more questions. Are there any present day NATs transporting both substrates equally efficiently? Or are there NATs that can transport L-ascorbate more efficiently than nucleobases? If cases of double specificity exist, could these represent intermediate versions in the evolutionary path from nucleobase to L-ascorbate recognition? Or did only extinct ancestral nucleobase-transporters had this double specificity? NATs that contain transitory variations of NAT signature motif might reflect extant intermediate steps in the evolution of L-ascorbate transporters.

Still many more questions remain to be answered in the direction opened by the present work. Has the specificity shift from nucleobase to L-ascorbate occurred via transporter intermediates that had specificities other than nucleobase or L-ascorbate (could the SVCT3 clade being such a putative case for example?), or even little or no function at all? Can transporters evolve the ability to recognize a new substrate before or after this substrate first appears and how this is related to emergence of substrates via evolution of biosynthetic pathways? Very recently, it was shown that NPF transporters are able to transport both cyanogenic glucosides and glucosinolates, before plants evolved the ability to synthesize glucosinolates [52]. Later in evolution, these dual-specificity transporters specialized to transport only glucosinolates. In addition, Jørgensen *et al*. [52] also showed that early glucosinolate transporters could move a broad variety of glucosinolates, but later evolved to only transport particular types. These findings support the idea that evolution of transporter specificity and biosynthetic pathways might evolve together. An analogous case concerning the evolution of specificity in vertebrate corticoid receptors was reported earlier

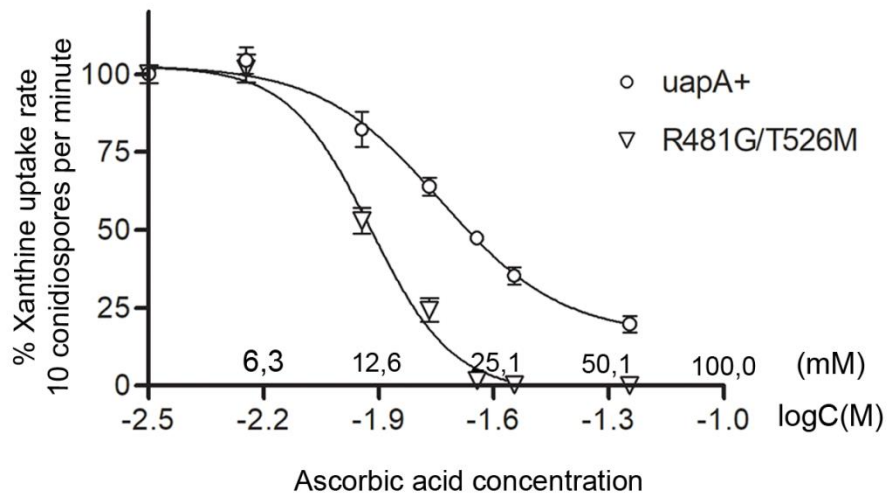


FIGURE 3: L-ascorbate inhibits UapA transport activity. The graph shows dose response curves of [^{3}H]-xanthine uptake rate in strains expressing either wild-type UapA or the isogenic mutant UapA-R481G/T526M, in the presence of increasing concentration of non-radiolabeled L-ascorbate (0-56 mM). In both cases, L-ascorbate inhibited UapA transport activity, with K_i values estimated at 17 and 11 mM for UapA and UapA-R481G/T526M, respectively. Results are averages of three measurements for each concentration point. SD was < 18%. For details of measurements and K_i estimation see Materials and methods.

[53, 54]. Corticoid and glucocorticoid receptors evolved post duplication of a dual specificity aldosterone and cortisol receptor basal to the jawed vertebrate. However, aldosterone biosynthesis did not arise before the advent of tetrapods suggesting that the ancestral receptor evolved affinity towards aldosterone before the hormone was present, possibly as a by-product of the receptors' affinity towards chemically similar ligands [54].

L-ascorbate biosynthesis seems to have evolved in plants and animals (prokaryotes do not synthesize it) via three major routes. The animal kingdom uses one of these pathways, and photosynthetic eukaryotes (plants, algae and some single-celled organisms) employ the other two. All three pathways proceed via different intermediate molecules, but the two from photosynthetic eukaryotes both use an enzyme called galactolactone dehydrogenase (GLDH), instead of gulonolactone oxidase (GULO), to catalyze the final step [55]. Some evidence exists that GULO and GLDH might have evolved from a common ancestor. Interestingly, certain animals, including guinea pigs, most bats, some fish and bird species, as well as humans and many other primates, have lost the ability to synthesize L-ascorbate. These animals obtain ascorbate from their diets [55].

Another important gene loss in vertebrates, possibly relative to NAT evolution, concerns purine catabolism. Most animals oxidize purines into allantoin, which is excreted from their body as the end product of purine catabolism. Humans and other primates are among some groups that lost the gene for uric acid oxidase (i.e. present as a pseudogene), which catalyzes of oxidation uric acid to allantoin [56]. This gene loss is partly responsible for the elevated levels of uric acid in blood plasma. Uric acid is a powerful antioxidant and is a scavenger of singlet oxygen and radicals, in a way similar, but not identical, with the action of L-ascorbic acid [57]. It has been argued that the loss of uric acid oxidase and elevation of uric acid levels in the body have played an important role in the development of the primate brain and human intelligence due to its protective role against oxidation related to neurotoxic activity. High uric acid levels have also been linked to human life span [57, 58]. Could it be that evolution of NAT specificity, which shifted from one outstanding antioxidant (uric acid) to another (L-ascorbic acid), and the loss of nucleobase-specific NATs in primates, be related to the co-evolution of relevant metabolic pathways? Relative to such questions, recent evidence supported the co-evolution of the uric acid transporter URAT1 (distinct from NAT transporters) and uric acid oxidase during primate evolution [59]. To most of the above questions, a future challenge will be to understand the detailed molecular changes that led to changes in specificity and possibly relate their outcome to the evolution of relative metabolic pathways.

MATERIALS AND METHODS

Media, growth conditions and strains

Standard media for *A. nidulans* were used. Media and supplemented auxotrophies were at the concentrations given in the Fungal Genetics Stock Center database (http://www.fgsc.net). Nitrogen sources were used at final concentrations of 10 mM sodium nitrate or 0.5 mM of purines. *A. nidulans* strains used are listed in Supplementary Table S1. The recipient for transformations was the *uapA*Δ *uapC*Δ *azgA*Δ *pabaA1 argB2* mutant strain lacking the endogenous *uapA* gene, as well as the genes encoding a secondary xanthine/uric acid transporter (UapC) and the major purine transporter AzgA. Plasmids carrying *gfp*-tagged wild-type (wt) or mutant versions of *uapA* (i.e. UapA-GFP, [60]) were introduced by standard genetic transformation into the recipient strain and transformants were selected on the basis of the *argB* arginine auxotrophy (*argB2*) complementation [61]. The strain expressing FurA-GFP, used as a control in Figure 2D, is described in [61]. GFP C-terminal tagging has previously shown to have absolutely no effect on UapA or FurA transport kinetics [51].

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). Restriction enzymes were from Takara Bio (Lab Supplies Scientific SA, Hellas). Conventional PCR reactions were done with KAPA-Taq DNA polymerase (Kapa Biosystems, Lab Supplies Sci- entific SA, Hellas). Cloning and amplification of products were done with Kapa HiFi (Kapa Biosystems, Lab Supplies Scientific SA, Hellas).

Construction and analysis of Aspergillus mutants

All *uapA* mutations appearing in this work were generated by standard site-directed mutagenesis, using specific oligonucleotides (Supplementary Table S2), the QuikChange Mutagenesis Kit and plasmid pAN510-GFP carrying a *gfp*-tagged *uapA* gene, as a template [47, 60]. The Q408P mutant was constructed previously [27, 47]. Mutant alleles of *uapA* were introduced by genetic transformation into the recipient strain *uapAD uapCD azgAD pabaA1 argB2*. Selected transformants were purified and those arising from single-copy plasmid integration, as evidenced by PCR and Southern blot analysis, identified [61]. Transformants were tested, at 25 or 37°C, pH 6.8, for their ability to grow on MM supplemented with uric acid or other purines as nitrogen sources, or for their general growth phenotype in the presence of L-ascorbate (concentration range from 0.1 to 1 %). Transformants were also tested, using epifluorescence inverted microscopy (using a Zeiss Observer Z1/AxioCam HR R3 camera/Zen lite 2012 software) in respect to UapA-GFP subcellular localization, both in the absence or presence of L-ascorbate (0.01-0.2%) for 15, 60 or 120 min. Samples for epifluorescence microscopy were prepared as previously described, after 16-18 h of growth in standard MM supplements with sodium nitrate as nitrogen source [34].

Transport assays

UapA mutants were functionally characterized by uptake assays performed in living cells. Radiolabelled [^{3}H]-xanthine (22.8 Ci mmol^{-1}, Moravek Biochemicals, CA, USA) uptake measurements were performed with *A. nidulans* germinating conidiospores concentrated at 10^7 conidiospores per 100 ml at 37°C, pH 6.8, at the end of the isotropic growth phase, in which UapA shows maximal expression [62]. Competition experiments of [^{3}H]-xanthine uptake were carried out with in-

creasing concentrations of unlabeled L-ascrobic acid (1–60 mM). All transport assays were carried out in at least three independent experiments, with three replicates for each concentration or time point and s.d. was < 20 % in all cases.

Phylogenetic analysis datasets and methods

We selected representative NAT sequences from organisms across all domains of life, representing all major taxa, with particular emphasis in characterized family members of known specificity. For the trees shown in Figure 1 and Supplementary Figure S2, which address in detail the plant and animal NATs (SVCT clades) and the evolution of L-ascorbate transporters, we started by detecting NAT domain containing sequences in complete sequenced genomes, coming from Ensembl release 91, Ensembl Metazoa release 38, and selected species from Ensembl Plants release 38. To detect NAT sequences in the genomes we used HMMER and the PF00860 (*Xan_ur_permease*) Pfam HMM profile, and performed an hmmsearch using the GA (gathering) curated threshold (cut_ga parameter). For the 728 detected sequences, a multiple sequence alignment was built using MAFFT v6.861b in the L-INS-i mode, designed for "set of sequences containing sequences flanking around one alignable domain". The resulting alignment was trimmed using trimAl v1.3 [63] using a gap score cut-off of 0.1. The best-fit evolutionary model selection was selected prior to the phylogenetic inference using ProtTest 3 [64]. 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 [65]. An ML tree was inferred with RAxML v8.2.11 [66] 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). We used the exact same methods for the phylogenetic tree of only the SVCT clade, a subset of the previous tree, consisting of 674 sequences, apart from the MAFFT algorithm used, which was the G-INS-I in this case (assumes that entire region can be aligned). ETE toolkit [42] was used for the execution of the phylogenetic pipelines and the visualization of all phylogenetic trees in Figure 1 and Supplementary Figures S1 and S2. ETE was also used for plotting the NCBI taxonomy tree for the 171 species, and the frequency profile of the subfamilies (Fig. 1). For the tree shown in Supplementary Figure S1, which includes all functionally characterized NATs from bacteria and fungi and their closest homologues from other prokaryotic and eukaryotic microorganisms, we obtained the relative sequences by blastP (https://blast.ncbi.nlm.nih.gov) and performed a multiple sequence alignment with MUSCLE v7.0.26 (http://www.megasoftware.net/). MEGA was used for testing the aligned sequences for optimal amino acid substitution model. According to the AIC, the LG+G+F model [67] was selected and the tree was created using a maximum likelihood (ML), and visualized by FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

ACKNOWLEDGEMENTS

We thank Ioannis Theologidis for critical assessment of the phylogenetic analysis presented herein. This work was supported by a "*Stavros Niarchos Foundation*" research grant.

SUPPLEMENTAL MATERIAL

All supplemental data for this article are available online at www.microbialcell.com.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

COPYRIGHT

© 2018 Kourkoulou *et al.* This is an open-access article released under the terms of the Creative Commons Attribution (CC BY) license, which allows the unrestricted use, distribution, and reproduction in any medium, provided the original author and source are acknowledged.

Please cite this article as: Anezia Kourkoulou, Alexandros A. Pittis and George Diallinas (**2018** Evolution of substrate specificity in the Nucleobase-Ascorbate Transporter (NAT) protein family. **Microbial Cell** 5(6): 280-292. doi: 10.15698/mic2018.06.636

REFERENCES

1. Ohno S (**1970**). Evolution by gene duplication. **Springer-Verlag, Berlin Heidelberg.** doi: 10.1007/978-3-642-86659-3

2. Taylor JS, and Raes J. (**2004**). Duplication and divergence: the evolution of new genes and old ideas. **Annu Rev Genet** 38:615-643. doi: 10.1146/annurev.genet.38.072902.092831

3. Bergthorsson U, Andersson DI, and Roth JR (**2007**). Ohno's dilemma: Evolution of new genes under continuous selection. **Proc Natl Acad Sci USA** 104 (43): 17004–17009. doi: 10.1073/pnas.0707158104

4. Bershtein S, and Tawfik DS (**2008**). Ohno's model revisited: measuring the frequency of potentially adaptive mutations under various mutational drifts. **Mol Biol Evol** 25(11):2311-2318. doi: 10.1093/molbev/msn174

5. Lynch M, and Conery JS (**2000**). The evolutionary fate and consequences of duplicate genes. **Science** 290 (5494):1151-5. doi: 10.1126/science.290.5494.1151

6. Force A, Lynch M, Pickett FB, Amores A, Yan YL, and Postlethwait J (**1999**). Preservation of duplicate genes by complementary, degenerative mutations. **Genetics** 151(4):1531-1545. PMID: 10101175

7. Grauer D, and Wen-Hsiung Li (**2000**). Fundamentals of molecular evolution. **Sunderland, MA. Sinauer**. pp 282-283. doi: 10.1002/jctb.375

8. Hittinger CT, and Carroll SB **(2007**). Gene duplication and the adaptive evolution of a classic genetic switch. **Nature** 449 (7163): 677–681. doi: 10.1038/nature06151

9. Averof M. (**2002**). Arthropod Hox genes: insights on the evolutionary forces that shape gene functions. **Curr Opin Genet Dev** 12(4):386-392. doi: 10.1016/s0959-437x(02)00314-3

10. Tokuriki N, and Tawfik DS **(2009**). Protein dynamism and evolvability. **Science** 324(5924):203-207. doi: 10.1126/science.1169375

11. Soskine M, and Tawfik DS (**2010**). Mutational effects and the evolution of new protein functions. **Nat Rev Genet** 11(8):572-582. doi: 10.1038/nrg2808

12. Goldsmith M, and Tawfik DS **(2012**). Directed enzyme evolution: beyond the low-hanging fruit. **Curr Opin Struct Biol** 22(4):406-412. doi: 10.1016/j.sbi.2012.03.010

13. Starr TN, and Thornton JW (**2016**). Epistasis in protein evolution. **Protein Sci** 25(7): 1204-1218. doi: 10.1002/pro.2897

14. Hochberg GKA, and Thornton JW (**2017**). Reconstructing Ancient Proteins to Understand the Causes of Structure and Function. **Annu Rev Biophys** 46:247-269. doi: 10.1146/annurev-biophys-070816-033631

15. Khersonsky O, and Tawfik DS (**2010**). Enzyme promiscuity: a mechanistic and evolutionary perspective. **Annu Rev Biochem**79:471-505. doi: 10.1146/annurev-biochem-030409-143718

16. Nyathi Y, Wilkinson BM, and Pool MR (**2013**). Co-translational targeting and translocation of proteins to the endoplasmic reticulum. **Biochim Biophys Acta** 1833(11): 2392-402. doi: 10.1016/j.bbamcr.2013.02.021

17. Foot N, Henshall T, and Kumar S (**2017**). Ubiquitination and the Regulation of Membrane Proteins. **Physiol Rev** 97(1):253-281. PMID: 27932395

18. MacGurn JA, Hsu PC, and Emr SD (**2012**). Ubiquitin and membrane protein turnover: from cradle to grave. **Annu Rev Biochem** 81:231-259. doi: 10.1146/annurev-biochem-060210-093619

19. 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(5):404-416. doi: 10.1039/b719777b

20. 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(5):363-372. PMID: 18981714

21. Dialinas G (**2016**). Dissection of Transporter Function: From Genetics to Structure. **Trends Genet** 32(9):576-590. doi: 10.1016/j.tig.2016.06.003

22. 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. PMID: 23097742

23. Bürzle M, Suzuki Y, Ackermann D, Miyazaki H, Maeda N, Clémençon B, Burrier R, and Hediger MA (**2013**). The sodium-dependent ascorbic acid transporter family SLC23. **Mol Aspects Med** 34(2-3):436-454. doi: 10.1016/j.mam.2012.12.002

24. Argyrou E, Sophianopoulou V, Schultes N, and Diallinas G (**2001**). Functional characterization of a maize purine transporter by expression in Aspergillus nidulans. **Plant Cell** 13(4):953-964. PMID: 11283348

25. Niopek-Witz S, Deppe J, Lemieux MJ, and Möhlmann T (**2014**). Biochemical characterization and structure-function relationship of two plant NCS2 proteins, the nucleobase transporters NAT3 and NAT12 from Arabidopsis thaliana. **Biochim Biophys Acta** 1838(12):3025-3035. doi: 10.1016/j.bbamem.2014.08.013

26. 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. doi: 10.1006/jmbi.2001.5087

27. 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(5):1121-1135. doi: 10.1016/j.jmb.2008.08.005

28. 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(3):808-819. doi: 10.1016/j.jmb.2005.12.070

29. 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(5):1132-1143. doi: 10.1016/j.jmb.2010.02.037

30. Västermark A, and Saier MH Jr (**2014**). Evolutionary relationship between 5+5 and 7+7 inverted repeat folds within the amino acid-polyamine-organocation superfamily. **Proteins** 82(2):336-346. doi: 10.1002/prot.24401

31. Yamamoto S, Inoue K, Murata T, Kamigaso S, Yasujima T, Maeda JY,Yukihiro Yoshida J, Ohta K, and Yuasa H (**2010**). Identification and Functional Characterization of the First Nucleobase Transporter in Mammals. **J Biol Chem** 285(9): 6522–6531. doi: 10.1074/jbc.M109.032961

32. Lu F, Li S, Jiang Y, Jiang J, Fan H, Lu G, Deng D, Dang S, Zhang X, Wang J, and Yan N (**2011**). Structure and mechanism of the uracil transporter UraA. **Nature** 14;472(7342):243-246. doi: 10.1038/nature09885

33. Alguel Y, Amillis S, Leung J, Lambrinidis G, Capaldi S, Scull NJ, Craven G, Iwata S, Armstrong A, Mikros E, Diallinas G, Cameron AD, and Byrne B (**2016**). Structure of eukaryotic purine/H(+) symporter UapA suggests a role for homodimerization in transport activity. **Nat Commun** 7:11336. doi: 10.1038/ncomms11336

34. 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(16):2679-2696. doi: 10.1016/j.jmb.2015.05.021

35. Yu X, Yang G, Yan C, Baylon JL, Jiang J, Fan H, Lu G, Hasegawa K, Okumura H, Wang T, Tajkhorshid E, Li S, and Yan N (**2017**). Dimeric structure of the uracil:proton symporter UraA provides mechanistic insights into the SLC4/23/26 transporters. **Cell Res** 27(8):1020-1033. doi: 10.1038/cr.2017.83

36. 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(1):129-145. doi: 10.1111/mmi.12646

37. Coudray N, L Seyler S, Lasala R, Zhang Z, Clark KM, Dumont ME, Rohou A, Beckstein O, and Stokes DL (**2007**). Structure of the SLC4 transporter Bor1p in an inward-facing conformation. **Protein Sci** 26(1):130-145. doi: 10.1002/pro.3061

38. Arakawa T, Kobayashi-Yurugi T, Alguel Y, Iwanari H, Hatae H, Iwata M, Abe Y, Hino T, Ikeda-Suno C, Kuma H, Kang D, Murata T, Hamakubo T, Cameron AD, Kobayashi T, Hamasaki N, and Iwata S (**2015**). Crystal structure of the anion exchanger domain of human erythrocyte band 3. **Science**. 350(6261):680-684. doi: 10.1126/science.aaa4335

39. Geertsma ER, Chang YN, Shaik FR, Neldner Y, Pardon E, Steyaert J, and Dutzler R (**2015**). Structure of a prokaryotic fumarate transporter reveals the architecture of the SLC26 family. **Nat Struct Mol Biol** 22(10):803-808. doi: 10.1038/nsmb.3091

40. Byrne B (**2017**). It takes two to transport via an elevator. **Cell Research 27**:965–966. doi: 10.1038/cr.2017.89

41. Wong S, and Wolfe KH (**2005**). Birth of a metabolic gene cluster in yeast by adaptive gene relocation. **Nat Genet** 37(7):777-782. doi: 10.1038/ng1584

42. Huerta-Cepas J, Serra F, Bork P (**2016**). ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. **Mol Biol Evol** 33(6):1635-1638. doi: 10.1093/molbev/msw046

43. Eddy SR (**2011**). Accelerated Profile HMM Searches. **PLoS Comput Biol**. 7(10):e1002195. doi: 10.1371/journal.pcbi.1002195

44. 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

45. Pantazopoulou A, and Diallinas G **(2006**). The first transmembrane segment (TMS1) of UapA contains determinants necessary for expression in the plasma membrane and purine transport. **Mol Membr Biol** 23(4):337-348. doi: 10.1080/09687860600738239

46. 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(7):e41939. doi: 10.1371/journal.pone.0041939

47. 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(3):499-513. doi: 10.3410/f.1026664.323301

48. Karatza P, Panos P, Georgopoulou E, and Frillingos S (**2006**). Cysteine-scanning analysis of the nucleobase-ascorbate transporter signature motif in YgfO permease of Escherichia coli: Gln-324 and Asn-325 are essential and Ile-329-Val-339 form an alpha-helix. **J Biol Chem** 281(52):39881-39890. doi: 10.1074/jbc.m605748200

49. Gournas C, Amillis S, Vlanti A, and Diallinas G (**2010**). Transport-dependent endocytosis and turnover of a uric acid-xanthine permease. **Mol Microbiol** 75(1):246-260. doi: 10.1111/j.1365-2958.2009.06997.x

50. Gournas C, Prévost M, Krammer EM, and André B (**2016**). Function and Regulation of Fungal Amino Acid Transporters: Insights from Predicted Structure. **Adv Exp Med Biol** 892:69-106. doi: 10.1007/978-3-319-25304-6_4

51. Krypotou E, Evangelidis T, Bobonis J, Pittis AA, Gabaldón T, Scazzocchio C, Mikros E, Diallinas G (**2015**). Origin, diversification and substrate specificity in the family of NCS1/FUR transporters. **Mol Microbiol** 96(5):927-950. doi: 10.1111/mmi.12982

52. Jørgensen ME, Xu D, Crocoll C, Ramırez D, Motawia M S , Olsen C E, Nour-Eldin HH, and Halkier BA (**2017**). Origin and evolution of transporter substrate specificity within the NPF family**. eLife** 6:1–30. doi: 10.7554/eLife.19466

53. Bridgham JT, Ortlund EA, and Thornton JW (**2009**). An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. **Nature** 461(7263):515-519. doi: 10.1038/nature08249

54. Carroll SM, Bridgham JT, and Thornton J C **(2008**). Evolution of Hormone Signaling in Elasmobranchs by Exploitation of Promiscuous Receptors. **Mol Biol Evol** 25(12): 2643–2652. doi: 10.1093/molbev/msn204

55. Wheeler G, Ishikawa T, Pornsaksit V, and Smirnoff N (**2015**). Evolution of alternative biosynthetic pathways for vitamin C following plastid acquisition in photosynthetic eukaryotes. **Elife** 13:4. doi: 10.7554/eLife.06369

56. Oda M, Satta Y, Takenaka O, and Takahata N (**2002**). Loss of urate oxidase activity in hominoids and its evolutionary implications. **Mol Biol Evol** 19(5):640-653. doi: 10.1093/oxfordjournals.molbev.a004123

57. Ames BN, Cathcart R, Schwiers E, and Hochstein P (**1981**). Uric acid provides an antioxidant defense in humans against oxidant- and radical-caused aging and cancer: a hypothesis. **Proc Natl Acad Sci** USA. 78(11):6858-6862. PMID: 6947260

58. Álvarez-Lario B, Macarrón-Vicente J (**2010**). Uric acid and evolution. **Rheumatology** (Oxford) 49(11):2010-2015. doi: 10.1093/rheumatology/keq204

59. Tan PK, Farrar JE, Gaucher EA, and Miner JN (**2016**). Coevolution of URAT1 and Uricaseduring Primate Evolution: Implications for Serum Urate Homeostasis and Gout. **Mol Biol Evol** 33(9):2193-200. doi: 10.1093/molbev/msw116

60. Pantazopoulou A, Lemuh ND, Hatzinikolaou DG, 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(7):627-640. doi: 10.1016/j.fgb.2006.10.003

61. Koukaki M, Giannoutsou E, Karagouni A, and Diallinas G (**2003**). A novel improved method for Aspergillus nidulans transformation. **J Microbiol Methods** 55(3):687-695. doi: 10.1016/s0167-7012(03)00208-2

62. 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. doi: 10.1016/j.fgb.2013.12.004

63. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (**2009**). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. **Bioinformatics** 25(15):1972-1973. doi: 10.1093/bioinformatics/btp348

64. Darriba D, Taboada GL, Doallo R, Posada D (**2011**). ProtTest 3: fast selection of best-fit models of protein evolution. **Bioinformatics** 27(8):1164-1165. doi: 10.1093/bioinformatics/btr088

65. Wagenmakers EJ, Farrell S (**2004**). AIC model selection using Akaike weights. **Psychon Bull Rev** 11(1):192-196. doi: 10.3758/bf03206482

66. Stamatakis A. (**2015**) Using RAxML to Infer Phylogenies. **Curr Protoc Bioinformatics** 51:6.14.1-14. doi: 10.1002/0471250953.bi0614s51

67. Le SQ, Gascuel O (**2008**). An improved general amino acid replacement matrix. **Mol Biol Evol** 25(7):1307-1320. doi: 10.1093/molbev/msn067

Specific Residues in a Purine Transporter Are Critical for Dimerization, ER Exit, and Function

Anezia Kourkoulou,* Pothos Grevias,* George Lambrinidis,† Euan Pyle,‡,§ Mariangela Dionysopoulou,* Argyris Politis,§ Emmanuel Mikros,† Bernadette Byrne,‡ and George Diallinas*,1

*Department of Biology and †Department of Pharmacy, National and Kapodistrian University of Athens, Panepistimioupolis, 15771, Greece and ‡Department of Life Sciences and §Department of Chemistry, King's College London, SE1 1DB, UK

ORCID IDs: 0000-0002-7070-4148 (P.G.); 0000-0002-2820-9338 (G.L.); 0000-0002-3426-726X (G.D.)

ABSTRACT Transporters are transmembrane proteins that mediate the selective translocation of solutes across biological membranes. Recently, we have shown that specific interactions with plasma membrane phospholipids are essential for the formation and/or stability of functional dimers of the purine transporter UapA, a prototypic eukaryotic member of the ubiquitous nucleobase ascorbate transporter (NAT) family. Here, we provide strong evidence that distinct interactions of UapA with membrane lipids are essential for *ab initio* formation of functional dimers in the ER, or ER exit and further subcellular trafficking. Through genetic screens, we identify mutations that restore defects in dimer formation and/or trafficking. Suppressors of defective dimerization restore *ab initio* formation of UapA dimers in the ER. Most of these suppressors are located in the movable core domain, but also in the core-dimerization interface and in residues of the dimerization domain exposed to lipids. Molecular dynamics suggest that the majority of suppressors stabilize interhelical interactions in the core domain and thus assist the formation of functional UapA dimers. Among suppressors restoring dimerization, a specific mutation, T401P, was also isolated independently as a suppressor restoring trafficking, suggesting that stabilization of the core domain restores function by sustaining structural defects caused by the abolishment of essential interactions with specific lipids. Importantly, the introduction of mutations topologically equivalent to T401P into a rat homolog of UapA, namely rSNBT1, permitted the functional expression of a mammalian NAT in *Aspergillus nidulans*. Thus, our results provide a potential route for the functional expression and manipulation of mammalian transporters in the model *Aspergillus* system.

KEYWORDS *Aspergillus nidulans*; UapA; nucleobase; NAT; rSNBT1

TRANSPORTERS are essential transmembrane proteins that catalyze the uptake or efflux of metabolites, nutrients, ions, and drugs across biological membranes. Transporter malfunction, due to genetic mutations or metabolic defects, results in significant cellular or organismal disruption (César-Razquin *et al.* 2015) (http://www.tcdb.org/). Despite their biological and apparent medical importance, knowledge on structure–function relationships in transporters is limited compared to extramembrane hydrophilic proteins. This is in part due to complexity associated with their translocation and cotranslational folding into a membrane lipid bilayer [the ER in eukaryotes or the plasma membrane (PM) in prokaryotes]. Additionally, in eukaryotes, transporters follow specific membrane trafficking, turnover, or recycling routes, which add further complications in understanding their mechanisms of regulating expression, function, and turnover (Lauwers *et al.* 2010; MacGurn *et al.* 2012; Herrmann and Spang 2015; Rabouille 2017; Bouris *et al.* 2019). A further contributory factor is their functional and structural dependence on specific membrane lipids, something has is only recently begun to be explored in detail (Bechara and Robinson 2015; Gupta *et al.* 2017). Further complications for transporter study arise from difficulties in expressing sufficient quantities for downstream structural studies, functional reconstitution in proteoliposomes, or in measuring their kinetics in intact cells where the presence of similar transporters with overlapping specificities complicates the analysis (Rigaud and Lévy 2003).

Copyright © 2019 by the Genetics Society of America
doi: https://doi.org/10.1534/genetics.119.302566
Manuscript received July 31, 2019; accepted for publication October 2, 2019; published Early Online October 14, 2019.
Supplemental material available at figshare: https://doi.org/10.25386/genetics.9177566.
[1]Corresponding author: Department of Biology, National and Kapodistrian University of Athens, Panepistimioupolis Zografou, Athens 15781, Greece. E-mail: diallina@biol.uoa.gr

One of the best studied eukaryotic transporters is the UapA xanthine–uric acid/H^+ symporter of the filamentous ascomycete *Aspergillus nidulans* (Diallinas 2014, 2016). This is due to the development of rigorous genetic, biochemical, and *in vivo* cellular approaches, uniquely available in the model system of *A. nidulans*, and, more recently, structural and biophysical studies. The high-resolution, inward-facing structure of a conformationally locked mutant of UapA ($G411V_{\Delta 1-11}$) revealed that UapA is a closely associated dimer (Alguel *et al.* 2016). Coexpression experiments with wild-type and nonfunctional UapA mutants revealed a dominant negative effect of the nonfunctional mutants, indicating functional interdependence of individual protomers in the UapA dimer and that the dimer is the functional unit (Alguel *et al.* 2016). In the UapA dimer each monomeric unit consists of a core domain that hosts the substrate-binding site and a dimerization domain that includes elements crucial for substrate specificity. Molecular dynamics (MD) simulations together with comparison with other related proteins suggested that UapA functions via an elevator mechanism, as reported for a range of other transporters (Arakawa *et al.* 2015; Geertsma *et al.* 2015; Thurtle-Schmidt and Stroud 2016; Yu *et al.* 2017; Huynh *et al.* 2018; Chang *et al.* 2019). Upon substrate binding, the core domain moves against the relatively immobile dimerization domain to transport the substrate from one side of the membrane to the other (Alguel *et al.* 2016; Drew and Boudker 2016). Interestingly, UapA substrate specificity is also determined by the proper formation of the dimer, as a specific Arg residue (Arg481) from one monomeric unit dynamically controls the substrate translocation trajectory in the opposite monomer (Alguel *et al.* 2016). Recently, native mass spectrometry (MS) of purified UapA, combined with MD, mutagenesis, and functional analyses, established that the membrane lipids, phosphatidylinositol (PI), and phosphatidylethanolamine (PE) have a critical role in stabilizing the functional UapA dimer (Pyle *et al.* 2018). More specifically, it has been shown that UapA delipidation during purification causes dissociation of the dimer into monomers, but subsequent addition of PI or PE both recovers lipid binding and reforms the UapA dimer. MD simulations predicted possible lipid-binding sites near the UapA dimer interface, and subsequent mutational studies confirmed that Arg287, Arg478, and Arg479 act as the lipid-binding residues involved in the formation of UapA dimers and are absolutely necessary for transport activity (Pyle *et al.* 2018). Importantly, however, while triple-alanine replacement of these arginines leads to total lack of UapA function, both native MS and bifluorescence complementation (BiFC) assays have indicated that a fraction of UapA can still dimerize (Martzoukou *et al.* 2015; Pyle *et al.* 2018). Thus, lipid binding might not be an absolute requirement for dimer formation, but is essential for the formation and/or stability of *functional* dimers. The total lack of function of the lipid-binding site UapA mutant (R287A/R478A/R479A) further suggests that specific interactions with lipids are also necessary for the mechanism of transport *per se*. Rather surprisingly, the GFP-tagged inactive triple arginine mutant can still properly traffic to the PM, which means that either monomeric UapA translocates normally to the PM, or that the UapA dimer is initially formed in the ER and traffics to the PM, but then becomes unstable and dissociates into nonfunctional monomers. Noticeably, the triple arginine mutant is also sufficiently folded and stable to allow isolation and native MS analysis, which is also compatible with sorting to the PM. As MD simulations further predicted that lipids could also bind to the outermost, membrane-facing regions of the core domains of the UapA dimer (Pyle *et al.* 2018), further investigation was needed to fully understand the role of membrane lipids in UapA folding, subcellular traffic, and transport function.

Here, we investigate further the role of residues Arg287, Arg478, and Arg479 in UapA dimer formation and/or stability, and study the role of additional putative interactions of UapA with membrane lipids. Using mutational and functional analyses we confirm that Arg287, Arg478, and Arg479 are essential for *ab initio* dimerization in the ER, but redundant for membrane traffic, whereas interactions between distinct peripheral residues (Lys73, Arg133, and Arg421) and lipids are essential for proper folding, ER exit, and membrane traffic. Additionally, using genetic screens for suppressors of mutations affecting putative lipid-binding residues, which could in principle lead to residues critical for structure–function relationships in UapA or to proteins assisting UapA sorting to the PM, we obtained intragenic suppressors in UapA that apparently compensate for the "lost" lipid interactions in the original mutants. Importantly, using information on a specific core residue that proved to have a key role in stabilizing UapA, Thr401, we genetically manipulate and achieve the functional expression, for the first time, of a mammalian homolog of UapA in *A. nidulans*. This opens the way for the functional expression of mammalian transporters, that belong to the nucleobase ascorbate transporter (NAT) family, including those essential for vitamin C transport in humans (Yamamoto *et al.* 2010; Yasujima *et al.* 2018), in the model *Aspergillus* system.

Materials and Methods

Media, strains, and growth conditions

Standard complete (CM) and minimal media (MM) for *A. nidulans* growth were used. Media and supplemented auxotrophies were used at the concentrations given in http://www.fgsc.net. Glucose 1% (w/v) was used as a carbon source, and 10 mM sodium nitrate (NO_3^-) or 10 mM proline was used as nitrogen source. Nucleobases and analogs were used at the following final concentrations: 5-fluorouracil (5FU) 100 μM, uric acid 0.5 mM, xanthine 1 mM, adenine 2 mM, thymine 2 mM, hypoxanthine 2 mM, and uracil 2 mM. All media and chemical reagents were obtained from Sigma (Sigma Chemical), St. Louis, MO (Life Science Chemilab SA, Athens, Hellas) or AppliChem (Bioline Scientific SA, Athens, Hellas). A Δ*azgA* Δ*uapA* Δ*uapC*::*AfpyrG pabaA1*

argB2 mutant strain, named Δ3, was the recipient strain in transformations with plasmids carrying *uapA* alleles, based on complementation of the arginine auxotrophy *argB2*. A Δ*furD*::*riboB* Δ*furA*::*riboB* Δ*fcyB*::*argB* Δ*azgA* Δ*uapA* Δ*uapC*::*AfpyrG* Δ*cntA*::*riboB pabaA1 pantoB100* mutant strain, named Δ7, was the recipient strain in transformations with plasmids carrying the *rSNBT1* wild-type and mutated alleles, based on complementation of the pantothenic acid auxotrophy strain *pantoB100* (Krypotou and Diallinas 2014). *A. nidulans* protoplast isolation and transformation were performed as previously described (Koukaki *et al.* 2003). Growth tests were performed at 25 or 37° for 48 hr at pH 6.8.

Standard molecular biology manipulations and plasmid construction

Genomic DNA extraction from *A. nidulans* was performed as described on the Fungal Genetics Stock Center website (http://www.fgsc.net). Plasmids, prepared in *Eschericia coli*, and DNA restriction or PCR fragments, were purified from agarose 1% gels with the Nucleospin Plasmid Kit or Nucleospin ExtractII kit, according to the manufacturer's instructions (Macherey-Nagel, Düren, Germany Lab Supplies Scientific SA, Athens, Hellas). Standard PCR reactions were performed using KAPATaq DNA polymerase (Kapa Biosystems). PCR products used for cloning, sequencing, and reintroduction by transformation in *A. nidulans* were amplified by a high-fidelity KAPA HiFi HotStart Ready Mix (Kapa Biosystems) polymerase. DNA sequences were determined by VBC-Genomics (Vienna, Austria). Site-directed mutagenesis was carried out according to the instructions accompanying the Quik-Change Site-Directed Mutagenesis Kit (Agilent Technologies and Stratagene, La Jolla, CA). The principal vector used for UapA mutants was pAN510-GFP carrying a *gfp*-tagged *uapA* gene as a template (Koukaki *et al.* 2005), and for *rSNBT1* mutants was a modified pGEM-T-easy vector carrying a version of the *gpdA* promoter, the *trpC* 39 termination region, and the *panB* selection marker (Pantazopoulou *et al.* 2007). For BiFC analyses, the N-terminal half of yellow fluorescent protein (YFPn; 154 amino acids of YFP) or the C-terminal half of YFP (YFPc; 86 amino acids of YFP) was amplified from plasmids PDV7 and PDV8 (Krypotou *et al.* 2015), and cloned into pAN510exp-$alcA_p$ or pAN520exp-$alcA_p$ (Takeshita *et al.* 2008) followed by cloning of the *uapA* ORF. UapA or *rSNBT1* mutations were constructed by oligonucleotide-directed mutagenesis, or appropriate forward and reverse primers (Supplemental Material, Table S2). Transformants arising from single-copy integration events with intact UapA ORFs were identified by PCR analysis.

Uptake assays

Kinetic analysis of UapA or rSNBT1 activity was measured by estimating uptake rates of ^{3}H-xanthine or ^{3}H-uracil uptake respectively (40–80 Ci mmol^{-1}; Moravek Biochemicals, Brea, CA, USA), as previously described in (Krypotou and Diallinas 2014). In brief, ^{3}H-xanthine or ^{3}H-uracil uptake, or competition by excess of other unlabeled substrates, was assayed in *A. nidulans* conidiospores germinating for 4 hr at 37°, at 140 rpm, in liquid MM, pH 6.8. Initial velocities were measured on 10^7 conidiospores/100 μl by incubation with concentrations of 0.2–2.0 μM of radiolabeled substrates at 37°. The time of incubation was defined through time-course experiments and the period of time when each transporter showed linear increased activity was chosen. All transport assays were carried out in triplicate. SD was <20%. Results were analyzed using GraphPad Prism software.

Isolation and characterization of suppressor mutations

Suppressor mutations of 10^9 conidiospores of strains R287A/R478A/R479A or K73A/R133A/R421A were obtained after 3 min and 45 sec exposure at a standard distance of 20 cm from an Osram HNS30 UV-B/C lamp, and subsequent selection of colonies capable of growing on MM containing uric acid as a sole nitrogen source, at 25°. Spores from positive colonies were collected after 6–8 days and further isolated on the same selective medium that was used to obtain the original colonies. Genomic DNA from 24 purified colonies was isolated, and the *uapA* ORF was amplified and sequenced. In all cases, the amplified fragments contained a new single missense mutation.

Epifluorescence microscopy

Samples for standard epifluorescence microscopy were prepared as previously described (Gournas *et al.* 2010; Karachaliou *et al.* 2013). In brief, sterile 35-mm l-dishes with glass bottoms (Ibidi, Gräfelfing, Germany) containing liquid minimal media supplemented with $NaNO_3$ and 0.1% glucose were inoculated from a spore solution, and incubated for 18 hr at 25° or for 8 hr at 37° (rSNBT1). The samples were observed on an Axioplan Zeiss (Carl Zeiss, Thornwood, NY) phase contrast epifluorescent microscope and the resulting images were acquired with a Zeiss-MRC5 digital camera using AxioVs40 V4.40.0 software. Image processing and contrast adjustment were made using ZEN 2012 software, while further processing of the TIFF files was done using Adobe Photoshop CS3 software for brightness adjustment, rotation, and alignment.

Homology modeling of rSNBT1

The construction of a structural model of rSNBT1 was based on the crystal structure of the UapA in the inward-open conformation [Protein Data Bank (PDB) entry 5I6C]. For this, we utilized the alignment shown in Figure S5. The final model was built using PRIME software with an energy-based algorithm (Jacobson *et al.* 2004). A loop refinement routine was also implemented.

Induced-fit docking of uracil on rSNBT1

Protein preparation using OPLS2005 force field (Banks *et al.* 2005) and molecular docking was performed with the Schrödinger Suite 2018. After protein structure alignment with the crystal structure of UapA (PDB 5I6C), the binding pocket was defined by residues Phe124, Glu347, Tyr395, Ser396, and

Glu397. Uracil was docked on the final structure from Homology Modeling, using the induced-fit docking protocol (Schrödinger Release 2018-1: Schrödinger Suite 2018-1 Induced Fit Docking protocol; Glide, Schrödinger, LLC, New York, NY, 2018; Prime, Schrödinger, LLC, New York, NY, 2018), 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).

MD

UapA (wild-type or mutated when discussed) or rSNBT1 homolog dimers were inserted into a lipid bilayer using the CHARMM-GUI tool (Wu *et al.* 2014, http://www.charmm-gui.org). The resulting system was explicitly solvated using the transferable intermolecular potential with 3 points (TIP3P) water model (Jorgensen *et al.* 1983), and neutralized by the addition of Na^+ and Cl^- counter ions at concentrations of 0.15 cM. The lipid bilayer utilized was composed of 20% ergosterol, 9% POPC, 12% DYPC, 9% YOPC, 6% POPE, 3% DYPE, 5% YOPE, 3% DOPE, 19% POPI, and 14% PYPI, as described previously (http://www.charmm-gui.org/?doc=archive&lib=lipid, Pyle *et al.* 2018). All UapA mutations were constructed by utilizing the CHARMM-GUI's initial step "PDB Manipulation Options." The N-terminal residues were always methylated and the C-terminal residues were always amidated. MD simulations were performed with GROMACS 2018 (Abraham *et al.* 2015) using the all-atom force field CHARMM36 (Huang and MacKerell 2013). Periodic boundary conditions were used. Long-range electrostatic interactions were treated with the Particle Mesh Ewald method. Nonbonded interactions were described with a Lennard–Jones potential, with a cut-off distance of 1 nm and an integration step of two femtoseconds (fs) was implemented. The system was progressively minimized and equilibrated using the GROMACS input scripts generated by CHARMM-GUI, and the temperature and pressure were held at 303.15° Kelvin and 1 bar, respectively (Lee *et al.* 2016). The resulting equilibrated structures were then used as an initial condition for the production runs of 100 nsec with all the constraints turned off. Production runs were subsequently analyzed using GROMACS tools, and all images and videos were prepared using VMD software (Humphrey *et al.* 1996).

In silico mutation of I157F, L192F, and L431F on UapA

Starting from the crystal structure of UapA, manual mutation of I157F, L192F, and L431F was performed using the "mutation" command on Maestro v11.5 (Schrödinger Release 2018-1). Each resulting structure was inserted to Protein Preparation Wizard Workflow as implemented on Maestro v11.5. Restrained minimization was converged when heavy atom root-mean-square deviation (RMSD) was >1 Å.

UapA K73A/R133A/R421A (T401P) expression and purification

Expression and purification of the UapA K73A/R133A/R421A and K73A/R133A/R421A/T401P mutants was carried out as described previously (Pyle *et al.* 2018). In brief, the mutants were individually expressed as C-terminally GFP8His-tagged constructs in *Saccharomyces cerevisiae* FGY217 cells (12L), using vector pDDGFP2. Yeast cells were incubated at 30° with shaking at 300 rpm to an OD_{600} of 0.6. Galactose was then added to the cultures to a final concentration of 2% to induce UapA expression. After incubation for a further 22 hr, the cells were harvested by centrifugation and resuspended. Cells were broken in a Constant Systems cell disruptor and the membranes isolated by centrifugation. The membranes were resuspended, flash-frozen, and stored at −80°. Membranes were solubilized for 1 hr in n-dodecyl-β-D-maltoside. Unsolubilized membranes were removed by centrifugation. The supernatant was incubated with Ni^{2+}-NTA resin for 2 hr. The His-tagged UapA bound to the resin was then washed with buffers containing 10 and 30 mM imidazole to remove contaminants. UapA was eluted with buffer containing 250 mM imidazole before overnight dialysis to dilute the imidazole from the protein sample. During dialysis, the protein was cleaved using a His-tagged TEV protease. The sample was run through a His-trap column, from which UapA was eluted in the flow through, to remove the His-tagged GFP and TEV. The sample was then loaded onto a size-exclusion chromatography column. Fractions containing monodisperse UapA were analyzed by SDS-PAGE and concentrated to 20 μM, flash frozen, and stored at −80°.

Native MS of UapA

Native MS of UapA K73A/R133A/R421A/T401P was carried out as described previously (Pyle *et al.* 2018). In brief, UapA was buffer exchanged into MS buffer [250 mM EDDA (pH 6.3), 0.014% DDMLA (v/v), and 10 mM L-serine] to a UapA concentration of 20 μM using Micro Bio-Spin 6 columns (Bio-Rad, Hercules, CA). UapA was loaded into gold-coated capillaries and the protein sprayed into a Synapt G2-Si (Waters Associates, Milford, MA) by nanoelectrospray ionization. The following conditions were used in the mass spectrometer for optimal peak resolution: capillary voltage +1.3–1.5 kV, sampling cone voltage 150 V, trap collision energy (CE) 200 V, transfer CE 0 V, backing pressure 3.88 mbar, trap and transfer pressure (argon) 1.72e-2 mbar, and ion mobility cell pressure (nitrogen) 2.58 mbar. The mass spectrometer was calibrated using cesium iodide. Spectra were recorded and processed using Masslynx 4.1 software (Waters Associates). The relative abundances of each oligomeric state were quantified by UniDec (Marty *et al.* 2015) as described previously (Pyle *et al.* 2018).

Data availability

Strains and plasmids are available upon request. The authors affirm that all data necessary for confirming the conclusions of the article are present within the article, figures, and tables. Supplemental material available at figshare: https://doi.org/10.25386/genetics.9177566.

Results

Residues Arg287, Arg 479, and Arg479 are crucial for ab initio dimerization of UapA in the ER

We have previously shown that arginine residues 287, 478, and 479 are essential for phospholipid-dependent functional dimerization of UapA at the PM (Pyle *et al.* 2018). To further understand the basis of the functional defect in the triple R287A/R478A/R479A mutant, we examined whether loss of dimerization occurs *ab initio* at the level of the ER or whether what we have previously observed was due to instability of UapA dimers at the PM. For this, we used a previously described BiFC assay (Martzoukou *et al.* 2015), adapted to follow the sorting and subcellular localization of *de novo*-made UapA. This was based on time-course experiments following the subcellular localization of *de novo*-made *alcAp*-UapA-GFP, which showed that after 1 hr of transcriptional derepression UapA-GFP was hardly visible, but at 2–3 hr it labeled the ER and at 4 hr appeared mostly in the PM (Bouris *et al.* 2019). Using this system, we followed reconstitution of split-YFP, via UapA dimerization, in a strain containing two copies of the *alcAp-uapA* gene, tagged at either the N- or the C-part of the *yfp* ORF (Martzoukou *et al.* 2015). To follow *de novo*-made UapA in young mycelia, we repressed the transcription of $alcA_p$-*uapA-yfp*$_n$ and *alcAp-uapA-yfp*$_c$ overnight (16 hr in MM at 25° in glucose MM), a period to allow conidiospore germination and young hyphae development, and then shifted the culture to fructose-derepression medium for 1–4 hr of growth. We performed this assay using wild-type UapA and the triple R287A/R478A/R479A mutant. Figure 1A shows that in the wild-type control strain, our assay detected early reconstitution of split-YFP fluorescence at the ER network at 3 hr, but failed to do so in the triple R287A/R478A/R479A mutant, where only a weak signal was observed. After 4 hr of expression, the totality of wild-type UapA fluorescent signal marks the PM, whereas a weak cytoplasmic fluorescent signal and very low cortical localization is observed in the R287A/R478A/R479A mutant. This shows that Ala substitutions of the three Arg residues led to significant reduction of apparent UapA dimerization at the ER membrane, and further suggests that specific contacts with ER lipids might be a prerequisite for dimerization. Surprisingly, the interactions of R287, R478, and R479 with lipids and dimerization proved redundant for sorting of the mutant UapA to the PM, as judged by the normal PM localization of R287A/R478A/R479A tagged with intact GFP (Figure 1B). These findings suggest that in the R287A/R478A/R479A mutant nonfunctional UapA monomers or partially misfolded dimers can still be secreted to the PM.

Genetic suppressors of R287A/R478A/R479A map in the core or dimerization domains

To further understand the molecular basis of how UapA–phospholipid interactions affect the functional dimerization of UapA, we isolated genetic suppressors restoring UapA-mediated growth on uric acid in the R287A/R478A/R479A

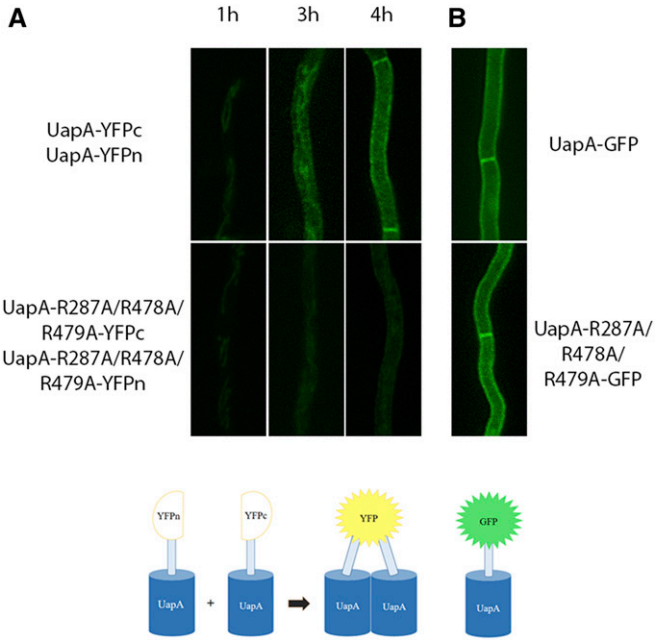


Figure 1 Residues Arg287, Arg 479, and Arg479 are crucial for *ab initio* dimerization of UapA in the ER. (A) Bimolecular complementation analysis based on reconstitution of split-yellow fluorescent protein (YFP) of *de novo*-expressed wild-type or R287A/R478A/R479A mutant, expressed via the $alcA_p$ regulatable promoter in the presence of derepressive carbon and nitrogen, and two copies of the *uapA* gene, one tagged with the N-terminal and the other with the C-terminal part of the *yfp* ORF (UapA-YFPc:UapA-YFPn and UapA-R287A/R478A/R479A-YFPc:UapA-R287A/R478A/R479A-YFPn). Notice the progressive appearance of a clear fluorescent signal in wild-type UapA, first associated with the ER membrane network (3 hr) and finally at the plasma membrane (4 hr). In contrast, in the R287A/R478A/R479A mutant, fluorescence remains extremely low, just above the level of detection. (B) Localization of wild-type or R287A/R478A/R479A mutant UapA tagged with GFP after 4 hr of transcription derepression, showing that the mutant can normally translocate in the plasma membrane despite a very low apparent ability to dimerize, as shown in (A). The lower panel shows schematically that dimerization of UapA is required for detection of a fluorescent signal from reconstituted split-YFP, while a fluorescent signal from GFP-tagged UapA does not distinguish monomers from dimers.

triple mutant at 25°, a temperature at which this mutant does not grow on media containing UapA substrates. We purified 38 apparent suppressors and sequenced the ORF of the *uapA* gene. All 38 contained the original mutation (R287A/R478A/R479A) plus an extra point mutation, apparently the one that suppresses the growth defect on uric acid. Table S1 summarizes the identity and frequency of isolation of all suppressors, which concerned 13 distinct single-amino acid substitutions in 11 different residues. Eight of the 13 suppressors were isolated more than once, showing that mutagenesis was fairly saturated and also confirming that the amino acid changes detected are responsible for suppression. Figure 2A shows the topology of suppressor mutations in the UapA structure.

The 13 distinct suppressors grew well, albeit slightly less than an isogenic strain expressing wild-type UapA, on uric acid or xanthine (Figure 2B). Among the suppressors, S119T

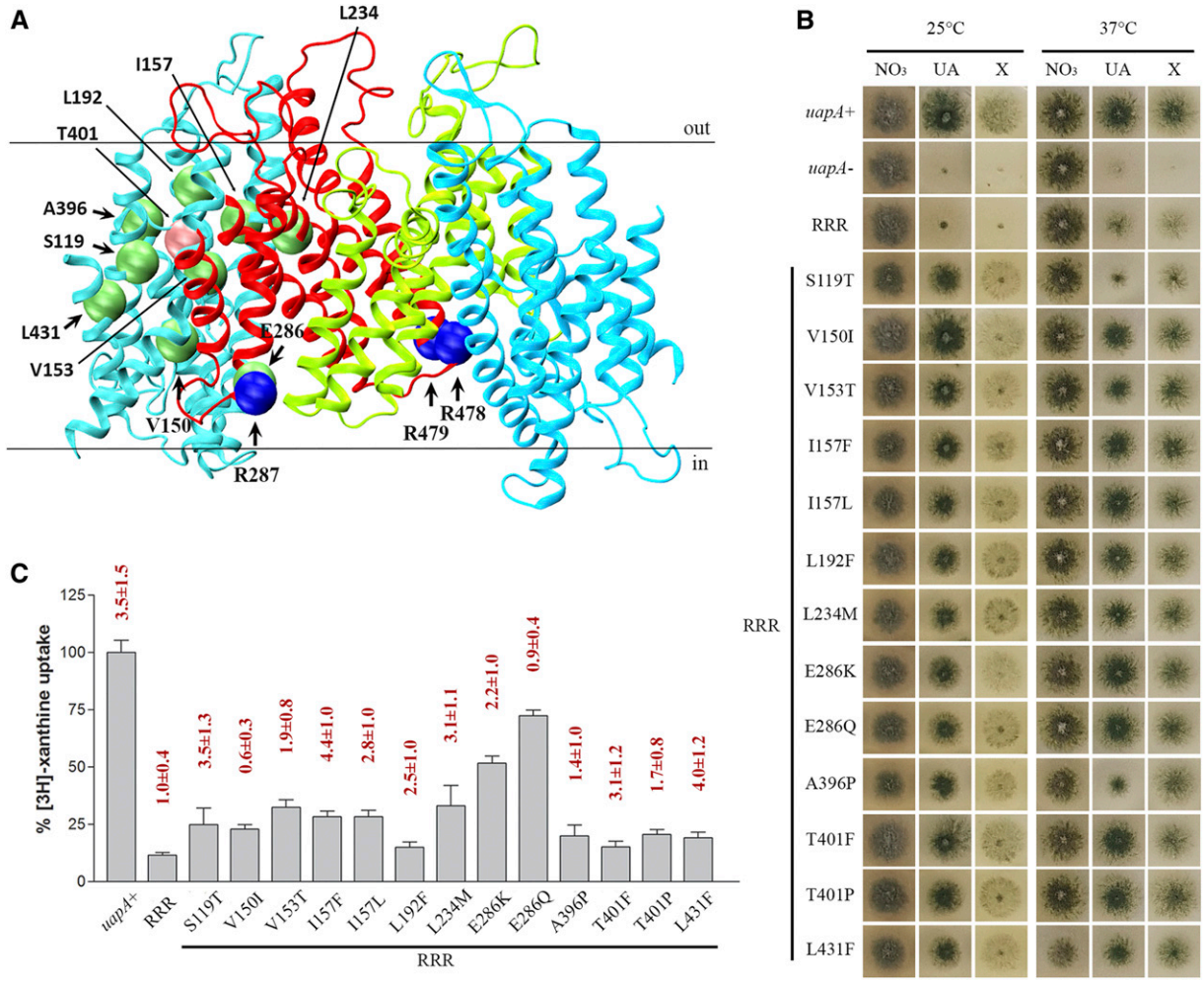


Figure 2 Genetic suppressors of R287A/R478A/R479A localized in the core and the dimerization domain of UapA partially restore UapA function. (A) Topology of amino acids modified in R287A/R478A/R479A suppressors. Core domains are colored light blue, and dimerization domains red and green for clarity. Mutated amino acids in the original mutant strain are shown with blue spheres, and in the suppressors with green and pink (T401) spheres. (B) Growth tests of R287A/R478A/R479A suppressors on UapA physiological substrates. Control strains are strains with total genetic deletions in all major purine transporters (ΔACZ; negative control), referred in the figure as *uapA*-, and a ΔACZ transformant expressing wild-type *uapA–gfp* (*uapA*+; positive control). All suppressor strains and the original R287A/R478A/R479A strain are isogenic to the negative and positive control strains, and express UapA from single-gene copies of *uapA* tagged with GFP. All strains were grown in minimal media containing 10 mM nitrate (NO_3), 0.5 mM uric acid (UA), or 1 mM xanthine (X) as a nitrogen source at 25° (left panel) or 37° (right panel). RRR depicts the R287A/R478A/R479A original genetic background. (C) Relative ^{3}H-xanthine transport rates of R287A/R478A/R479A and R287A/R478A/R479A suppressors expressed as percentages of initial uptake rates (V) compared to the wild-type (*uapA*+) rate. K_m values (μM) for xanthine are shown at the top of histograms. Results are averages of three measurements for each concentration point. SD was 20%.

scored as a thermosensitive mutant, growing very weakly on both UapA substrates at 37°, similar to the original R287A/R478A/R479A mutant. Direct uptake assays, measuring the transport rate of radiolabeled xanthine (Krypotou and Diallinas 2014), were used to estimate the effect of the suppressor mutation on UapA transport kinetics. Figure 2C shows that in most suppressors, UapA transport rates were reestablished at ∼15–30% of the wild-type protein, a level known to be the threshold for conferring visible UapA-mediated growth on uric acid or xanthine. Highest transport rates were obtained in suppressors E286Q (∼70%) and E286K (∼51%). In fact, the relative uptake differences of some suppressors (L192F, A396P, T401P, T401F, or L431F) compared with the original R287A/R478A/R479A mutant were marginal. It should be noted that in growth tests, purines were added to concentrations at the 1–2 mM level to be used as nitrogen sources, while in uptake experiments, radiolabeled xanthine was used at a submicromolar range (0.3–0.5 μM) for technical reasons. Thus, any mutation that causes significant reduction in substrate affinity might score as an apparent loss-of-function mutation in uptake experiments, but still can allow normal growth on the relative substrate when this is supplied at millimolar concentrations. To test whether the low apparent transport capacity of suppressors was due to reduced affinity for xanthine, we estimated the approximate K_m of several of the suppressors relative to the original

R287A/R478A/R479A mutation or a wild-type UapA control. We found no significant reduction of affinity for xanthine in all suppressors tested (Figure 2C, on top of histograms).

Genetic suppressors of R287A/R478A/R479A reestablish UapA dimerization

As phospholipid binding has been shown to be essential for the formation of functional UapA dimers, the suppressors of R287A/R478A/R479A could either restore functional dimerization or confer transport activity to UapA monomers. To test these two alternatives, we performed BiFC assays to follow UapA dimerization of *de novo*-made UapA in selected suppressors (I157F, L234M, and T401P) and control strains, as described earlier for the original R287A/R478A/R479A mutant. Figure 3 shows significant reconstitution of split-YFP fluorescence, and thus apparent dimerization, at both the ER and the PM in all of the three suppressors studied, a picture similar to what is seen in wild-type cells. This confirms that suppressors restored function by restoring early dimerization of UapA at the ER.

MD provides a structural rationale for the effect of suppressor mutations on UapA stability and function

Rather surprisingly, 10 out of the 13 mutations concerned residues located in transmembrane segments (TMSs) of the movable core domain (*i.e.*, in TMS2, TMS3, TMS4, TMS9, TMS10, or TMS11). The core domain consists of two layers of transmembrane helices: TMS1, TMS3, TMS8, and TMS10 in the inner part in contact with the dimerization domain, and TMS2, TMS4, TMS9, and TMS11 in the outer part in contact with the membrane lipids. These two layers are stabilized through an extended network of hydrophobic interactions and certain key polar interactions, which are mainly related to Asp388. Most suppressor mutations are related to the network of the hydrophobic interactions and are located on a virtual diagonal intersecting the core domain in the space between the two transmembrane layers (Figure 4, A and B). The UapA crystal structure (PDB 5i6c) shows that Leu192 interacts with Tyr189 (TMS4), Thr401 (TMS10), Ile157 (TMS3), Phe165 (TMS3), Ile101 (TMS1), Ile193 (TMS4), and Ile346 (TMS8). Similarly, Ile157 interacts with Ile101 (TMS1), Thr401 (TMS10), Pro402 (TMS10), Val349 (TMS8), Ile346 (TMS8), and Leu192 (TMS4). Thr401 is located in the center between the two TMS layers of the core domain and interacts with Tyr189 (TMS4), Leu192 (TMS4), Ile193 (TMS4), Val153 (TMS3), Ile157 (TMS3), and Ile101 (TMS1). Val153 is surrounded by Thr401 (TMS10), Pro97 (TMS1), Val94 (TMS1), Met400 (TMS10), Ser119 (TMS2), Met403 (TMS10), and Cys123 (TMS2). On the opposite side, the Ser119 side chain is located in the middle between Val153 (TMS3), Val94 (TMS1), Met400 (TMS10), and Cys123 (TMS2). Finally, Leu431 is located between Ala87 (TMS1), Met90 (TMS1), and Leu120 (TMS2). Contrary to the above-mentioned residues, Leu234 is located in the dimerization domain, interacting mainly with Ile158 in TMS3. The suppressor mutations—and more particularly L192F,

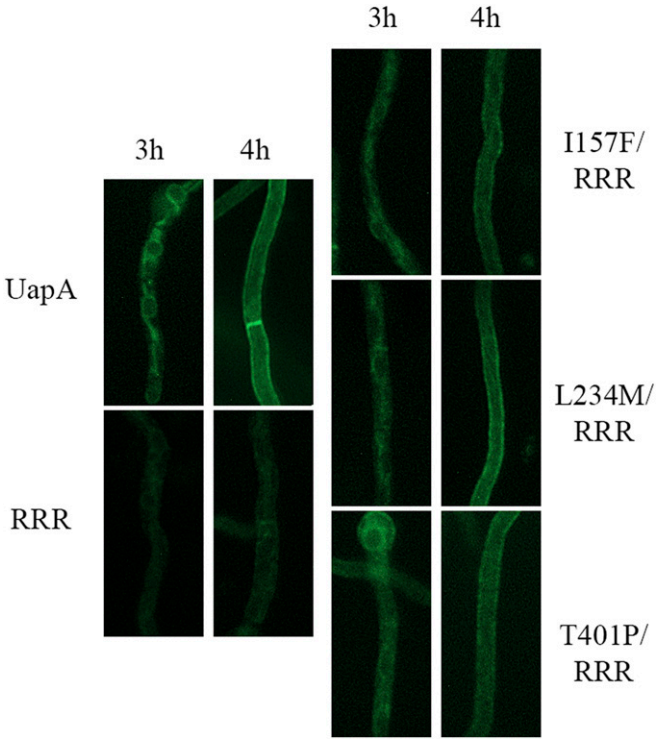


Figure 3 Genetic suppressors R287A/R478A/R479A reestablish UapA dimerization. Bimolecular complementation analysis based on reconstitution of split-YFP of *de novo*-expressed UapA, R287A/R478A/R479A, and selected R287A/R478A/R479A suppressors, namely I157F, L234M, and T401P. Details of expression are as in Figure 1. Notice that all suppressors analyzed clearly restore, at least partially, the strength of the fluorescent signal detected in the ER and the plasma membrane, when compared to the original R287A/R478A/R479A mutant (RRR).

T401F, L431F, S119T, and V153T—seem to enhance the above-mentioned interactions and mainly stabilize the core domain. To verify this hypothesis, models of UapA, including selected suppressor mutations, were constructed and subjected to geometry optimization and short MD calculations. As shown in Figure 4, C–F, the phenyl moieties of the mutated residues I157F, T401F, and L192F are accommodated in the space between the other lipophilic residues, increasing hydrophobic interactions between TMS8, TMS3, TMS4, and TMS10. The introduction of a methyl group in the S119T mutant will also enhance, albeit to a lower degree, interactions with Val153 (TMS3) and Met400 (TMS10). Finally, the polar mutation V153T introduces a new hydrogen bond with Ser119, which already interacts with the backbone of Val94, creating a hydrogen bond network between TMS3, TMS2, and TMS1.

Arg133 and Arg421 are essential for ER exit, function, and lipid-dependent stability

In addition to Arg287, Arg478, and Arg479 forming a lipid-binding site at the dimer interface of UapA, MD simulations identified other, cytosolic-facing residues located on the outside of the core domain that have the potential to form interactions with specific or annular lipids (Pyle *et al.*

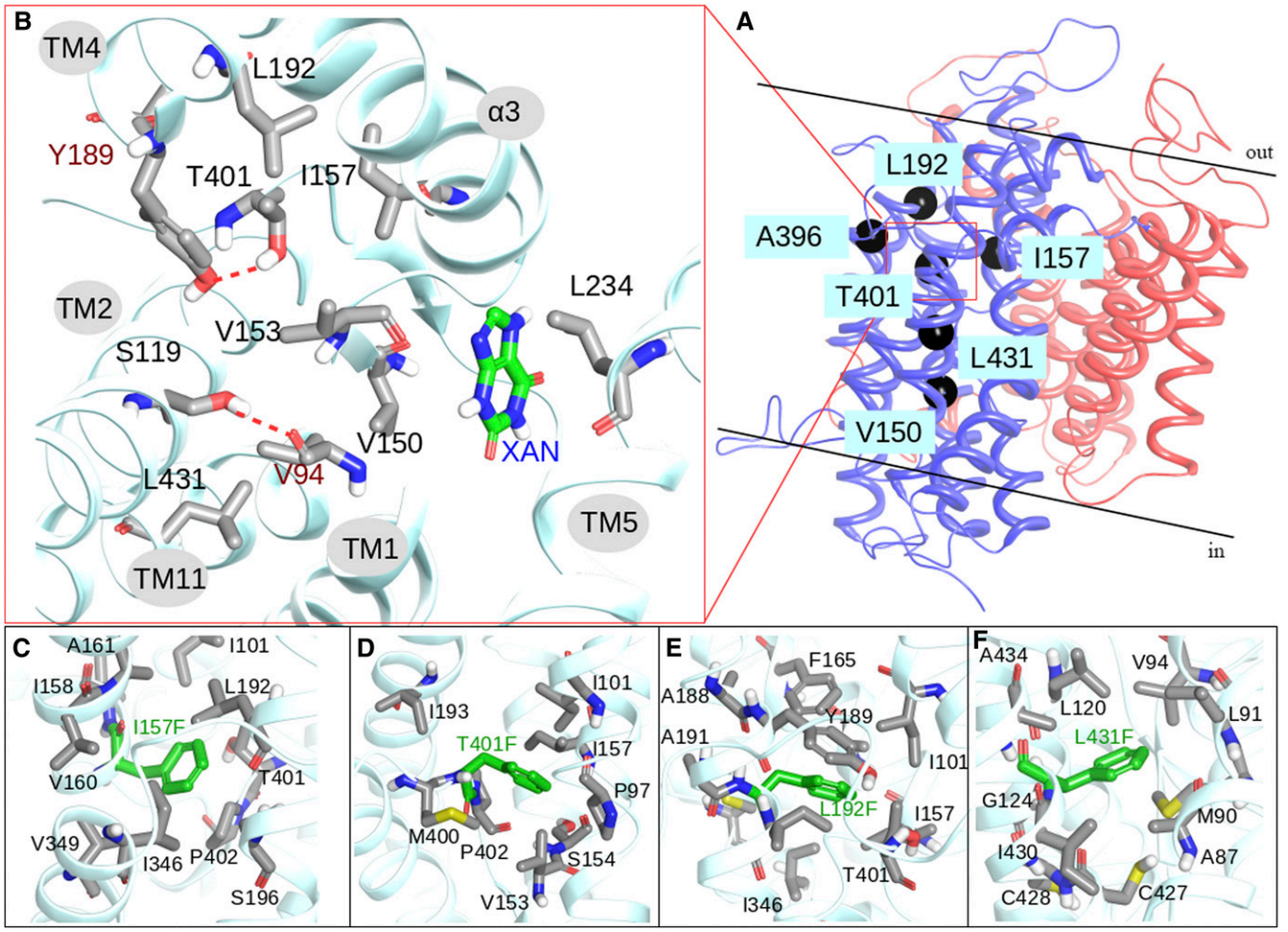


Figure 4 Molecular dynamics provide a structural rationale for the effect of suppressor mutations on UapA stability and function. (A) Ribbon representation of UapA monomer. The core domain is colored blue and the gate domain red. Topologies of type I suppressors on UapA crystal structure are depicted with black spheres. (B) Detailed view of the topology of T401 and amino acids within 4 Å, including residues mutated in suppressors in black lettering (see text). The substrate (xanthine) is also shown. (C)–(F). Detailed view of I157F, T401F, L192F, and L431F mutations and amino acids within 4 Å.

2018). These residues, shown in Figure 5A, are Lys73 in the N-terminus just upstream of TMS1; Arg133, Tyr137, and Lys138 in loop L2; Lys212 in helix ^{1}H of L4; and Arg421 in the border of L10 with TMS11. Of these residues, Arg421 is highly conserved in all NATs, whereas Lys73, Arg133, and Lys212 are well conserved in fungal NATs. Tyr137 and Lys138 are not conserved in other NATs (Figure S1).

To investigate the potential functional role of the above residues, we constructed mutants expressing all single-, and selected double- and triple-Ala substitutions. Figure 5B shows that the single mutations did not affect growth on the UapA substrates (xanthine or uric acid). Of the four double mutants constructed, Y137A/K138A and K73A/R421A exhibited no UapA-related growth defect, while mutant K73A/R133A specifically showed significantly reduced growth on xanthine. Mutant R133A/R421A scored as an apparent total loss-of-function mutant, as it did not grow on either uric acid or xanthine. Finally, the triple K73A/R133A/R421A mutant also scored as a total loss-of-function mutant. Overall, Arg133 and Arg421 proved very important for UapA function, while Lys73 (when present in the context of R133A) was critical for UapA specificity for xanthine, but not for uric acid. For each mutant version of UapA we also assessed localization to the PM, compared to the wild-type UapA, using the GFP epitope attached to UapA. This analysis (right panel in Figure 5B) showed that most single mutations and Y137A/R138A, which led to no defect in UapA transport activity, allowed normal localization of UapA in the PM, as was expected. On the other hand, R133A and K73A/R421A mutants showed partial UapA retention in perinuclear ER membranes. Finally, mutants with apparently defective (K73A/R133A) or lost (R133A/R421A and K73A/R133A/R421A) transport activity showed partial or total retention in the ER. Growth tests and subcellular localization were in good agreement with measurements of rates of radiolabeled xanthine accumulation, which confirmed that the simultaneous presence of Arg133 and Arg421 is essential for transport activity, whereas Lys73 is critical for xanthine uptake only when Arg133 is also

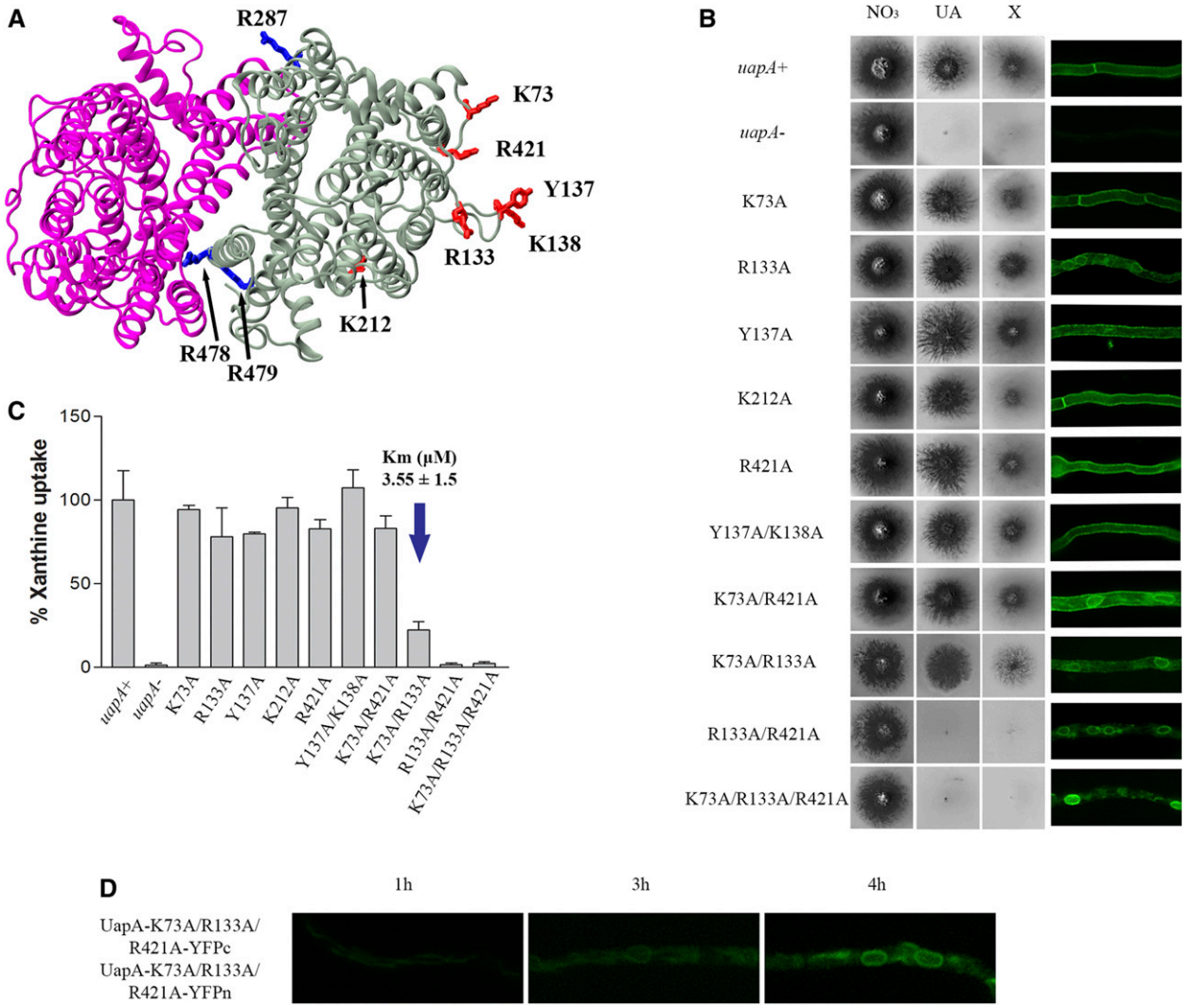


Figure 5 Arg133 and Arg421 are essential for ER exit, function, and lipid-dependent stability. (A) Topology of the predicted lipid-binding sites near the dimer interface (blue) and those in the membrane-facing regions (red) of the UapA dimer. (B) Growth tests of *A. nidulans* strains in minimal media supplemented with nitrate (NO_3), uric acid (UA), or xanthine (X) as a nitrogen source at 37° (left panel). Control strains and concentrations of supplements are as in Figure 2. All UapA mutant strains are isogenic to the negative and positive control strains, expressing *uapA* alleles tagged with *gfp*. Inverted fluorescence microscopy images show localization of the GFP-tagged UapA constructs (ight panel). (C) Relative ^{3}H-xanthine transport rates of UapA mutants expressed as percentages of initial uptake rates (V) compared to the wild-type (*uapA*$^+$) rate. ^{3}H-xanthine uptakes were performed at 37°. The K_m value (μM) for xanthine for mutant K73A/R133A is indicated by a blue arrow. The results are averages of three measurements for each concentration point. SD was 20%. (D) Bimolecular complementation analysis of the K73A/R133A/R421A UapA mutant, performed as previously described in Figure 1 and Figure 3. Notice that the K73A/R133A/R421A mutant retains the ability to reconstitute a fluorescent signal in the ER, but not in the plasma membrane.

replaced by Ala (Figure 5C). For the double mutant K73A/R133A, which showed reduced growth on xanthine, we also measured the K_m for xanthine and showed that this was very close to that of the wild-type UapA (3.6 *vs.* 5 ± 2 μM), suggesting that reduced growth on xanthine is not assigned to reduced substrate binding.

We investigated whether the lack of UapA sorting out of the ER in the triple K73A/R133A/R421A mutant is related to its ability to dimerize. For this, we employed BiFC assays, as described before for the R287A/R478A/R479A mutant. Figure 5D shows that, rather surprisingly, the K73A/R133A/R421A mutant could apparently dimerize in the ER, as a strong fluorescence signal was reconstituted associated with the ER membrane network. This was in line with western blot analysis that detected persisting dimeric forms of K73A/R133A/R421A (Figure S2). These findings suggested that UapA dimerization is not sufficient for ER exit and further sorting to the PM. This is also in agreement with the observation that a lack or reduction of dimerization, seen in the R287A/R478A/R479A mutant (Pyle *et al*. 2018), did not interfere with proper ER exit and sorting to the PM. Thus, functional dimerization and trafficking processes, seemingly affected by distinct lipid-interacting residues, are not necessarily related. Noticeably, attempts to purify the

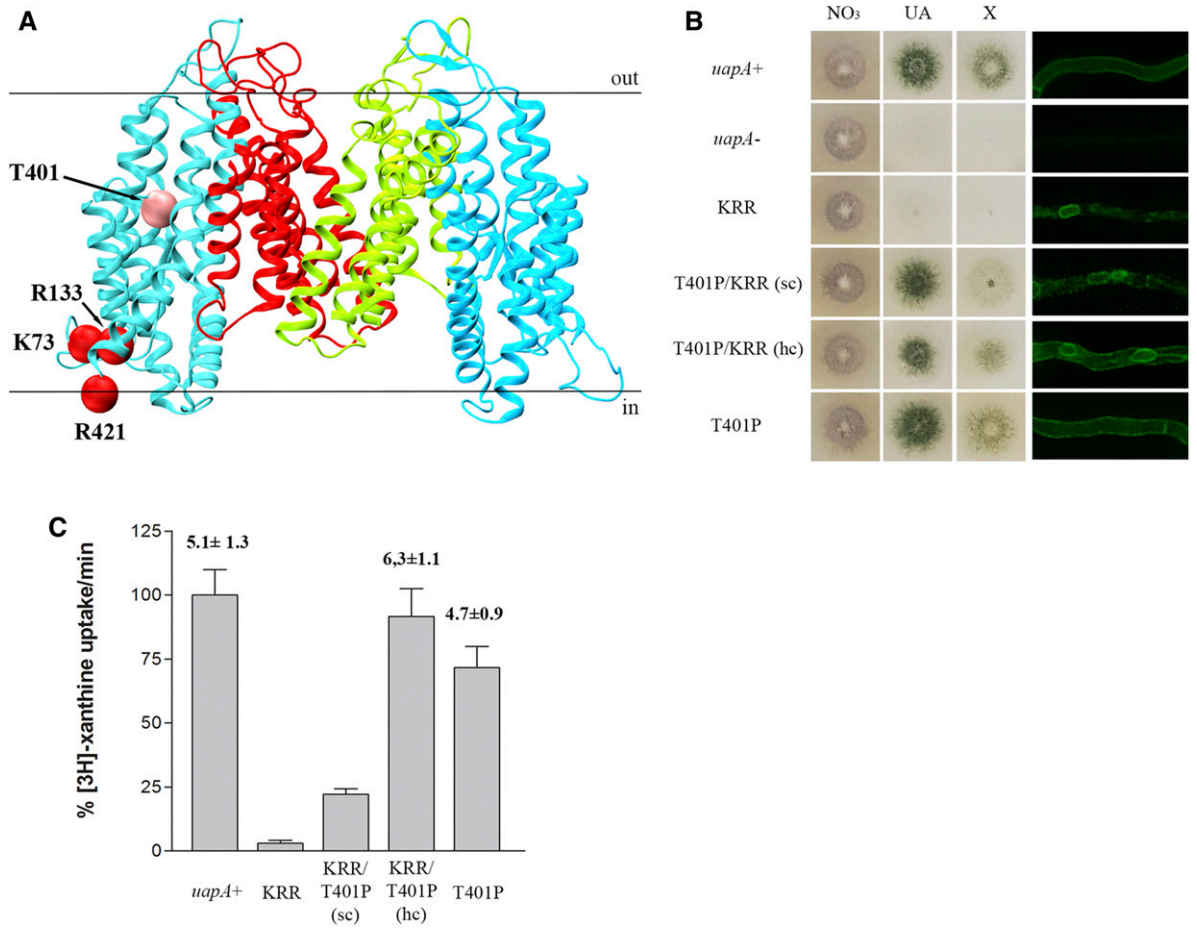


Figure 6 Substitution T401P partially restores the lipid-dependent functional defects of K73A/R133A/R421A. (A) Topologies of amino acids modified in K73A/R133A/R421A suppressors. Core domains are colored light blue, and dimerization domains red and green. Mutated amino acids in the original strain are shown with red spheres and in the suppressor with pink spheres. (B) Growth tests of K73A/R133A/R421A suppressors in minimal media supplemented with nitrate (NO_3), uric acid (UA), and xanthine (X) as nitrogen sources at 37° (left panel). Control strains and supplement concentrations are as in previous Figures. All mutants are isogenic to the negative and positive control strains, expressing *uapA* alleles tagged with GFP. Inverted fluorescence microscopy images show localization of the GFP-tagged UapA constructs (right panel). KRR depicts K73A/R133A/R421A, sc depicts single copy transformants and hc depicts high copy transformants. (C) Relative ^{3}H-xanthine transport rates of UapA mutant versions expressed as percentages of initial uptake rates (V) compared to the wild-type (*uapA+)* rate. ^{3}H-xanthine uptakes were performed at 37°. K_m values (μM) for xanthine are shown. Results are averages of three measurements for each concentration point. SD was 20%.

K73A/R133A/R421A protein showed that this version of UapA is prone to aggregation and is highly unstable (Figure S3). Given that the K73A/R133A/R421A protein is very stable in total extracts, the instability observed upon purification is consistent with the removal of lipids.

Substitution T401P partially restores the lipid-dependent functional defects of K73A/R133A/R421A

We used a genetic approach to understand how residues Lys73, Arg133, and Arg421 might affect UapA sorting to the PM by isolating suppressor mutations that restored UapA-mediated growth on uric acid in the mutant K73A/R133A/R421A, as described earlier for the isolation suppressors of the R287A/R478A/R479A mutant. We obtained, purified, and sequenced the *uapA* ORF from nine suppressors. Rather surprisingly, all proved to include the same single mutation, namely T401P, in addition to the original K73A/R133A/R421A triple mutation (Figure 6A). Noticeably, T401P was also isolated among the suppressors of the dimerization-defective R287A/R478A/R479A mutant. Growth tests showed that although T401P confers normal growth on xanthine and uric acid in the context of K73A/R133A/R421A (Figure 6B), this occurs by only partial restoration of UapA-mediated transport of these purines (Figure 6C). By targeted mutagenesis, we also constructed plasmid vectors carrying K73A/R133A/R421A/T401P and T401P alone, introduced them in the *A. nidulans* strain lacking endogenous nucleobase transporters, and analyzed several purified transformants. Those carrying the quadruple mutation K73A/R133A/R421A/T401P behaved as the original suppressor, confirming that T401P is the causative mutation suppressing the lack of function in R287A/R478A/R479A. Transformants expressing UapA-T401P behaved nearly as well as a wild-type UapA control, showing 70% transport rates, and normal

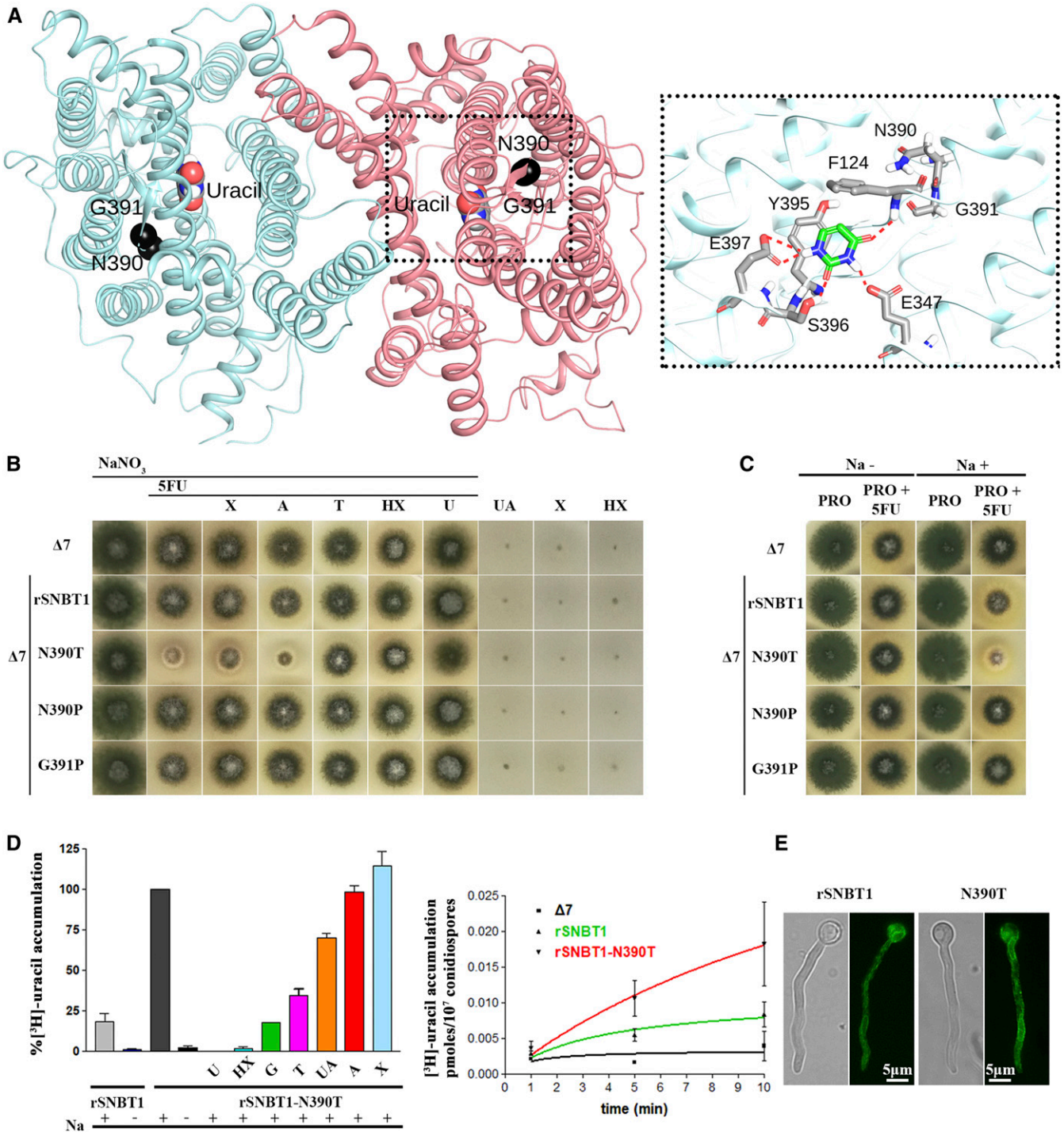


Figure 7 Manipulation of a residue topologically equivalent to T401P leads to functional expression of a mammalian NAT homolog in *A. nidulans*. (A) Homology modeling of the topology of rSNBT1, constructed using, as described in the *Materials and Methods*, the inward-facing conformation of the crystal structure of the UapA dimer. The two mutated and functionally analyzed residues, N390 and G391, are shown as black spheres. The location of uracil, the major substrate of rSNBT1, is also depicted, as determined by dynamic docking (left panel). In the right panel, a zoomed-out picture of the substrate-binding site depicting the major interactions of uracil with specific residues. (B) Growth tests of isogenic *A. nidulans* strains expressing single-copy wild-type rSNBT1 or its mutated versions rSNBT1-N390T, rSNBT1-N390P, and rSNBT1-G391P. A negative control strain (*i.e.*, the recipient Δ7 strain that has null activity for nucleobase transport; see text) is included for comparison. Growth tests were performed at 37° on minimal media supplemented with Na^+ (100 mM NaCl); 10 mM $NaNO_3$ was used as a control nitrogen source unrelated to purine transport activities in all tests scoring resistance/sensitivity to 5FU (rows 1–7). Rows 3–7 represent *in vivo* competition assays scoring the ability of excess purines (2 mM) to compete with the uptake of 5FU (100 μM), and thus revert 5FU sensitivity. X is xanthine, A is adenine, T is Thymine, HX is hypoxanthine, and U is uracil. Notice that T, HX, and U competed with 5FU uptake and suppressed sensitivity. Growth was also scored on minimal media containing UA (uric acid), X, or HX as sole nitrogen sources, none of which supported growth of the strains tested (three last rows). (C) Growth tests of *A. nidulans* of the same strains as in (A),

growth on xanthine or uric acid. Epifluorescence microscopy of K73A/R133A/R421A/T401P and T401P was in line with growth tests and uptake transport measurements. In particular, T401P, in the genetic context of K73A/R133A/R421A, partially restored sorting of UapA to the PM, while when present by itself it did not affect UapA localization to the PM (right panel in Figure 6B). High-copy transformants expressing K73A/R133A/R421A/T401P further confirmed that a major fraction of the mutant UapA translocates to the PM, despite some persistent retention in the ER (see Figure 6B). The positive effect of T401P in the context of K73A/R133A/R421A was further shown by the isolation of fairly stable, detergent-solubilized K73A/R133A/R421A/T401P protein. Native MS further showed that the purified K73A/R133A/R421A/T401P protein, despite being mostly monomeric, also formed a distinct population of dimers (Figure S4). As mentioned above, although attempts were made to isolate and analyze the K73A/R133A/R421A mutant, the construct aggregated completely during purification indicating partial or total misfolding of this protein. Thus, all evidence showed that T401P not only favors functional UapA dimerization in the R287A/R478A/R479A context, but also partially restores ER exit, sorting, and function in the K73A/R133A/R421A context.

Manipulation of a residue topologically equivalent to T401P leads to functional expression of a mammalian NAT homolog in A. nidulans

While *A. nidulans* or *S. cerevisiae* have been successfully used to functionally express plant solute transporters [see early examples in Schachtman *et al.* (1997), von Wirén *et al.* (2000), and Fischer *et al.* (2002)], including a UapA homolog (Argyrou *et al.* 2001), long-standing efforts of our group and many others have failed to functionally express metazoan solute transporters in model fungi. In all cases, metazoan transporters are retained in the ER of fungi and often elicit an unfolded protein response. The simplest explanation for this outcome is that the membrane environment of the fungal ER is incompatible with packaging metazoan transporters into COPII secretory vesicles (Borgese 2016). Based on this idea, and the fact that T401 might be key in restoring lipid-dependent defects in UapA subcellular sorting and function, seemingly by increasing the compactness of the core domain of UapA, we thought we might achieve the functional expression of metazoan NATs in *A. nidulans* by manipulating similar key residues involved in interactions with lipids.

To test this idea, we used a characterized NAT homolog from rat, rSNBT1, which similarly to UapA acts as nucleobase transporter (Yamamoto *et al.* 2010; Yasujima *et al.* 2018). Noticeably, UapA and rSNBT1 have different specificities and cation dependence, as rSNBT1 is a rather promiscuous Na^+ symporter specific for pyrimidines (uracil and thymine) and most purines (hypoxanthine, guanine, xanthine, and uric acid), while UapA is quite specific for xanthine and uric acid, and utilizes H^+ for cotransport. Previous attempts to express rSNBT1 or rSNBT1/UapA chimeric transporters in *A. nidulans* or *S. cerevisiae* have failed, always due to total ER retention (A. Kourkoulou, C. Gournas, S. Amillis, B. Byrne, and G. Diallinas, unpublished data). To putatively identify residue(s) in rSNBT1 that are topologically and functionally equivalent to T401 of UapA, we built a structural model by homology threading using the available UapA crystal structure (Figure 7A and Figure S5). Candidate residues, equivalent to T401 in UapA, proved to be Asn390 and Gly391. Both are part of a conserved, short sequence motif, Gly-Thr-Gly-**Asn**^{390}-**Gly**^{391}, which is present in all metazoan NAT members irrespective of their specificity (Kourkoulou *et al.* 2018). In ascomycetes, the analogous sequence motif is Φ-Thr-Φ-**Thr-Pro** (Φ stands for aliphatic amino acid), while in basidiomycetes and other more primitive fungi it is less well conserved, being Φ-Thr-Φ-Thr/Ser/Ala/Pro-Pro. In other words, what clearly distinguishes metazoans from fungi in this region is the replacement of the Asn residue (390 in rSNBT1) with Thr, and the last Gly (391 in rSNBT1) with a Pro.

Based on the above observations, we constructed and analyzed isogenic *A. nidulans* strains expressing the wild-type tSNBT1, or mutated versions with substitutions N390P, N390T, or G391P, in a genetic background lacking all endogenous transporters related to purine or pyrimidine uptake (see *Materials and Methods*). Strains expressing the wild-type form or mutations N390P and G391P did not grow on purines, and were resistant to 5FU (*i.e.*, a test for scoring uracil uptake), similarly to the recipient negative control strain lacking all endogenous nucleobase transporters. In contrast, the strain expressing rSNBT1-N390T showed clear sensitivity to 5FU and, although it could not grow on any purine, 5FU sensitivity could be competed in the presence of excess purines or pyrimidines that are known rSNBT1 substrates (*e.g.*, hypoxanthine, uracil, or thymine) (Figure 7B). Importantly, rSNBT1-dependent 5FU sensitivity was Na^+-dependent (Figure 7C), compatible with the physiological mechanism of

on minimal media plus proline as a sole nitrogen source, supplemented or not with 100 mM NaCl. Notice that rSNBT1-N390T-mediated 5FU sensitivity is dependent on the presence of Na^+ supplementation. Notice also that in the presence of proline as a nitrogen source, the wild-type rSNBT1 allele confers very moderate sensitivity to 5FU. (D) Left panel: ^{3}H-uracil (0.1 μM) accumulation in strains expressing *rSNBT1* and *rSNBT1-N390T*, performed in the presence or absence of 100 mM Na^+, and in the presence or absence of excess (2 mM) unlabeled nucleobases, after a period of 10 min incubation with radiolabeled substrate. ^{3}H-uracil accumulation in rSNBT1-N390T in the presence of 100 mM Na^+ and absence of unlabeled nucleobase is arbitrarily taken as 100%. Right panel: relative ^{3}H-uracil (0.1 μM) transport accumulation in Δ7 (negative control), *rSNBT1*, or *rSNBT1-N390T* strains as a time course. Uptake results are averages of three measurements for each concentration point. SD was 20%. (E) Inverted fluorescence microscopy images showing the subcellular localization of the GFP-tagged *rSNBT1* and *rSNBT1-N390T* constructs. Notice that the strains used for microscopy are identical to those used in growth tests and uptake assays, as in all cases rSNBT1 sequences were tagged C-terminally with GFP (see *Materials and Methods*).

functioning of rSNBT1 in rat (Yamamoto *et al.* 2010; Yasujima *et al.* 2018).

To further confirm that the phenotype observed in the relative transformants is due uniquely to the genetically introduced rSNBT1-N390T protein, we also analyzed the meiotic progeny of an rSNBT1-N390T transformant. *A. nidulans* undergoing meiosis during a process called "selfing" (van den Brink-van der Laan *et al.* 2004) are prone to high recombination rates that often lead to the loss of sequences introduced by transformation. Figure S6 shows that in an analysis of 28 meiotic progenital colonies of an original rSNBT1-N390T transformant, 21 colonies conserved the original sensitivity to 5FU, while 7 colonies appeared to be 5FU-resistant. Subsequent epifluorescence analysis of selected colonies showed that in all cases 5FU sensitivity was conserved, a fluorescent signal from the rSNBT1-N390T protein tagged with GFP was also conserved. In contrast, all selected colonies that acquired resistance to 5FU lost the fluorescent signal of the rSNBT1-N390T-GFP (not shown).

Finally, we also performed direct measurements of radiolabeled uracil accumulation or competition in the strain expressing rSNBT1-N390T, which further confirmed the functionality of the rat transporter in *A. nidulans* (Figure 7D). Noticeably, the low apparent transport capacity of rSNBT1-N390T with respect to some of its substrates (*e.g.*, hypoxanthine or uric acid) in *A. nidulans* is very probably due by the observation that mutation N390T restores sorting of rSNBT1 to the PM only partially (Figure 7E**).**

Discussion

It is becoming well established that the physicochemical nature of lipid bilayers and the specific lipid composition of membranes affect transporter folding, oligomerization, subcellular trafficking, function, and turnover (van den Brink-van der Laan *et al.* 2004; Koshy *et al.* 2013; Laganowsky *et al.* 2014; Koshy and Ziegler 2015; Martens *et al.* 2016; Gupta *et al.* 2017; Henrich *et al.* 2017; Landreh *et al.* 2017). For transporters conforming to the 5 + 5 inverted repeat or LeuT fold, similarities in structurally resolved lipid–protein interactions suggest common ways in which transporter structure and function are supported by lipid interactions (Drew and Boudker 2016). These are likely to include stabilization of the inverted repeat topology, but also mechanistic roles as major determinants of the alternating access mechanism of secondary transporters. Noticeably however, the great majority of studies on transporter–lipid interactions have to date focused on prokaryotic transporters.

To our knowledge, our previous study on UapA–lipid interactions still remains the only one focusing on a eukaryotic transporter (Pyle *et al.* 2018). In brief, we have shown that UapA, which primarily exists as a dimer, dissociates into monomers upon removal of tightly bound lipids, and that the dimer can be recovered by the addition of PI or PE. Furthermore, as mutagenesis of tentative lipid-binding Arg residues 287, 478, and 479, predicted by MD, abolished lipid binding and function, we have proposed that PI and PE bind at specific sites in the dimer interface, and thus stabilize the dimeric functional form of UapA (Pyle *et al.* 2018). The total lack of transport activity in R287A/R478A/R479A, despite the fact that in the mutant a degree of dimerization and normal sorting to the PM was still evident, suggested that lipid binding may also be directly essential for the mechanism of transport. Here, we studied further the role of Arg287, Arg478, and Arg479, and in parallel investigated the role of binding of lipids at specific peripheral residues of UapA. We showed that Arg287, Arg478, and Arg479 are essential for early *de novo* formation in the ER membrane, a process absolutely essential for transport activity, albeit not for sorting to the PM. In parallel, we identified distinct positively charged residues (Lys73, Arg133, and Arg421), exposed to the PM membrane bilayer, which are essential for ER exit, sorting to the PM, and transport activity, but apparently not essential for the initial formation of dimers in the ER. Thus, the two sets of positively charged residues define two distinct sites of interaction of UapA with membrane lipids, both essential for function, albeit due to different reasons. The distinct defects caused by Ala substitutions at the dimer interface or those exposed to the inner side of the PM bilayer are well supported by epifluorescence microscopy, BiFC assays, and native MS. Thus, while substitutions of Arg287, Arg478, and Arg479 did not affect UapA stability and sorting to the PM, substitutions of Lys73, Arg133, and Arg421 led to significant protein instability and ER retention. Interestingly, in no case did the mutant UapA versions studied elicit an unfolded protein response (results not shown), suggesting that they probably do not lead to significant misfolding. This is particularly interesting in the case of the K73A/R133A/R421A mutant, which is totally blocked within the ER membrane. This observation further suggests that interactions with specific peripheral lipids might be crucial for specific packaging into COPII secretory vesicles and ER exit.

The most original finding of the present work stems from the isolation of genetic suppressors that partially restore defects caused due to modified interactions of UapA with specific lipids at the dimer interface. In principle, genetic reversion of the lost apparent activity of UapA in the original mutants could have led us to either intragenic or extragenic suppressors, the latter probably concerning proteins assisting ER exit and/or sorting of UapA to the PM. However, we did not obtain extragenic suppressors. The reason for that might be found in our recent findings that show that the only currently known factors needed for UapA sorting to the PM are COPII components, actin, and the clathrin heavy chain, all proteins essential for cell viability. Most of the suppressors isolated mapped in the center of the core domain, and less frequently in the dimerization domain. No rational approach or MD studies could have predicted the functional importance of the residues identified via unbiased genetics. How these residues might correct defects in lipid binding became possible *a posteriori* with the help of MD and by taking into account the biophysical nature of residues introduced by suppressor mutations.

We can classify suppressors of the original dimerization mutant, R287A/R478A/R479A, into three types. Type I, which contains the majority of suppressors, map in the center of the core domain in TMS2, TMS3, TMS4, TMS9, TMS10, and TMS11. All, except one, introduce residues with increased hydrophobicity and/or aromaticity (V150I, I157F, I157L, L192F, A396P, T401P, T401F, or L431F). Only V153T introduces a polar residue, while S119T replaces a polar residue with a longer residue of similar properties. MD suggested that all these changes increase the strength of the relative TMS interactions, and thus the compactness of the core domain and the stability of the protein. Type II includes L234M in TMS5 in the middle plane of the dimerization domain. This mutation could increase the strength of interactions between the dimerization and core domains. Type III includes E286Q and E286K at the end of the cytoplasmic-facing part of TMS6. The most logical scenario for these last two suppressors is that they replace directly the interactions with lipids of the nearby mutated Arg287 (*i.e.*, in R287A). This is also in line with the fact that these are the strongest isolated suppressors in respect to UapA transport activity. Thus, our findings, especially those concerning type I and II suppressors, strongly suggest that by stabilizing the core, which is the motile part of the monomeric units that undergoes dynamic up-and-down elevator-like sliding, the dimer is also stabilized and thus function is restored.

Interestingly, all isolated suppressors of the trafficking mutant K73A/R133A/R421A proved to correspond to substitution T401P, a mutation that also restored the dimerization mutant R287A/R478A/R479A. How this is achieved remains quite unclear, mostly because we still do not understand the molecular basis of the trafficking defect in the original mutant. Based on BiFC assays, western blot analysis, and native MS, R287A/R478A/R479A was shown to form dimers, but these seem much more unstable and less abundant with respect to monomers compared to wild-type UapA (Pyle *et al.* 2018). Thus, despite the distinct defects caused by K73A/R133A/R421A and R287A/R478A/R479A triple mutations in UapA subcellular localization, both seem to affect the stability and function of UapA via abolition of essential but distinct interactions with specific lipids. Apparently, mutation K73A/R133A/R421A was more critical than R287A/R478A/R479A for the packaging of UapA into COPII secretory vesicles and ER exit (Bouris *et al.* 2019), which might in turn suggest that specific peripheral lipid interactions are more important for the trafficking of UapA and other structurally similar eukaryotic transporters. In line with our results, it has been recently suggested that lipid binding around domain interfaces of the prokaryotic NhaA Na^+/H^+ exchanger are also involved in stabilizing the core domain during the conformational transitions required for transport by the elevator mechanism (Landreh *et al.* 2017). Thus, it has been speculated that elevator-type antiporters use a subset of specific peripheral or annular lipids as structural support to facilitate large-scale conformational changes within the membrane. Other recent studies using native MS and functional assays have also demonstrated that protein–lipid interactions play a crucial role in stabilizing the dimer form of prokaryotic transporters conforming to the 7 + 7 inverted repeat topology found in transporters using the sliding elevator mechanism of transport (Arakawa *et al.* 2015; Gupta *et al.* 2017).

An impressive finding of our work was that a single mutation, T401P, proved to be a key residue in restoring defective interactions with specific lipids either at the dimer interface or peripherally. We made use of this information and achieved, for the first time, the functional expression of a rat homolog of UapA, rSNBT1, in *A. nidulans*. The successful expression of rSNBT1 in *A. nidulans* strongly supports the idea that the bottleneck in expressing metazoan transporters in fungi is proper folding in an environment of heterologous membrane lipid composition. Our achievement opens the way for further manipulations, via rational design or unbiased genetic screens, of lipid-binding residues in transporters for their functional expression and manipulation in genetically tractable model fungal systems, such as *A. nidulans* or *S. cerevisiae*. Additionally, the successful functional expression of metazoan homologs in *A. nidulans* leads to new routes for identifying and studying the evolution of novel functions and substrate specificities in the NAT family, such as, for example, understanding how human NAT homologs have evolved to become specific for vitamin C rather than nucleobases (Kourkoulou *et al.* 2018).

Acknowledgments

This work was supported by a Stavros S. Niarchos Foundation grant to A.K. and G.D., by computational time granted from the Greek Research and Technology Network in the National High-Performing Computing facility (ARIS) under project number NCS1_Mechanism (pr006040), and a Biotechnology and Biosciences Research Council grant (B.B./N016467/1). E.P. is the recipient of an Imperial College London Institute of Chemical Biology Engineering and Physical Sc.

Literature Cited

Abraham, M. J., T. Murtola, R. Schulz, S. Páll, J. C. Smith *et al.*, 2015 GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1–2: 19–25. https://doi.org/10.1016/j.softx.2015.06.001

Alguel, Y., S. Amillis, J. Leung, G. Lambrinidis, S. Capaldi *et al.*, 2016 Structure of eukaryotic purine/H(+) symporter UapA suggests a role for homodimerization in transport activity. Nat. Commun. 7: 11336. https://doi.org/10.1038/ncomms11336

Arakawa, T., T. Kobayashi-Yurugi, Y. Alguel, H. Iwanari, H. Hatae *et al.*, 2015 Crystal structure of the anion exchanger domain of human erythrocyte band 3. Science 350: 680–684. https://doi.org/10.1126/science.aaa4335

Argyrou, E., V. Sophianopoulou, N. Schultes, and G. Diallinas, 2001 Functional characterization of a maize purine transporter by expression in Aspergillus nidulans. Plant Cell 13: 953–964. https://doi.org/10.1105/tpc.13.4.953

Banks, J. L., H. S. Beard, Y. Cao, A. E. Cho, W. Damm *et al.*, 2005 Integrated modeling program, applied chemical theory (IMPACT). J. Comput. Chem. 26: 1752–1780. https://doi.org/10.1002/jcc.20292

Bechara, C., and C. V. Robinson, 2015 Different modes of lipid binding to membrane proteins probed by mass spectrometry. J. Am. Chem. Soc. 137: 5240–5247. https://doi.org/10.1021/jacs.5b00420

Borgese, N., 2016 Getting membrane proteins on and off the shuttle bus between the endoplasmic reticulum and the Golgi complex. J. Cell Sci. 129: 1537–1545. https://doi.org/10.1242/jcs.183335

Bouris V, O. Martzoukou, S. Amillis, and G. Diallinas, 2019 Nutrient transporter translocation to the plasma membrane via a Golgi-independent unconventional route. bioRxiv. Available at: https://doi.org/10.1101/540203. https://doi.org/10.1101/540203

César-Razquin, A., B. Snijder, T. Frappier-Brinton, R. Isserlin, G. Gyimesi *et al.*, 2015 A call for systematic Research on solute carriers. Cell 162: 478–487. https://doi.org/10.1016/j.cell.2015.07.022

Chang, Y. N., E. A. Jaumann, K. Reichel, J. Hartmann, D. Oliver *et al.*, 2019 Structural basis for functional interactions in dimers of SLC26 transporters. Nat. Commun. 10: 2032. https://doi.org/10.1038/s41467-019-10001-w

Diallinas, G., 2014 Understanding transporter specificity and the discrete appearance of channel-like gating domains in transporters. Front. Pharmacol. 5: 207. https://doi.org/10.3389/fphar.2014.00207

Diallinas, G., 2016 Dissection of transporter function: from genetics to structure. Trends Genet. 32: 576–590. https://doi.org/10.1016/j.tig.2016.06.003

Drew, D., and O. Boudker, 2016 Shared molecular mechanisms of membrane transporters. Annu. Rev. Biochem. 85: 543–572. https://doi.org/10.1146/annurev-biochem-060815-014520

Fischer, W. N., D. D. Loo, W. Koch, U. Ludewig, K. J. Boorer *et al.*, 2002 Low and high affinity amino acid H+-cotransporters for cellular import of neutral and charged amino acids. Plant J. 29: 717–731. https://doi.org/10.1046/j.1365-313X.2002.01248.x

Geertsma, E. R., Y. N. Chang, F. R. Shaik, Y. Neldner, E. Pardon *et al.*, 2015 Structure of a prokaryotic fumarate transporter reveals the architecture of the SLC26 family. Nat. Struct. Mol. Biol. 22: 803–808 [corrigenda: Nat. Struct. Mol. Biol. 23: 462 (2016)]. https://doi.org/10.1038/nsmb.3091

Gournas, C., S. Amillis, A. Vlanti, and G. Diallinas, 2010 Transport-dependent endocytosis and turnover of a uric acid-xanthine permease. Mol. Microbiol. 75: 246–260. https://doi.org/10.1111/j.1365-2958.2009.06997.x

Gupta, K., J. A. C. Donlan, J. T. S. Hopper, P. Uzdavinys, M. Landreh *et al.*, 2017 The role of interfacial lipids in stabilizing membrane protein oligomers. Nature 541: 421–424. https://doi.org/10.1038/nature20820

Henrich, E., O. Peetz, C. Hein, A. Laguerre, B. Hoffmann *et al.*, 2017 Analyzing native membrane protein assembly in nanodiscs by combined non-covalent mass spectrometry and synthetic biology. Elife 6: e20954. https://doi.org/10.7554/eLife.20954

Herrmann, J. M., and A. Spang, 2015 Intracellular parcel service: current issues in intracellular membrane trafficking. Methods Mol. Biol. 1270: 1–12. https://doi.org/10.1007/978-1-4939-2309-0_1

Huang, J., and A. D. MacKerell, Jr., 2013 CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J. Comput. Chem. 34: 2135–2145. https://doi.org/10.1002/jcc.23354

Humphrey, W., A. Dalke, and K. Schulten, 1996 Vmd - visual molecular dynamics. J. Mol. Graph. 14: 33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Huynh, K. W., J. Jiang, N. Abuladze, K. Tsirulnikov, L. Kao *et al.*, 2018 CryoEM structure of the human SLC4A4 sodium-coupled acid-base transporter NBCe1. Nat. Commun. 9: 900. https://doi.org/10.1038/s41467-018-03271-3

Jacobson, M. P., D. L. Pincus, C. S. Rapp, T. J. Day, B. Honig *et al.*, 2004 A hierarchical approach to all-atom protein loop prediction. Proteins 55: 351–367. https://doi.org/10.1002/prot.10613

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, 1983 Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79: 926–935. https://doi.org/10.1063/1.445869

Karachaliou, M., S. Amillis, M. Evangelinos, A. C. Kokotos, V. Yalelis *et al.*, 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. https://doi.org/10.1111/mmi.12184

Koshy, C., and C. Ziegler, 2015 Structural insights into functional lipid-protein interactions in secondary transporters. Biochim. Biophys. Acta 1850: 476–487. https://doi.org/10.1016/j.bbagen.2014.05.010

Koshy, C., E. S. Schweikhard, R. M. Gärtner, C. Perez, O. Yildiz *et al.*, 2013 Structural evidence for functional lipid interactions in the betaine transporter BetP. EMBO J. 32: 3096–3105. https://doi.org/10.1038/emboj.2013.226

Koukaki, M., E. Giannoutsou, A. Karagouni, and G. Diallinas, 2003 A novel improved method for Aspergillus nidulans transformation. J. Microbiol. Methods 55: 687–695. https://doi.org/10.1016/S0167-7012(03)00208-2

Koukaki, M., A. Vlanti, S. Goudela, A. Pantazopoulou, H. Gioule *et al.*, 2005 The nucleobase-ascorbate transporter (NAT) signature motif in UapA defines the function of the purine translocation pathway. J. Mol. Biol. 350: 499–513. https://doi.org/10.1016/j.jmb.2005.04.076

Kourkoulou, A., A. A. Pittis, and G. Diallinas, 2018 Evolution of substrate specificity in the Nucleobase-Ascorbate Transporter (NAT) protein family. Microb. Cell 5: 280–292. https://doi.org/10.15698/mic2018.06.636

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

Krypotou, E., T. Evangelidis, J. Bobonis, A. A. Pittis, T. Gabaldón *et al.*, 2015 Origin, diversification and substrate specificity in the family of NCS1/FUR transporters. Mol. Microbiol. 96: 927–950. https://doi.org/10.1111/mmi.12982

Laganowsky, A., E. Reading, T. M. Allison, M. B. Ulmschneider, M. T. Degiacomi *et al.*, 2014 Membrane proteins bind lipids selectively to modulate their structure and function. Nature 510: 172–175. https://doi.org/10.1038/nature13419

Landreh, M., E. G. Marklund, P. Uzdavinys, M. T. Degiacomi, M. Coincon *et al.*, 2017 Integrating mass spectrometry with MD simulations reveals the role of lipids in Na+/H+ antiporters. Nat. Commun. 8: 13993. https://doi.org/10.1038/ncomms13993

Lauwers, E., Z. Erpapazoglou, R. Haguenauer-Tsapis, and B. André, 2010 The ubiquitin code of yeast permease trafficking. Trends Cell Biol. 20: 196–204. https://doi.org/10.1016/j.tcb.2010.01.004

Lee, J., X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman *et al.*, 2016 CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. J. Chem. Theory Comput. 12: 405–413. https://doi.org/10.1021/acs.jctc.5b00935

MacGurn, J. A., P. C. Hsu, and S. D. Emr, 2012 Ubiquitin and membrane protein turnover: from cradle to grave. Annu. Rev. Biochem. 81: 231–259. https://doi.org/10.1146/annurev-biochem-060210-093619

Martens, C., R. A. Stein, M. Masureel, A. Roth, S. Mishra *et al.*, 2016 Lipids modulate the conformational dynamics of a

secondary multidrug transporter. Nat. Struct. Mol. Biol. 23: 744–751. https://doi.org/10.1038/nsmb.3262

Marty, M. T., A. J. Baldwin, E. G. Marklund, G. K. Hochberg, J. L. Benesch *et al.*, 2015 Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. Anal. Chem. 87: 4370–4376. https://doi.org/10.1021/acs.analchem.5b00140

Martzoukou, O., M. Karachaliou, V. Yalelis, J. Leung, B. Byrne *et al.*, 2015 Oligomerization of the UapA purine transporter is critical for ER-exit, plasma membrane localization and turnover. J. Mol. Biol. 427: 2679–2696. https://doi.org/10.1016/j.jmb.2015.05.021

Pantazopoulou, A., N. D. Lemuh, D. G. Hatzinikolaou, C. Drevet, G. Cecchetto *et al.*, 2007 Differential physiological and developmental expression of the UapA and AzgA purine transporters in Aspergillus nidulans. Fungal Genet. Biol. 44: 627–640. https://doi.org/10.1016/j.fgb.2006.10.003

Pyle, E., A. C. Kalli, S. Amillis, Z. Hall, A. M. Lau *et al.*, 2018 Structural lipids enable the formation of functional oligomers of the eukaryotic purine symporter UapA. Cell Chem. Biol. 25: 840–848.e4. https://doi.org/10.1016/j.chembiol.2018.03.011

Rabouille, C., 2017 Retriever fetches integrins from endosomes. Nat. Cell Biol. 19: 1144–1146. https://doi.org/10.1038/ncb3612

Rigaud, J. L., and D. Lévy, 2003 Reconstitution of membrane proteins into liposomes. Methods Enzymol. 372: 65–86. https://doi.org/10.1016/S0076-6879(03)72004-7

Schachtman, D. P., R. Kumar, J. I. Schroeder, and E. L. Marsh, 1997 Molecular and functional characterization of a novel low-affinity cation transporter (LCT1) in higher plants. Proc. Natl. Acad. Sci. USA 94: 11079–11084. https://doi.org/10.1073/pnas.94.20.11079

Sherman, W., T. Day, M. P. Jacobson, R. A. Friesner, and R. Farid, 2006 Novel procedure for modeling ligand/receptor induced fit effects. J. Med. Chem. 49: 534–553. https://doi.org/10.1021/jm050540c

Takeshita, N., Y. Higashitsuji, S. Konzack, and R. Fischer, 2008 Apical sterol-rich membranes are essential for localizing cell end markers that determine growth directionality in the filamentous fungus Aspergillus nidulans. Mol. Biol. Cell 19: 339–351. https://doi.org/10.1091/mbc.e07-06-0523

Thurtle-Schmidt, B. H., and R. M. Stroud, 2016 Structure of Bor1 supports an elevator transport mechanism for SLC4 anion exchangers. Proc. Natl. Acad. Sci. USA 113: 10542–10546. https://doi.org/10.1073/pnas.1612603113

van den Brink-van der Laan, E., J. A. Killian, and B. de Kruijff, 2004 Nonbilayer lipids affect peripheral and integral membrane proteins via changes in the lateral pressure profile. Biochim. Biophys. Acta 1666: 275–288. https://doi.org/10.1016/j.bbamem.2004.06.010

von Wirén, N., F. R. Lauter, O. Ninnemann, B. Gillissen, P. Walch-Liu *et al.*, 2000 Differential regulation of three functional ammonium transporter genes by nitrogen in root hairs and by light in leaves of tomato. Plant J. 21: 167–175. https://doi.org/10.1046/j.1365-313x.2000.00665.x

Wu, E. L., X. Cheng, S. Jo, H. Rui, K. C. Song *et al.*, 2014 CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. J. Comput. Chem. 35: 1997–2004. https://doi.org/10.1002/jcc.23702

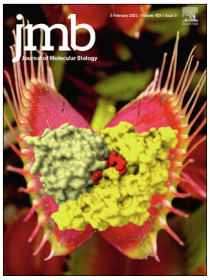
Yamamoto, S., K. Inoue, T. Murata, S. Kamigaso, T. Yasujima *et al.*, 2010 Identification and functional characterization of the first nucleobase transporter in mammals: implication in the species difference in the intestinal absorption mechanism of nucleobases and their analogs between higher primates and other mammals. J. Biol. Chem. 285: 6522–6531. https://doi.org/10.1074/jbc.M109.032961

Yasujima, T., C. Murata, Y. Mimura, T. Murata, M. Ohkubo *et al.*, 2018 Urate transport function of rat sodium-dependent nucleobase transporter 1. Physiol. Rep. 6: e13714. https://doi.org/10.14814/phy2.13714

Yu, X., G. Yang, C. Yan, J. L. Baylon, J. Jiang *et al.*, 2017 Dimeric structure of the uracil proton symporter UraA provides mechanistic insights into the SLC4/23/26 transporters. Cell Res. 27: 1020–1033. https://doi.org/10.1038/cr.2017.83

Communicating editor: M. Schuldiner

Research Article



Context-dependent Cryptic Roles of Specific Residues in Substrate Selectivity of the UapA Purine Transporter

Anezia Kourkoulou [1], Iliana Zantza [2], Konstantina Foti [1], Emmanuel Mikros [2] and George Diallinas [1,3]*

*1 - **Department of Biology,** National and Kapodistrian University of Athens, Panepistimioupolis, Athens 15784, Greece*
*2 - **Department of Pharmacy,** National and Kapodistrian University of Athens, Panepistimioupolis, Athens 15771, Greece*
*3 - **Institute of Molecular Biology and Biotechnology,** Foundation for Research and Technology, Heraklion, Greece*

Correspondence to George Diallinas:*Department of Biology, National and Kapodistrian University of Athens, Panepistimioupolis, Athens 15784, Greece. *diallina@biol.uoa.gr (G. Diallinas)*
https://doi.org/10.1016/j.jmb.2021.166814
Edited by Randy Stockbridge

Abstract

Members of the ubiquitous Nucleobase Ascorbate Transporter (NAT) family are H^+ or Na^+ symporters specific for the cellular uptake of *either* purines and pyrimidines *or* L-ascorbic acid. Despite the fact that several bacterial and fungal members have been extensively characterised at a genetic, biochemical or cellular level, and crystal structures of NAT members from *Escherichia coli* and *Aspergillus nidulans* have been determined pointing to a mechanism of transport, we have little insight on how substrate selectivity is determined. Here, we present systematic mutational analyses, rational combination of mutations, and novel genetic screens that reveal cryptic context-dependent roles of partially conserved residues in the so-called NAT signature motif in determining the specificity of the UapA transporter of *A. nidulans*. We show that specific NAT signature motif substitutions, alone and in combinations with each other or with distant mutations in residues known to affect substrate selectivity, lead to novel UapA versions possessing variable transport capacities and specificities for nucleobases. In particular, we show that a UapA version including the quadruple mutation T405S/F406Y/A407S/Q408E in the NAT signature motif (UapA-SYSE) becomes incapable of purine transport, but gains a novel pyrimidine-related profile, which can be further altered to a more promiscuous purine/pyrimidine profile when combined with replacements at distantly located residues, especially at F528. Our results reveal that UapA specificity is genetically highly modifiable and allow us to speculate on how the elevator-type mechanism of transport might account for this flexibility.

© 2021 Elsevier Ltd. All rights reserved.

Introduction

Solute transporters are transmembrane proteins that mediate the controlled translocation of nutrients, metabolites and drugs across biological membranes, and thus their activity is essential for cell nutrition, detoxification, signalling and homeostasis. The activity of transporters, similar to enzymes, is characterized by kinetic parameters such as affinity for substrate(s), rate of transport, and specificity. In addition, the apparent functional capacity of transporters, also similar to enzymes, depends on regulation of cellular expression and turnover, and in many cases also on oligomerization and allostery.[1] Being polytopic transmembrane proteins, transporter function is also challenged by their continuous and dynamic interaction with membrane lipids, which drives their folding, subcellular trafficking, transport activity and turnover. This in turn signifies that transporter func-

Please cite tts article as: A. Kourkoulou, I. Zantza, K. Foti, et al., Context-dependent Cryptic Roles of Specific Residues in Substrate Selectivity of the UapA Purine TransporterJournal of Molecular Biology, https://doi.org/10.1016/j.jmb.2021.166814

tional evolution should, in principle, run in parallel with the evolution of lipid composition of cellular membranes, specific chaperones necessary for folding, or other effectors mediating proper subcellular trafficking and turnover.[2] Surprisingly, little effort has been dedicated in understanding how the function and specificity of transporters evolve. Addressing the molecular basis of transporter functional evolution, especially in respect to specificity, is interesting not only for understanding basic mechanisms governing protein–ligand interactions in distinct cellular contexts, but is also expected to be crucial in genome mining and predicting novel transporter functions of high biotechnical or biomedical interest. In that direction, we are interested in understanding how the function and specificity of members of the ubiquitous Nucleobase Ascorbate Transporter (NAT) family evolve.

The NAT family (also known as Nucleobase Cation Symporter 2 or NCS2 family) constitutes a large group of transporters present in all domains of life.[3–6] Functionally characterised NATs from bacteria, fungi and plants are specific for nucleobases, but rather surprisingly, mammals and other vertebrates possess NAT homologues that are specific for L-ascorbate transport, defining the evolutionary distinct SVCT1/SVCT2 (SLC23A1/SLC23A2) clade, in addition to nucleobase-specific members, grouped in the SVCT4/SLC23A4 clade.[7–9] Interestingly, humans and higher apes have lost the nucleobase-specific SVCT4/SLC23A4 group, conserving only the SVCT1/2 L-ascorbate-specific homologues.[9] Vertebrates include a third, phylogenetically distinct, highly conserved, clade (the SVCT3/SLC23A3-like group) that is expressed in several tissues, but mostly in the intestine and kidney (http://www.informatics.jax.org/), which remains of unknown function and specificity. Plants also include numerous NATs of unknown specificity, some of which seem to transport substrates other than nucleobases or L-ascorbate.[10,11] In conclusion, NATs constitute an excellent group for studying transporter molecular and functional evolution because: a) they are present in all major taxa, b) significant specificity shifts have occurred within the family, and c) a fungal NAT member (namely UapA from *Aspergillus nidulans;* see below) is among the best-studied *eukaryotic* transporters in respect to regulation of expression and turnover, subcellular trafficking, structure–function relationships, mechanism of transport, specificity, and role of lipid interactions.[12,13]

The UapA protein is a specific xanthine-uric acid/H^+ symporter, which can also transport xanthine analogues, including the commonly used drugs allopurinol and oxypurinol.[14–16] Specific single, double or triple mutations can convert UapA into a more promiscuous nucleobase transporter, able to transport or bind, in addition to uric acid and xanthine, adenine, hypoxanthine, guanine or uracil, as well as, several nucleobase analogues.[13,17,18] A high-resolution crystal structure of UapA dimer, where each monomeric unit consists of a movable *core* and a rather stable *dimerization* domain, trapped in a cytoplasm-facing conformation, has been resolved.[12] This structure confirmed conclusions drawn from genetic studies in respect to residues involved in substrate binding and transport and led to a possible mechanism of transport via sliding of the core domain along the dimerization scaffold.[13]

In the most probable scenario, the mammalian L-ascorbate transporters have evolved from a clade of ancient nucleobase transporters.[9] This shift in substrate specificity, from nucleobases to L-ascorbate, which probably occurred before the divergence of cartilaginous fishes, is a dramatic one as these solutes have very different molecular structures and properties. Nucleobases are planar molecules with relatively low solubility, whereas L-ascorbate is non-planar and highly soluble in water. This difference should necessitate prominent changes in the architecture of the substrate-binding site and substrate translocation trajectory, but maybe also in other regions of the relevant NATs. We still have no insight on the molecular trajectories and mechanisms that led to this specificity shift. Here, we present rational and genetic approaches, combined with relative Molecular Dynamics, addressing the functional role of a conserved NAT motif that forms part of the substrate binding site, within the core domain of UapA.[12,19,20] Our results reveal an unexpected context-dependent cryptic role of specific residues in determining the specificity of UapA towards nucleobases. Our findings also show that genetic conversion of UapA to an L-ascorbate transporter might not be simply achieved by modifications of the NAT signature motif.

Results

Role of the NAT signature motif and rationale of the current study

Previous systematic mutational analysis in UapA have shown that a conserved amino acid motif, named the *NAT signature motif,* located in the 10th transmembrane segment (TMS10) of all NATs includes residues critical for substrate binding and specificity (Q408, R417) or transport catalysis (N409, G411, T416). These five residues are practically irreplaceable for wild-type UapA transport activity, although specific conserved substitutions might still possess low activity or substrate binding, albeit with modified kinetics.[13,17,19,21] Random genetic screens have further shown that a sixth residue of the motif, F406, can also moderately contribute to specificity.[18] In all UapA homologues that are specific for nucleobases the NAT signature motif conforms to the consensus sequence **T/S**/A/V^{405}-**F/Y/**S/L^{406}-**A/S/**T/G/V^{407}-**Q/E**^{408}-**N**^{409}-X-**G**^{411}-X-X-X-X-**T**^{416}-**R/K/G**^{417} (numbering refers to UapA), where highly

conserved residues are shown in bold (Figure 1(A)). This motif has been shown, via modelling and crystallography, to be a major element of the substrate binding site located in the core domain of UapA[12,20] or UraA, an *Escherichia coli* homologue specific for uracil.[22,23] In the UapA crystal, Q408 forms a strong H-bond via its side chains with the substrate (xanthine). A407 and F406 also interact with xanthine via a backbone nitrogen and π-π stacking, respectively, while N409 and T416 seem to be involved in a network of dynamic intramolecular interactions with specific residues in TMS1, TMS3 and TMS8. The flexibility conferred by G411 also proved essential for transport as replacement by Val traps UapA in a non-functional substrate-occluded conformation facing the cytoplasm.[12] Finally, specific substitutions of R417 proved crucial in determining the high affinity of UapA for uric acid compared to xanthine. The aforementioned studies on UapA have been corroborated by mutational and modelling studies with a number of bacterial NATs.[5,24,25]

Interestingly, mammalian NATs that evolved to become highly specific for L-ascorbic acid and their vertebrate orthologues (SVCT1/SVCT2 group), but also homologues of the vertebrate SVCT3/SLC23A3 group that remains functionally uncharacterized, possess modified versions of the canonical NAT signature motif present in nucleobase-specific NATs.[9] As shown in Figure 1 (A), the two critical differences concern substitutions in the first part of the motif. More specifically, the aromatic Phe/Tyr residue (F406 in UapA) is substituted by Ser, while the polar residue Gln/Glu (Q408 in UapA) is replaced by a Pro. Mutant versions of UapA with single (F406S and Q408P), double (F406S/Q408P) or triple substitutions (T405S/F406S/A407S and F406S/A407S/Q408P) introducing residues present in the L-ascorbate transporters proved to have practically no transport activity for either purines or L-ascorbate.[9] These findings pointed out that additional amino acids, possibly outside the major substrate binding site, might need to be substituted for genetically converting UapA into an efficient L-ascorbate transporter.

Interestingly, the most prominent specificity mutations in UapA, selected by direct genetic

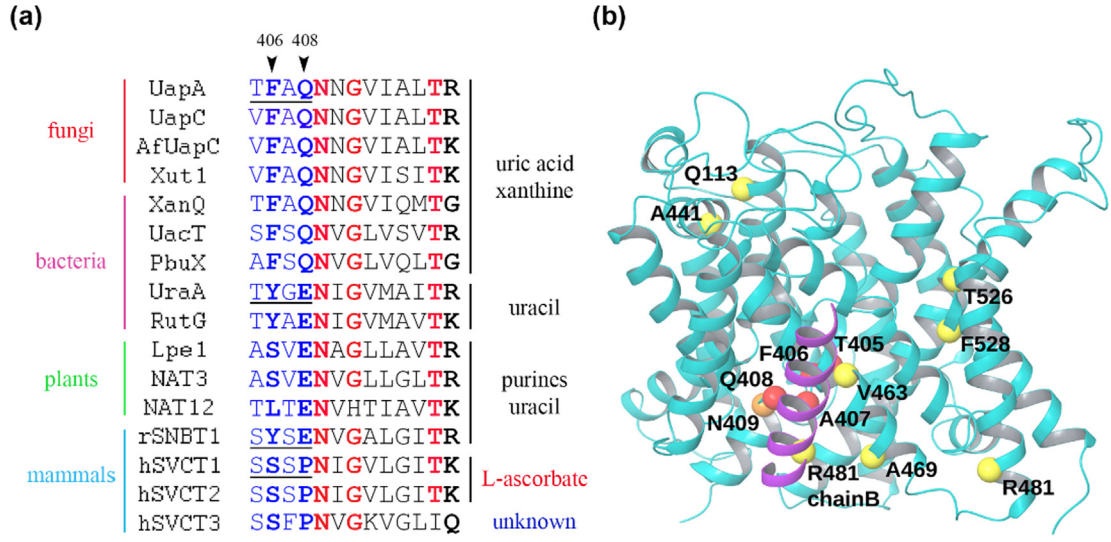


Figure 1. Sequence of the NAT signature motif and topological model of UapA highlighting residues critical for specificity. (A) Alignment of UapA, UapC, AfUapC, Xut1, XanQ, UacT, PbuX, UraA, RutG, Lep1, NAT3, NAT12, rSNBT1 (SVCT4), SVCT1, SVCT2 and SVCT3. UapC (XP_664334), AfUapC (XP_748919) and Xut1 (XP_715538) are xanthine-uric acid transporters from *A. nidulans, Aspergillus fumigatus* and *Candida albicans*. XanQ (P67444), UacT (Q46821) and PbuX (P42086) are xanthine and/or uric acid transporters from *Escherichia coli* and *Bacillus subtilis*. RutG (AAC74091) and UraA (P0AGM7) are uracil transporters from *E. coli*. NAT3 (NP_180219) and NAT12 (NP_850108) are purine-uracil transporters from *Arabidopsis thaliana*. Lpe1 (Q41760) is a maize xanthine-uric acid transporter also showing low affinity for uracil. rSNBT1 (SVCT4/SLC23A4) is a rat purine-uracil transporter. hSVCT1 (Q9UHI7) and hSVCT2 (NP_005107) are human L-ascorbate transporters. hSVCT3 (NP_001138361) is a human transporter of unknown specificity, unrelated to nucleobases, nucleosides or L-ascorbic acid.[9] (B) Topology of the UapA monomeric subunit showing residues affecting (enlarging) UapA specificity: V463, A469 located in TMS12, T526 and F528 at end of TMS14 (outer gate), R481 in start of TMS13 (inner gate), and Q113 and A441 at two flexible helical hinges linking the core and dimerization domains. UapA function as a homodimer (monomers A and B). Notice that R481 of monomeric subunit B (only part of TMS13 of this subunit shown in purple) affects specificity by regulating the translocation of substrates via the trajectory of the opposite monomeric subunit A. All other specificity mutations are located in the interface of the core and dimerization domain in each monomer, and thus affect the sliding of the core, including the substrate binding site, along the dimerization domain (for details see[12]).

screens, map outside the major substrate binding site and the NAT signature motif. These concern little conserved residues located at the interface of the core and dimerization domain (V463, A469 in TMS12), at putative outer (T526 and F528 at the end of TMS14) or inner (R481 at the start of TMS13) *gates*, or at flexible helical hinges linking the core and dimerization domains (Q113, A441) (Figure 1(B)).[12,13] Most of these mutations enlarge the specificity of UapA so that it can transport, besides its natural substrates (xanthine and uric acid), all natural purines, several purine-analogues and uracil. Notably however, no mutation converted UapA into a protein able to transport L-ascorbic acid.

In the present work we a) developed a *novel* strategy for the systematic genetic analysis of specific UapA residues that are variably conserved in the NAT signature motif of UapA, b) used unbiased genetic screens to isolate revertants of a specific UapA NAT signature motif mutant incapable of transporting its physiological substrates, c) run Molecular Dynamics (MDs) for corroborating major conclusions drawn from mutational and functional analyses.

Context-depended role of the NAT signature motif in determining UapA specificity

For the functional analysis of NAT signature motif mutants we used, for the first time, a genetic system that allows the characterization of an extended specificity profile of UapA. This system makes use of a) the strong $gpdA_p$ promoter for expressing *uapA* alleles, some of which might exhibit too little transport activity (<10% of wild-type) to be measured through expression via the native *uapA* promoter, and b) an *A. nidulans* strain, called Δ7, which genetically lacks all seven major endogenous nucleobase-related transporters, thus allowing direct functional assessment of any *uapA* allele introduced by transformation. More specifically, the Δ7 strain includes total deletions of the endogenous *uapA*, *azgA, uapC, fcyB, furD, furA* and *cntA* genes. UapC is a paralogue of UapA exhibiting similar specificity for xanthine and uric acid, albeit distinct transporter kinetics (i.e. lower affinity for uric acid relative to xanthine, and lower transport capacity in general).[26] AzgA is a high-affinity hypoxanthine/adenine/guanine transporter.[27] FcyB is a high-affinity cytosine transporter that also contributes to purine uptake,[28,29] FurD and FurA are high-affinity uracil and allantoin transporters[30,31] and CntA is the single general nucleoside transporter.[32] Due to deletions in the aforementioned genes, strain Δ7 has no detectable transport activity for purines, pyrimidines, nucleosides, allantoin and nucleobase toxic analogues.[33] Thus, Δ7 offers the advantage of scoring, based on simple growth tests and subsequently direct transport assays, whether UapA mutant versions, introduced by genetic transformation, can recognize and transport any nucleobase-related solute. Since pyrimidines, unlike purines, cannot be used as N sources in *A. nidulans*, their accumulation by UapA mutants can be scored indirectly in growth tests based on sensitivity of pyrimidine toxic analogues (e.g. 5-fluorouracil/5FU or 5-fluorocytosine/5FC).[33] Allantoin and purine-nucleosides can be used as N sources in *A. nidulans* and thus UapA mutations can be scored by conferring growth on these substrates. Pyrimidine-nucleoside transport can be scored by sensitivity to 5-fluorouridine (5FUd). Expression via the $gpdA_p$ promoter and the use of the Δ7 strain have been previously used to reveal cryptic functions in the Fur and Fcy groups of the NCS1 family of transporters.[29,31,33]

Using the above system, we functionally analyzed seven UapA versions mutated in the first four variable residues of the NAT signature motif. These are: UapA-T**Y**AQ, -TFA**E**, -T**YAE**, -**SYS**Q, -**SYSE**, -**SSS**Q and -**SSSP**, where in bold are substitutions relative to the TFAQ (T^{405}-F^{406}-A^{407}-Q^{408}) motif present in the wild-type UapA. Among them, UapA-TFA**E, -**T**Y**AQ and -**SSS**Q have been previously functionally analyzed via expression from the *uapA* native promoter and in a genetic background (*uapA*$^-$ *azgA*$^-$ *uapC*$^-$) that did not permit testing of nucleobase-related substrates other than purines. This previous analysis showed UapA-TFA**E** sequence is functional and additionally has acquired the ability to bind (but not to transport) hypoxanthine/guanine, albeit with relatively high affinity,[19] whereas UapA-T**Y**AQ has acquired ability for recognition of all natural purines, but with extremely low binding affinities (K_i-$\geq$ 1 mM).[18] Thus, before the present work, it was known that residues at positions 406 and especially 408 (F/Y^{406} or Q/E^{408}) are critical for UapA specificity. Finally, the UapA-**SSS**Q mutant has been shown to lack any measurable purine transport capacity.[9]

UapA versions carrying T**Y**AQ, TFA**E**, T**YAE**, **SYSE, SYS**Q, **SSS**Q or **SSS**P NAT motif versions were used to examine the role of specific residues at critical positions 406 and 408 (i.e. F/Y/S^{406} and Q/E/P^{408}) in substrate recognition and transport in an otherwise wild type UapA context, but also in variable combinations with neighboring substitutions at positions 405 or 407, basically introducing amino acids present in NATs specific for a L-ascorbic acid (i.e. Ser or Pro) (see Figure 1(A)). Notice that **SYSE** is a sequence frequently present in fungal and metazoan NATs of unknown function and the one predicted to exist in phylogenetically reconstructed ancestral fungal and metazoan NATs (George Diallinas, unpublished analysis). Notice also that **SSSP** introduces the full set of residues found in the signature motif of L-ascorbic acid NATs.[9] By analyzing these mutants we primarily asked which residues are critical for restricting or enlarging UapA

specificity towards nucleobase-related metabolites, but also tried to understand whether changes in the NAT signature motif can lead to a relative shift of specificity from nucleobases to L-ascorbate.

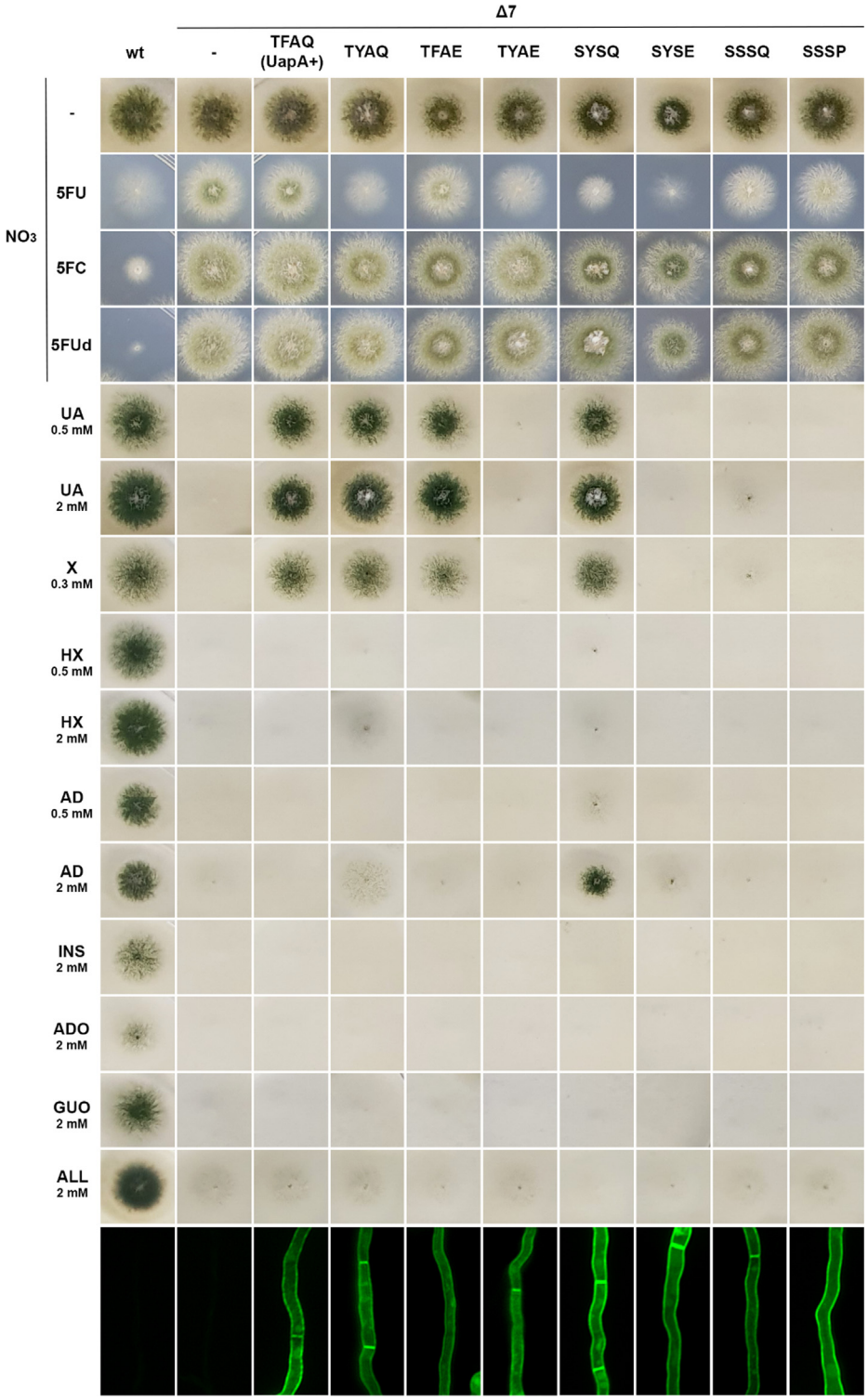
For each of the mutations studied we selected, purified and analyzed several single-copy transformants. In all cases *uapA* mutations were constructed on a plasmid that carries a fully functional chimeric *uapA-gfp* gene so that UapA subcellular localization and stability could also be analyzed by fluorescent microscopy. Figure 2 shows growth tests of mutants and control strains on purines, purine-nucleosides or allantoin as N sources, and on media containing toxic concentrations of 5FU, 5FC and 5FUd. As expected, the negative control strain Δ7 does not grow on purines, purine-nucleosides or allantoin and is resistant to 5FU, 5FC and 5Fud, which is nearly the mirror image of the growth phenotype of a standard wild-type *A. nidulans* strain. The positive control strain (Δ7::*uapA*$^{+}$) grows on uric acid and xanthine, but not on other nucleobase-related compounds tested, and is resistant to 5FU, 5FC and 5FUd, as expected given the specificity of UapA. Importantly, none of the mutations analyzed conferred instability or lack of proper localization of UapA in the plasma membrane (Figure 2, lowest panel), showing that the mutations studied have not affected UapA folding and subcellular trafficking, which in turn strongly suggested that changes in growth reflect direct effects on UapA activity.

The NAT motif versions that allowed wild-type (i.e. TFAQ-like) growth on physiological substrates of UapA (i.e. uric acid or xanthine) were those including the T**Y**AQ, TFA**E** and **SYS**Q sequences. Notably, **SYS**Q conferred growth also on high adenine concentration, while T**Y**AQ led to very weak growth on hypoxanthine. In addition, UapA-T**Y**AQ and more prominently UapA-**SYS**Q, led to increased sensitivity to 5FU. This confirmed that position 406 in UapA (F/Y^{406}) is critical for specificity, and showed for the first time that the presence of Tyr residue increases recognition of 5FU. Combinations of mutations further revealed the contribution of all NAT residues studied in specificity. For example, the single substitution Q408E (e.g. TFA**E**) led to no effect on transport specificity, at least within the limit of growth tests, but in combination with other neighboring substitutions (i.e. in T**Y**A**E** or **SYSE**) led to distinct outcomes. UapA-T**Y**A**E** could not confer growth on any purine or other nucleobase-related solutes and led to resistance to toxic nucleobase analogues, except from some evidence for relatively increased 5FU sensitivity, as compared to wild-type UapA (UapA-TFAQ). UapA-**SYSE,** despite showing no apparent transport capacity for purines, nucleosides or 5FC and 5FUd, conferred high sensitivity to 5FU, suggesting that it acts as a 5FU carrier. Notably, UapA versions carrying NAT sequences mimicking the motif of L-ascorbic acid transporters showed no apparent capacity for nucleobase transport (e.g. **SSSP**), although the one that conserves Q408 (e.g. **SSS**Q) seemed to confer some growth on high concentration of UA. Both, however, seem to confer some low level accumulation of 5FU not seen in the wild-type UapA control (Figure 2). Thus, all mutant versions of UapA studied seem to retain some transport capacity, albeit for distinct substrates. In addition, the growth tests strongly suggested that all four partially conserved residues of the first part of the NAT signature motif of UapA contribute, albeit at different degrees, to specificity.

We performed direct uptake assays, using radiolabelled xanthine, in all mutants studied. Figure 3(A) shows that, in line with growth tests, UapA-T**Y**AQ and UapA-**SYS**Q conserve high capacity for xanthine transport (>70% of wild-type UapA), whereas all other mutants show very low relative transport (≤10% of wild-type UapA). In the UapA mutants that conferred highest sensitivity to 5FU, namely UapA-**SYS**Q and UapA-**SYSE,** we also measured their capacity for accumulation of radiolabelled uracil, which is structurally very similar to 5FU. Surprisingly, none of the two mutants showed measurable UapA-mediated accumulation of uracil. This is discussed further later. We estimated the K_i values for purines and uracil in **SYS**Q and T**Y**AQ mutants, which conserved considerable transport levels of radiolabeled xanthine, but also in TFA**E**. Figure 3 (B) shows that the binding affinity for the physiological substrates xanthine and uric acid in all three mutants did not differ significantly from those of wild-type UapA (i.e. at the 2–15 μM range). However, unlike wild-type UapA, the mutants exhibited variable, mostly very low but measurable, binding affinity for adenine, hypoxanthine or uracil.

We took advantage of the fact that several mutants tested are sensitive to 5FU to investigate whether some of them, despite being unable to transport efficiently purines, can still bind purines or purine-related compounds. For this, we performed *in vivo* competition growth assays where we scored the reversion of 5FU inhibition in the presence of excess purines or pyrimidines in the growth medium (Figure 3(C)). This test showed that 5FU sensitivity mediated by UapA-**SYS**Q is significantly reversed by most nucleobases tested, while that of UapA-**SYSE** is also reduced, but at a lower degree. Notice also that UapA-T**Y**AQ, which, otherwise, is very similar in phenotype with UapA-**SYS**Q (Figs. 2, 3A and 3B), appears to be less efficient in reversing the 5FU sensitivity with some nucleobases (e.g. xanthine). Finally, the moderate 5FU sensitivity conferred by UapA-**SSS**Q and UapA-**SSSP** was also reversed by uracil and less so by uric acid.

In an effort to investigate whether any of the mutants analyzed also alters the capacity of UapA in recognizing L-ascorbic acid, we tested the growth of all mutants in the presence of relatively high concentration of L-ascorbic acid (50 mM or 1%). Notice that L-ascorbic acid is not taken up by *A. nidulans* cells as there is neither a specific endogenous transporter, nor evidence for passive diffusion (G. Diallinas, unpublished observations). Thus, we thought that, if any of the UapA mutations conferred a capacity for L-ascorbate accumulation, this might be reflected in an effect



on the growth phenotype of relative mutants, given the well-known anti- or pro-oxidant activity of L-ascorbate.[34] None of the UapA mutants altered the mild toxic effect (i.e. reduced conidiation) of L-ascorbic acid on *A. nidulans* growth (not shown). This result was not however conclusive on whether any of the UapA mutants can transport or at least bind L-ascorbic acid(discussed later).

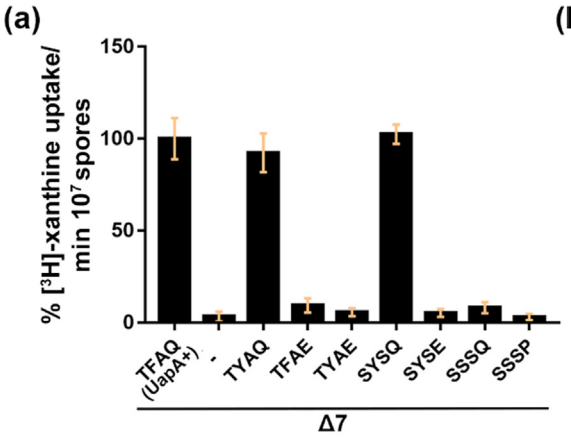
Based on all aforementioned results, Table 1 summarizes the function and specificity of the different UapA versions analyzed. Notice that UapA-**SYS**Q is the most efficient and promiscuous UapA version, being able to recognize and transport purines and 5FU, and it can also at least bind uracil. In contrast, UapA-**SYSE,** which differs only at a single residue with UapA-**SYS**Q (i.e. Q408E), proves to be a mutant that has lost its capacity for efficient purine transport, but acquired a prominent high capacity for 5FU accumulation. However, also notice that if we compare the relative transport function and specificity of UapA-TFAQ *versus* UapA-TFA**E**, which differ in the same residue as UapA-**SYS**Q differs from UapA-**SYSE**, the outcome is different, as replacement Q408E does not lead to total loss of purine uptake or acquisition of high capacity for 5FU transport. This observation clearly shows that UapA specificity is determined by complex interactions of residues in the NAT signature motif, difficult to predict *a priori.*

Combination of the SYSE sequence with selected UapA mutations leads to distinct transport profiles

A **SYSE** consensus sequence is naturally present in several NATs, including the metazoan rSNBT1 that functions as promiscuous nucleobase transporter.[7] However, a **SYSE** sequence within the context of UapA led to inability for purine transport and acquisition of a capacity for 5FU accumulation, strongly suggesting that the NAT motif interacts functionally and in a context-dependent manner with other residues in NATs. To address this issue, we combined the **SYSE** sequence with mutations R481G, T526L, F528M, R481G/T526L and R481G/F528M. These mutations enlarge UapA specificity to transport, besides its natural substrates uric acid and xanthine, all purines, several purine analogues and uracil, albeit with low affinity.[17,18] Topological evidence based on the UapA crystal structure and Molecular Dynamics support that the relevant residues are parts of outer (T526 and F528) or inner (R481) gating elements located at the two sides of the substrate translocation trajectory in UapA.[12]

Figure 4(A) summarizes growth tests of combinations of **SYSE** with these mutations. Notice that in all cases the combined mutations did not affect the proper localization of UapA to the PM (see right column). As a general picture, the increased promiscuity of UapA in respect to purine accumulation caused by F528M, T526L, R481G, R481G/F528M or R481G/T526L (see rows 3–7) was significantly suppressed in the presence of **SYSE** (see rows 9–13). Only F528M/**SYSE** conferred detectable growth on purines, especially on high concentrations of adenine. In respect of 5FU toxicity, which was very high in UapA-**SYSE** and some specificity mutants (T526L, R481G/T526L or R481G/F528M), the simultaneous presence of the **SYSE** sequence and specificity mutations led to variable toxicity levels that could have not been predicted *a priori*. For example, the presence of R481G led to partial suppression of the ability of UapA-**SYSE** to accumulate 5FU, as evidenced by increased resistance to this toxic analogue exhibited by UapA-R481G/**SYSE**. The triple mutant UapA-R481G/T526L/**SYSE** showed moderate sensitivity to 5FU, similar to UapA-**SYSE,** thus 'masking' the hypersensitivity of the parental double UapA-R481G/T526L mutant, while UapA-R481G/F528M/**SYSE** remained hypersensitive to 5FU, similar to the relative double mutants UapA-R481G/F528M or UapA-F528M/S**YSE**. Furthermore, mutation F528M increased sensitivity to 5FU in the context of **SYSE** or/and R481G, while when it stands alone conferred reduced sensitivity to this drug. Notice also that

Figure 2. Context-depended role in UapA specificity of variably conserved residues in the NAT signature motif. The image shows growth tests of UapA mutants and a control strain (referred by the sequence of their NAT motif, e.g. wild-type UapA is TFAQ, etc.) on toxic nucleobase/nucleoside analogues at 100 μM 5FU, 50 μM 5FC, 10 μM 5FUd (5FU is 5-fluorouracil; 5FC is 5-flurocytosine; 5FUd is 5-fluorouridine), or on different purines (UA is uric acid; X is xanthine, HX is hypoxanthine; AD is adenine), nucleosides (INS is inosine; ADO is adenosine, GUO) is guanosine) or allantoin (ALL), as sole N sources. Concentrations of purine-related substrates used as N sources are in the range of 0.3–2.0 mM. Toxic analogues are scored in the presence of 10 mM sodium nitrate (NO_3) as nitrogen source. The lower panel shows the subcellular localization of UapA mutants, as evidenced via a GFP tag, in all strains analyzed. All UapA mutant versions shown are analyzed in the isogenic background of Δ7 strain (see text). A standard wild-type (wt) strain possessing all relative endogenous nucleobase-related transporters is also shown for comparison in the first column. Growth tests were performed at 37 °C and pH 6.8. Lower panel: Inverted fluorescence microscopy images showing localization of the GFP-tagged UapA constructs. Samples were grown for 18 h at 25 °C on MM with NO_3 as N source.



(b)

Substrate binding profile				
	$K_{m/i}$ (μM)			
Substrate/ligand	TFAQ (UapA+)	TFAE	TYAQ	SYSQ
Xanthine	4±1.5	9±1.1	3±0.8	2±1.1
Uric acid	6±1.2	15±1.2	11±1.3	5±1.0
Adenine	NI	>1000	NI	>1000
Hypoxanthine	NI	71±1.2	>1000	~1000
Uracil	NI	>1000	>1000	~1000

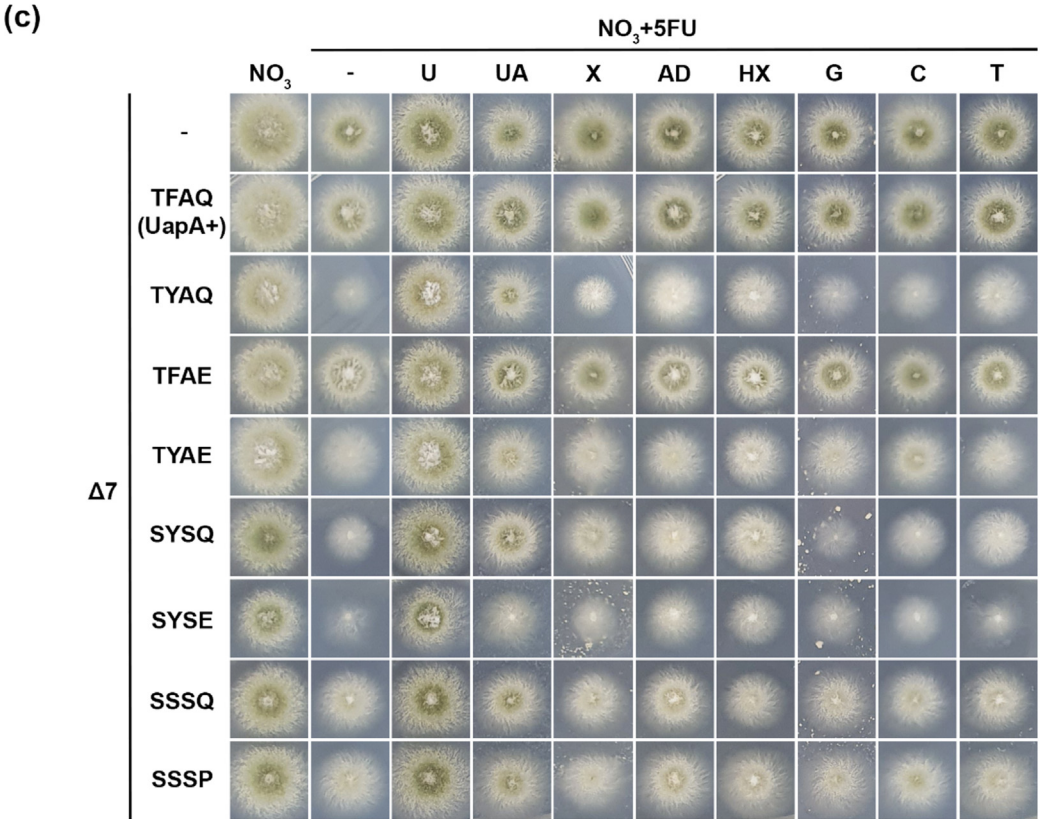


Figure 3. Transport activity and specificity of UapA NAT motif mutant versions (A) Comparative transport of 0.1 μM of radiolabelled ^{3}H-xanthine expressed as percentages of initial uptake rates (V) compared to the wild-type (UapA) rate taken as 100%, performed at 37 °C (see Materials and methods). Results shown are the average of triplicate measurements from 3 independent assays with a SD shown as error bars. (B) K_m (xanthine) or K_i (uric acid, adenine, hypoxanthine, uracil) values estimated via competitive inhibition assays of radiolabelled xanthine uptake (Krypotou and Diallinas[26]; see also Materials and methods) in selected mutants. NI stands for no evidence of radiolabeled xanthine transport inhibition in the simultaneous presence of excess 'competitor' nucleobase supplied to up to 1 mM. >1000 means detectable inhibition, but the K_i value could not be determined given that nucleobases are little soluble at concentrations ≥2 mM. Results shown are averages of triplicate measurements from 3 independent assays with SD values shown. (C) *In vivo* competition assays scoring the reversion of 5FU (100 μM) growth inhibition in the presence of excess purines or pyrimidines (2 mM) in the growth medium. Strains and growth conditions are as described in Figure 2. U is uracil, G is guanine, C is cytosine and T is thymine.

mutants UapA-F528M, -R481G and -**SYSE** are significantly less sensitive to 5FU than their relative double or triple mutants. Finally, while single mutation T526L confers hypersensitivity to 5FU, this is compromised in the context of R481G/**SYSE.** In conclusion, the functional effect

Table 1 Transport function and specificity of UapA NAT motif mutants

NAT motif version	UA	XA	AD	HX	5FU (U)	5FC (C)	5Fud (Ud)	AA	Comments on UapA function
TFAQ	+++	+++	–	–	–	–	–	–	Wt function/specificity
TYAQ	+++	+++	+	+	++	–	–	–	increased *binding/transport* promiscuity
TFAE	++	++	+/–	++	+/–	–	–	–	increased *binding/transport* promiscuity, reduced affinity and transport of UA, XA
TYAE	–	–	–	–	+/–	–	–	–	loss of purine *binding/transport*, but gain of moderate 5FU transport
SSSQ	+/–	–	–	–	+	–	–	–	loss/reduction of purine *binding/transport*, but gain of 5FU transport
SYSE	–	–	–	–	+++	–	–	–	loss/reduction of purine transport but gain of efficient 5FU transport
SYSQ	+++	+++	++	+/–	++	–	–	–	increased promiscuity for purines and 5FU
SSSP	–	–	–	–	+/–	–	–	–	loss of purine, some gain of very low 5FU transport

Underlined are amino acid replacements in mutant versions of UapA.
+++ signifies wild type-like substrate binding and transport activity. – stands for ≤5% transport capacity of wild-type UapA and no indication of binding of nucleobases (up to 1 mM) or toxic nucleobase analogues (50–100 μM).
In between values, +/–, +, ++ reflect compromised binding/transport (10–50% of wild-type UapA).
UA is uric acid, XA is xanthine, AD is adenine, HX is hypoxanthine, , 5FU is 5-flurouracil, U is uracil, 5FC is 5-flurocytosine, C is cytosine, 5FUd is 5-fluorouridine, Ud is uridine, AA is L-ascorbic acid (1%).
This table is based on functional studies (growth tests, subcellular localization uptake measurements and transport competition assays) of different mutants expressed in a Δ7 genetic background (see Figs. 2 and 3).
None of the above mutant confers AA, 5FC, 5FUd or nucleoside or allantoin transport

of the **SYSE** sequence in the context of UapA proved highly dependent on the presence of mutations at residues on R481, T526 or F528. Notably, none of the above combinations of mutations conferred to UapA an ability to transport nucleosides or allantoin. Additionally, we did not detect any change in *A. nidulans* growth of strains containing the aforementioned combinations of mutations in media containing 0.1–1% (w/v) of L-ascorbic acid (not shown).

To better rationalize the above observations, we performed radiolabeled xanthine transport assays in selected mutants (Figure 4(B)). The presence of the **SYSE** sequence dramatically reduced UapA-mediated xanthine transport, despite the presence of specificity mutations, which when present in an otherwise wild-type NAT context (i.e. TFAQ instead of **SYSE**) not only enlarged specificity, but also lead to increased xanthine transport rates. Given that R481G, T526L, F528M, **SYSE** mutations, alone or in specific combinations, conferred increased capacity for 5FU accumulation, we tested whether the corresponding strains also have a capacity for UapA-mediated transport of uracil, which is very similar in structure with its toxic analogue 5FU. Results in Figure 4(C) showed that the **SYSE** sequence does not confer to UapA measurable transport activity for uracil, and also seems to suppress detectable uracil transport in UapA-T526L, -R481G/T526L or -R481G/F528M mutants, which possess a wild-type NAT signature motif. Taken together, results shown in Figure 4 (A–C) suggest that introducing the **SYSE** sequence in the NAT signature motif of UapA leads to reduced activity for purine or uracil transport. This negative effect of **SYSE** seems only partially suppressed in the context of F528M. On the other hand, versions of UapA carrying **SYSE** alone or in combination with other mutations lead to accumulation of 5FU, thus revealing that the transporter is transport-active for this toxic pyrimidine analogue.

The most probable explanation for the observed apparent inability of UapA-**SYSE** to transport uracil, despite its capacity to lead to significant 5FU accumulation, is that this UapA version has high-affinity specifically for 5FU, but low affinity for uracil and nucleobases. Similar situations, where transporters have increased affinities for chemical analogues of physiological substrates are known to exist, and seem to be due to additional interactions of substituted positions in the analogues.[35,36] To test this possibility, and given that radiolabelled 5FU is not available, we performed *in vivo* competition assays in growth tests where we follow the level of 5FU reversion of toxicity by increasing concentrations of uracil. For this, we first established the concentration range that 5FU toxicity is best visualized. We performed these tests not only for UapA-**SYSE**, but also for several of the specificity mutants used in this work. The left panel of Figure 5(A) shows that most mutants (T526L, R481G/T526L, R481G/F528M, **SYSE**, T526L/**SYSE**, F528M/**SYSE**, R481G/F528M/**SYSE**) were very sensitive to 5FU at concentrations as low as 10 μM. Some mutants (F528M, R481G, R481G/**SYSE** or R481G/T526L/**SYSE**) were more mildly sensitive to 5FU, showing significantly reduced growth only at 100 μM of the toxic analogue. In gen-

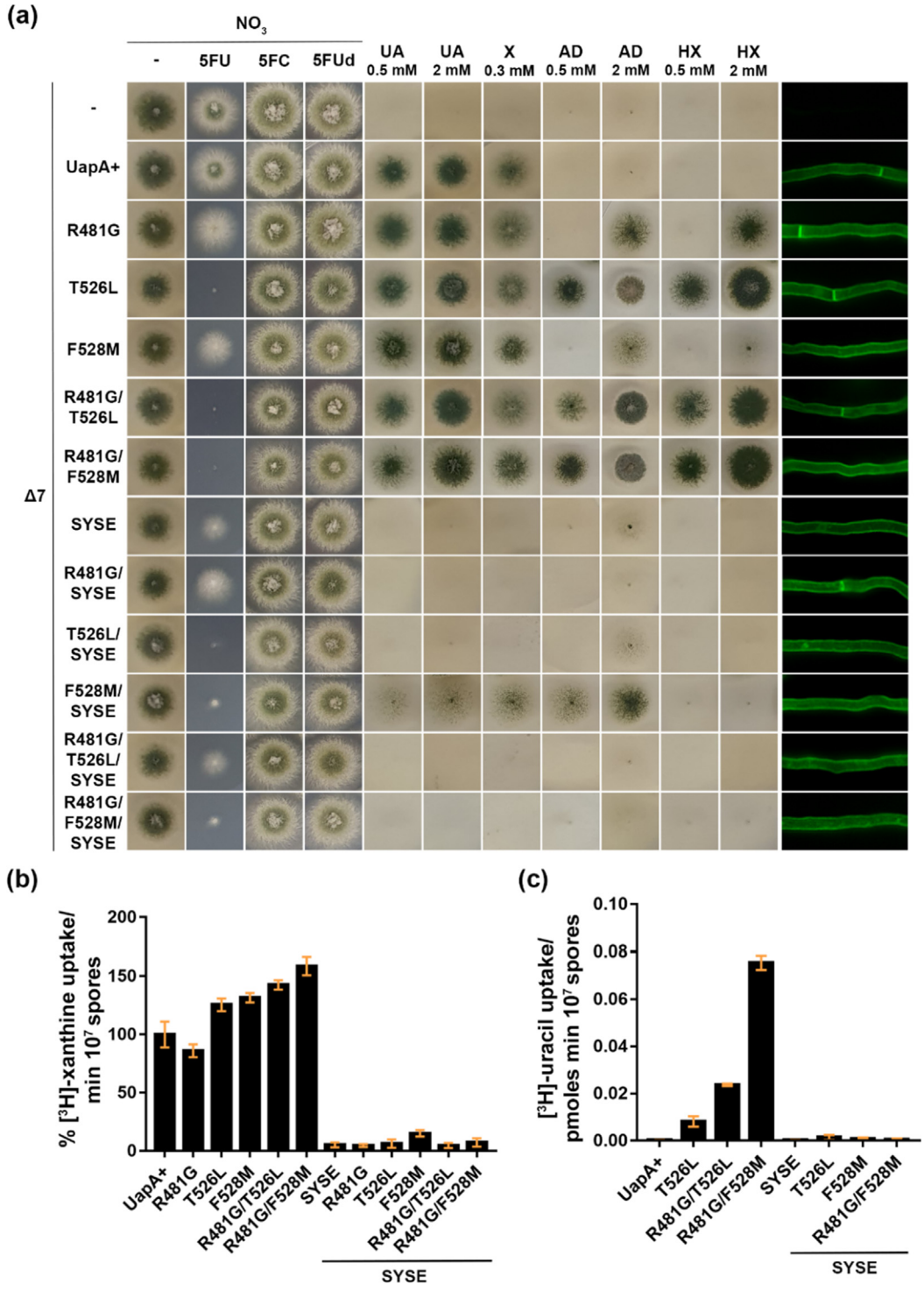
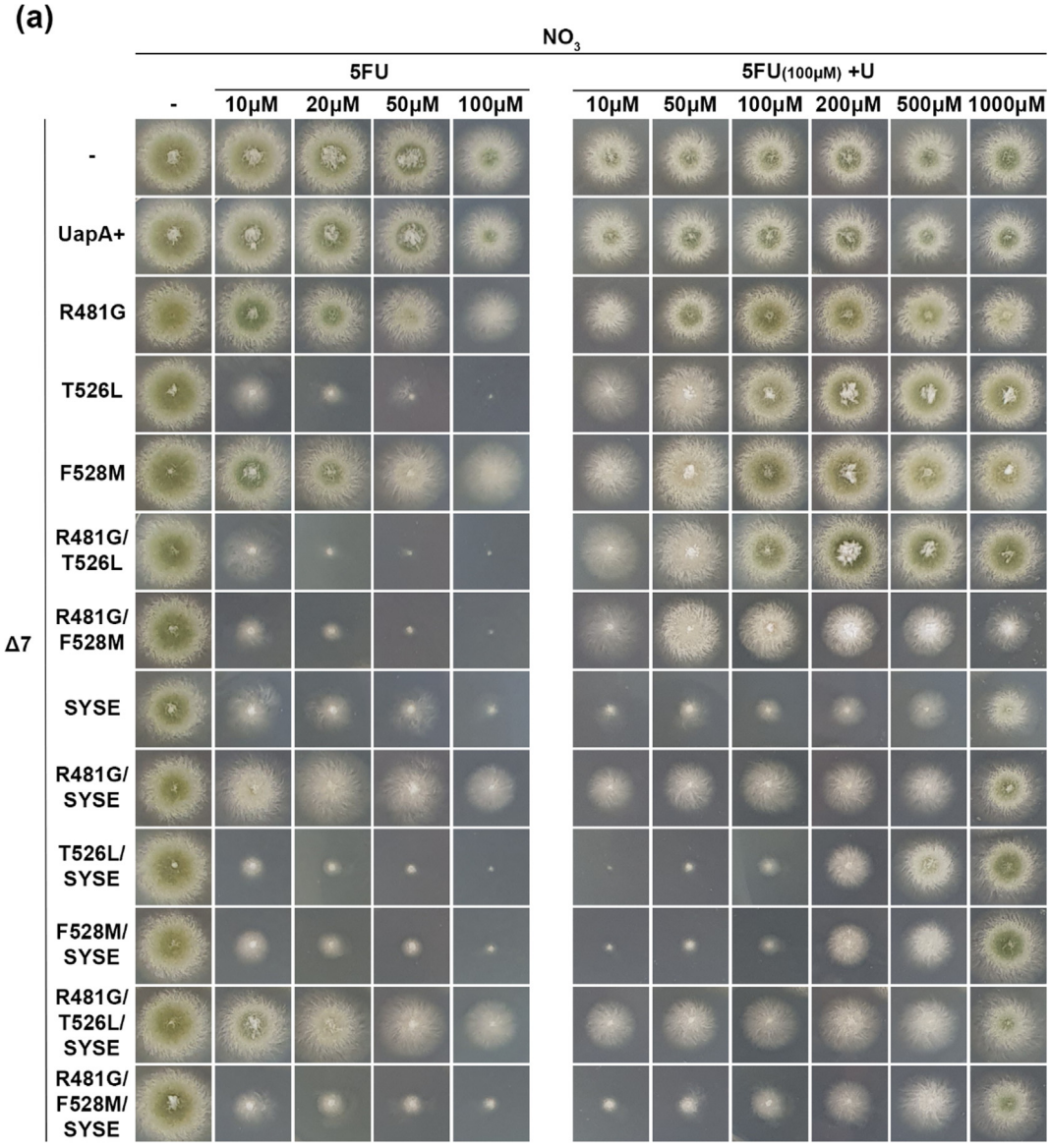


Figure 4. Combination of the **SYSE** NAT motif sequence with specificity mutations leads to functionally distinct UapA versions (A) Growth tests of strains expressing UapA versions carrying the **SYSE** sequence in the NAT motif combined with selected specificity mutations (F528M, T526L, R481G, R481G/F528M or R481G/T526L). Growth conditions and control strains are as described in Figure 2. Inverted fluorescence microscopy images confirming the proper localization of all UapA mutant versions to the PM is shown on the last right column. (B) Comparative ^{3}H-xanthine (0.1μM) transport rates in strains expressing UapA versions combining **SYSE** with selected specificity mutations (F528M, T526L, R481G, R481G/F528M or R481G/T526L). Details are as in Figure 3A. Standard deviation is depicted with error bars. (C) Comparative ^{3}H-uracil (0,1 μM) transport rates in strains expressing UapA mutant versions that conferred increased capacity for 5FU accumulation. Details are as in Figure 3A. Standard deviation is depicted with error bars.

(a)



(b)

Binding affinity constants			
	$K_{m/i}$ (μM)		
Substrate/ligand	UapA$^+$	F528M	F528M/SYSE
Xanthine	4	4	2
Uric acid	6	24	15
Adenine	NI	NI	275
Hypoxanthine	NI	NI	NI
Uracil	NI	NI	102
5-FU	NI	NI	271

Figure 5. Relative recognition of 5FU *versus* uracil as established by growth tests in UapA mutants. (A) Left panel: Concentration-dependent 5FU toxicity in the mutants. Right panel: *In vivo* competition of 100 μM 5FU toxicity by increasing concentration of uracil (range 10–1000 μM) in the mutants. Growth conditions are as described in Figure 2. (B) K_m (xanthine) or K_i (uric acid, adenine, hypoxanthine, uracil, 5FU) values estimated via competitive inhibition assays of radiolabelled xanthine uptake ([26]; see also Materials and methods) in selected mutants. Details are as described in Figure 3B. Results shown are averages of triplicate measurements from 3 independent assays with a SD of < 20%.

eral, mutants possessing the **SYSE** sequence are more sensitive to 5FU, compared to wild-type UapA. We thus used 100 μM 5FU to perform *in vivo* competition assays by uracil. The right panel of Figure 5(A) shows that uracil competed 5FU toxicity, in all cases, but the level of reversion of toxicity was different in the mutants analysed. In mutants that are either highly (T526L, R481G/T526L or R481G/F528M) or moderately (R481G or F528M) sensitive to 100 μM 5FU due to mutations outside the NAT motif, uracil suppressed toxicity at a concentration range of 50–100 μM, which suggested that 5FU and uracil are recognized with comparable affinities. However, in most mutants having the **SYSE** sequence only 10-fold excess uracil (1000 μM) could significantly suppress the toxicity of 100 μM of 5FU. These results suggested that the presence of the SYSE sequence might lead to either higher affinity for 5FU or lower recognition of uracil, or both.

To address this issue directly, we estimated the K_i values for purines, uracil and 5FU in F528M/**SYSE**, which was the only combination of **SYSE** with other mutations that possessed measurable transport of radiolabeled xanthine, thus enabling relative competition assays. As a control for the effect **SYSE** on substrate recognition we also estimated the K_i values of the single mutant F528M and wild-type UapA. Figure 5(B) shows that UapA-F528M/**SYSE** conserves high affinity (low μM range) for physiological substrates (xanthine and uric acid), similar to wild-type UapA, but additionally recognizes adenine, uracil and 5FU with moderate affinities (275, 102 271 μM, respectively). Notice that F528M alone can lead to very low apparent accumulation of adenine or hypoxanthine (see Figure 4(A)), but in this case transport should operate via very low affinity, as suggested by the lack of competition of radiolabeled xanthine by excess of these purines (see Figure 5(B)). Thus, the presence of **SYSE**, at least in the context of F528M, contributes to better binding of several non-physiological substrates (adenine, uracil, 5FU), but does not make a significant distinction between uracil and 5FU. Thus, the simplest explanation for the observation that 10-fold excess uracil did not fully suppress 5FU toxicity in mutants possessing **SYSE** (seen in Figure 5A, right panel), is that moderate affinity binding and low transport of 5FU is still sufficient to cause cytotoxicity.

We also combined the quadruple **SYSE** substitution (residues 405–408) with mutations N410I and N410V. The rationale for this was the following. We observed and verified via correlated sequence conservation algorithms (BIS2Analyzer, see Materials and methods) that in NAT proteins the residue analogous to position 410 of UapA is conserved as an Asn *only* when the residue at position 408 is occupied by a Gln (see Figure 1 (A)). In these cases, the NAT protein functions as a transporter specific for xanthine or uric acid and xanthine. In cases where the position equivalent to 410 in UapA is occupied by variable aliphatic amino acids, while positions 408 is a Glu, the proteins show reduced specificity towards uric acid/xanthine, being able to transport most purines and uracil. In other words, position 410 might have a 'cryptic' role in determining the specificity and in particular in contributing to acquisition of increased affinity for uric acid/xanthine binding in a context-dependent manner. The functional importance of N410 in UapA has not been tested before, while mutational studies in the bacterial XanQ homologue have suggested that residue N326 (equivalent to N410 in UapA) is not essential for xanthine transport.[24,25,37] Here we show that the single N410V mutation led to loss of apparent UapA transport activity of purines, while N410I had no effect on UapA activity and specificity, as based on growth assays with uric acid, adenine and hypoxanthine as sole N sources. However, either mutation, when combined with the **SYSE** sequence, led to loss of any apparent transport of nucleobases, including a significant reduction in 5FU toxic accumulation (Supplementary Figure S1). This shows that the presence of aliphatic residues at 410 is not compatible with SYSE in the UapA sequence context, contrasting what is observed in other NATs which naturally possess a SYSE sequence. Essentially "mirror-image" replacements at the position equivalent to 410 have been studied in bacterial xanthine/uric acid-transporting homologues that have a Q408-V410 combination. In those cases, replacement of the aliphatic residue Val with Asn leads to functional inactivation: the single-replacement mutants V312N in UacT[37] and V320N in SmXUacT[38] are present at wild-type levels in the membrane, but are functionally inactive. Thus, re-establishment of the combination Q408-N410 in the sequence context of those bacterial UapA-homologs is not compatible with (xanthine or uric acid) transport activity. This further highlights the complexity of epistatic interactions that control the fine functioning of the substrate binding site in NATs.

Suppressor mutations conferring uric acid transport in UapA-SYSE support the functional importance of F528 and reveal new residues critical for specificity

To further understand the role of the NAT signature motif in specificity, we took advantage of the fact that UapA-**SYSE** has no capacity to transport purines to isolate revertants, after U.V. mutagenesis, that confer growth on uric acid as sole N source. Table 2 and Figure 6(A) summarize the molecular identity and position of 33 revertants, which all proved to be within the *uapA* orf. Revertants concerned 7 amino acid changes in four residues: L234, E408, A519 and F528. All mutations, except F528I, were isolated

Table 2 Profile of SYSE suppressors

Mutation	TMS*	Domain	Codon change	Number of isolates
L234M	5	Dimerization	CTG → ATG	2
E408Q	10	Core	GAG → CAG	5
A519P	14a	Dimerization	GCG → CCG	3
F528S	14b	Dimerization	TTT → TCT	10
F528S	14b	Dimerization	TTT → TCA	1
F528C	14b	Dimerization	TTT → TGT	6
F528V	14b	Dimerization	TTT → GTT	5
F528I	14b	Dimerization	TTT → ATT	1

* Based on.[12]

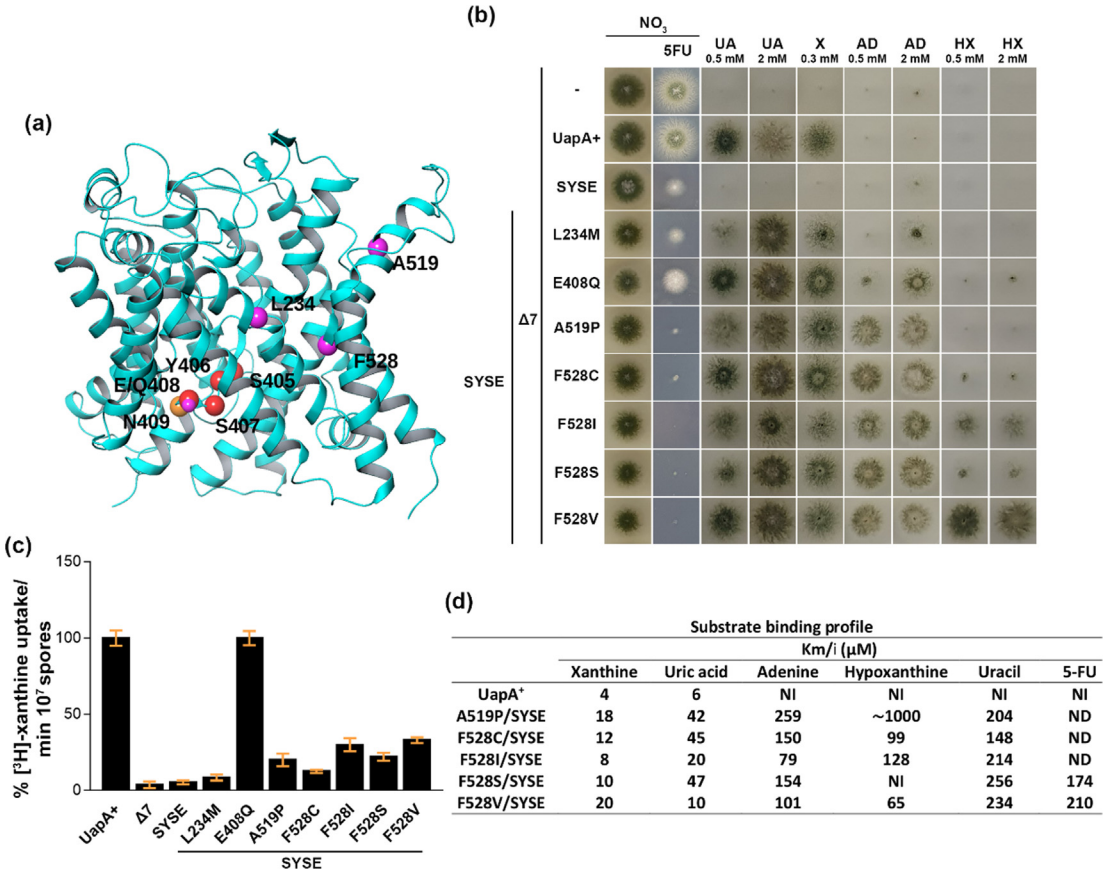


Figure 6. Characterization of mutations allowing uric acid transport in the UapA-**SYSE** context (A) Topological model showing the position of suppressor mutations relative to the NAT motif in UapA (B) Growth tests of suppressors. Growth conditions and control strains are as described in Figure 2. (C) Comparative ^{3}H-xanthine (0.1μM) transport rates in suppressors. Details are as described in Figure 3A. Standard deviation is depicted with error bars. (D) $K_{i/m}$ values (μM) determined using ^{3}H-xanthine uptake competition in suppressors. Details are as described in Figure 3B. ND indicates not determined. Results shown are the averages of triplicate measurements from 3 independent assays with a SD of <20%.

several times (2–10 times), showing that mutagenesis was rather saturated. Mutations in E408 introduced a Gln residue, thus reverting the **SYSE** sequence to **SYS**Q, which is a mutation expected to be isolated in this screen given the fact that UapA-**SYS**Q was earlier shown to confer growth on uric acid. Mutations in F528 were: F528C, F528I, F528S and F528V.

F528 has been previously shown to act as major specificity element as mutations introducing aliphatic (Ala, Met) or polar amino acids (Ser, Thr, Gln or Asn) converted UapA into a promiscuous purine-uracil transporter (Papageorgiou *et al.*[17] and Amillis *et al.*[39] but also herein). Notably, while F528 specificity mutations conferred very low affinity transport (>1 mM) for non-physiological purine

substrates, when combined with Q408E increased significantly the affinity of UapA for these 'new' substrates. Based on the crystal structure of UapA and Molecular Dynamics[12,20] (see also later) F528 is a residue located at the outward-facing gate of the substrate translocation trajectory, in the middle of TMS14 of UapA. Mutation A519P concerned also a residue in TMS14, but one that has never appeared in previous genetic screens for specificity mutants. Introduction of a Pro residue at this position might well modify the local structure to TMS14, and in particular the positioning of T526 and F528, both residues being critical for specificity.[17] Interestingly, the mutation concerning residue L234 (L234M in TMS5) has been isolated before in a screen searching for revertants of UapA mutations that led to loss-of-function because of inefficient dimerization.[40] It has then been proposed that L234M stabilizes the structure of monomeric UapA units, and thus lead to increased steady state levels of UapA. How such a mutation can also affect specificity opens new issues on the role of dimerization in finely regulating specificity, as already discussed in [1,12].

We analysed further the specificity profile of the revertants isolated by growth tests and uptake assays, as described earlier. Figure 6(B) shows that all revertants grow on uric acid (0.5 mM), as expected given their selection was carried out on this purine, but showed different levels of growth. E408Q, F528C and F528V were similar to the control strain expressing wild-type UapA. F528S, F528I, A519P and L234M showed relatively reduced growth, which however improved to nearly wild-type levels when uric acid was supplied at 4-fold higher concentration (2 mM). This suggests that the latter set of mutants might exhibit reduced uric acid binding affinity (see later). Most revertants grew equally well on xanthine, except from L234M, which seemed to be less efficient in xanthine accumulation. Most mutants also grew well on adenine, except from E408Q and L234M, which showed little growth on 0.5 mM adenine. Increased supply of adenine restored significant growth of E408Q, but only partially that of L234M. Most mutants, except F528V showed very little or moderate capacity for growth on hypoxanthine, both at standard and high concentrations. In fact, L234M and A519P mutants showed no evidence for hypoxanthine growth at all, scoring similarly to the original **SYSE** mutant. Finally, some revertants increased the apparent capacity for 5FU accumulation (F528S, F528V, F528C, F528I and A519P) or did not seem to affect significantly (L234M) the existing level of sensitivity of the original UapA-**SYSE** strain, with only E408Q leading to decreased capacity for 5FU accumulation.

We performed transport assays with radiolabelled xanthine in all revertants isolated. Figure 6(C) shows that E408Q has wild-type transport rates justifying its full growth on uric acid or xanthine (Figure 6B). Most other mutants analyzed, except L234M, grew relatively well on uric acid or xanthine, despite reduced apparent transport rates of xanthine (~15–28% of wild-type levels), suggesting that this level of transport is apparently sufficient for growth (Notice also that in growth tests the concentration of substrates is 1000-fold higher that the concentration of radiolabeled substrates in transport assays; for details see [26]). Lastly, L234M that showed the lowest transport rate (8% of wild-type), also showed the lowest capacity to grow on xanthine and uric acid. We also estimated the K_i values for purines, uracil and 5FU in suppressors (Figure 6(D)). In general, all suppressors conserved relatively high or moderate affinity for xanthine or uric acid (8–47 M), albeit reduced compared to wild-type UapA. The moderate reduction in affinity for the physiological substrates is probably due, at least in the case of suppressors concerning F528 substitutions, to the presence of the **SYSE** sequence. This is concluded due to previously analyzed single F528 mutations that do not affect significantly the affinity of UapA for xanthine and uric acid,[17,41] but also when considering reduced growth on xanthine or uric acid of relative mutants including **SYSE**, as shown in Figure 4(A). Most suppressors also showed moderate affinities for other purines, uracil and 5FU (in the range of 65–259 μM), with few exceptions of certain mutants towards specific substrates. Given that single mutations in F528 do not modify significantly the affinity for substrates,[17,41] the **SYSE** sequence should, in principle, be the reason of acquisition of moderate/low affinity binding of purines, uracil and 5FU. Thus, while the **SYSE** sequence when it stands alone in an otherwise wild-type UapA context leads to a transporter that is only able to accumulate low but toxic amounts of 5FU, when combined with suppressor mutations removing the aromatic Phe residue at position 528, it may lead to generation of UapA versions recognizing and transporting all purines, uracil and 5FU.

A similar, but not identical, promiscuous profile has been previously reported when specific mutations in F528, T526 or R481 are combined with mutation Q408E (i.e. sequence TFAE in the NAT motif; see[17,41]). Thus, the critical substitution that converts UapA from being highly specific for xanthine/uric acid to become a more general nucleobase transporter is mutation Q408E, a replacement topologically located in the substrate binding site and specifically concerning a residue known to interact strongly with substrates.[12,13] However, the potential effect of substitution Q408E to specificity is revealed only in the simultaneous presence of specific substitutions of other residues, distant from the *bona fidae* binding site, in this case amino acids R481, T526 or F528.

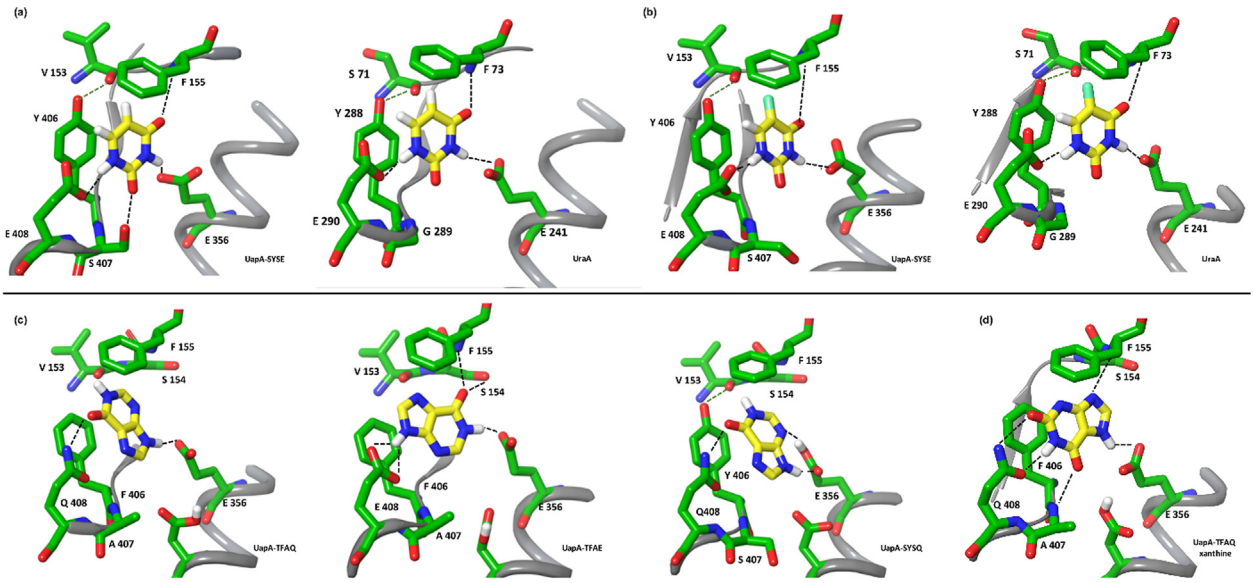


Figure 7. UapA mutants-substrate interactions as resulted by Molecular Simulations. (A) UapA-**SYSE**-uracil and UraA-uracil interactions. The UapA-**SYSE**-uracil interaction is stabilized by the formation of four H-bonds (depicted as dashed lines) between N1H and N3H with E408 and E356 carboxylates, respectively, C2=O with S407 hydroxyl group and C4=O with F155 backbone, as well as, a p-p stacking formed between uracil with both F155 and Y406. Notice also that Y406 hydroxyl group forms a H-bond with V153 backbone. The UraA-uracil crystallographic structure is also presented on the right for comparison (23). (B) UapA- **SYSE**-5FU and UraA-5FU interactions. In both structures 5FU is oriented similarly to uracil with the F atom located in a hydrophobic cleft and C-F bond parallel to both F155 and Y406 rings. (C) Hypoxanthine interactions with UapA, UapA-TFA E and UapA- SYSQ. Hypoxanthine forms two H-bonds with the UapA binding cavity, one between N9H and E356 carboxylate and the second between C6=O and Q408 amide group. In UapA-TFA E, hypoxanthine N9H and N1H both form H-bonds with E408, and E356 carboxylates, respectively, while C6=O forms H-bond with S154 hydroxyl group and F155 backbone. In UapA-SYSQ-hypoxanthine interaction E356 is protonated. Hypoxanthine H-bond interactions with the binding cavity are between N9H, N3 and E356, as well as, between C6=O and Q408. (D) Wild-type UapA-xanthine crystal structure is included for comparison.

Rationalizing specificity modifications via alteration in NAT signature motif

To better understand the specificity alterations of UapA, molecular simulations of protein-substrate interactions were performed (Figure 7). To this end, the models of UapA-**SYSE**, UapA-**SYS**Q, and UapA-TFA**E** have been constructed on the basis of the structure of UapA crystallized in the inward-open conformation. Important differences were observed in the modelled structures. These are related to the binding cavity of the transporter due to the presence of the F406Y mutation but also to the calculated pKa of acidic residues critical for substrate binding (e.g. the invariable E356 and D360 in TMS8 and E408 in TMS10 of UapA-**SYSE** or UapA-TFA**E**). More specifically, in UapA-**SYSE** the calculated pKa of E356 rises from 7.5 in the wild-type (UapA-TFAQ) to 10.0, while that of D360 drops from 9.4 to 7.4. Similar pKa values to UapA-**SYSE** were calculated in UapA-**SYS**Q. Thus, in both UapA-**SYSE** and UapA-**SYS**Q E356 is considered protonated, while D360 is thought to be negatively charged (given transport assays are performed at pH 6.8). On the other hand, E408 in both UapA-**SYSE** and UapA-TFA**E** exhibits a theoretical pKa of 6.8. Alterations in the acidic character in side chains of UapA-**SYS**Q and UapA-**SYSE** as compared to wild-type (UapA-TFAQ) or UapA-TFA**E** are expected to modify the polar nature of the substrate binding side, and thus might affect substrate binding and transport. The other important observation related to the binding cavity of these mutants versus the wild-type UapA concerns the formation of a hydrogen bond when F406 is mutated to Tyr between the phenolic hydroxyl and the V153 backbone carbonyl (see Figure 7(A), (B)). This hydrogen bond stabilizes the main part of the binding cavity between TMS3 and TMS10, which is known to form the characteristic beta-sheet structure in the middle of the transmembrane domain where the substrate is embedded. This stabilization does not seem to affect transport activity (see Figure 3(A)), but apparently contributes to enlarged specificity as UapA-T**Y**AQ recognizes 5FU, hypoxanthine and adenine

significantly better than UapA-TFAQ (see Figs. 2 and 3(B), (C))

In a second step, we used Induced Fit Docking (IFD) calculations to try to understand the nature of increased apparent affinity of UapA-**SYSE** for uracil and 5FU. The most energetically favourable binding structure of uracil exhibits a remarkable similarity to the crystal structure of uracil bound by its natural transporter UraA,[23] as presented in the left and right panels of Figure 7(A). In UapA-**SYSE**, uracil is well stabilized between the two acidic side chains E408 and E356 forming also p-p stacking interactions with F155 and Y406 plus one more H-bond with S407 side chain. The latter is not present in UraA having a glycine at this position while UapA has an alanine. A quite similar positioning resulted by the IFD calculations also for 5FU, in both UapA-**SYSE** and UraA (Figure 7(B), left and right panels, respectively). Notice that the F atom is located in a hydrophobic cleft of the binding pocket where the C-F bond is parallel to both F155 and Y406 rings, without affecting proper orientation of the uracil ring. Importantly, in wild-type UapA, which does not seem to transport uracil or 5FU, these ligands are not similarly positioned in the binding pocket, when compared to the docking position in UraA or UapA-**SYSE**, as in their best fit position seem to interact via a mirror-image orientation (Supplementary Figure S2). Thus, it becomes evident that uracil/5FU docking in UraA or UapA-**SYSE**, both of which recognize and can transport these ligands, is distinct from that in wild type UapA, which cannot transport these pyrimidines. Our docking analysis cannot however rationalize why, despite similar docking positions, UraA recognizes uracil and 5FU with very high affinity (sub micromolar concentration), whereas UapA-**SYSE** recognizes these two substrates with moderate affinities (148–259 μM, depending on the combination of the **SYSE** sequence with different F528 substitutions). A probable explanation for this difference, other than considering that binding affinities cannot be well predicted by modeling approaches, is that residues outside the binding site, crucial for recognition, might be distinct in UraA and UapA-**SYSE**.

In an effort to further rationalize the specificity of UapA mutants against hypoxanthine IFD calculations were performed considering wild type (UapA-TFAQ) and mutant versions UapA-**SYS**Q and UapA-TFA**E**. Notice again here that functional assays show that wild type (UapA-TFAQ) does not bind hypoxanthine, whereas UapA-TFA**E** binds with relatively high affinity (60–70 μM) but does not transport hypoxanthine, and UapA-**SYS**Q recognizes hypoxanthine with low affinity (~1 mM). The lowest energy binding modes of hypoxanthine of all three forms are presented in Figure 7(C). It can be easily noticed that none of them is similar to the xanthine-UapA binding mode where Q408 forms two H-bonds with N1H and C2=O (Figure 7(D)). In the wild type UapA-hypoxanthine structure N9H, but not N1H, interacts with E356, while Q408 forms a H-bond with C6=O forcing the hypoxanthine scaffold in a completely different orientation compared to the orientation of physiological substrates (e.g. xanthine). In UapA-TFA**E**, hypoxanthine appears to be oriented in such a way that the two positive charges on N9H and N1H[42] interact with the two negative charges of E408 and E356, respectively, while the C6=O group forms a H-bond with S154. This could explain the relatively high binding affinity of hypoxanthine in this UapA mutant. However, probably due to its topological disorientation within the binding site, this ligand is not transported. Finally, in UapA-**SYS**Q, where E356 is protonated, as deduced from pKa calculations, can now interact with N9H and N3 while C6=O interacts with Q408 amide group. This might explain the binding of hypoxanthine by UapA-**SYS**Q, even with low affinity, as compared to wild-type UapA that makes fewer interactions.

F528 dynamically interacts with the substrate binding site

The structural explanation of the functional role of all different UapA residues studied herein is beyond the scope of this article and under investigation in our group. However, in order to rationalize the importance of F528 mutations isolated several times in this work, but also previously, as a key element affecting specificity alone and in combination with NAT signature motif substitutions, the model of the outward conformation of UapA has been constructed by homology modelling to the anion exchanger domain of human erythrocyte Band3 crystal structure.[43] Band3 is a human structural homologue of UapA that functions as bicarbonate ($HCO3^-$) transporter, which has been crystalized in the outward conformation and represents an excellent model, if not unique, for the outward conformation of transporters following the elevator mechanism. Figure 8 shows the position of F528 relatively to the position of E356 and Q408 in the modelled outward conformation and the inward crystal structure of UapA. As can be seen, in the outward topology the two residues are in very close distance interacting through p-electrons. The MD calculations shows that the distance of the two side chains remains stable ~3.50 Å through the 100 ns simulation, suggesting a quite stable interaction. In contrast, in the inward crystal structure this distance becomes ~6.00 Å. This suggests a dynamic interaction of F528 with the substrate binding site that could explain its critical role in specificity (see Discussion).

Discussion

Despite the fact that UapA and bacterial NATs have been extensively analysed by a plethora of

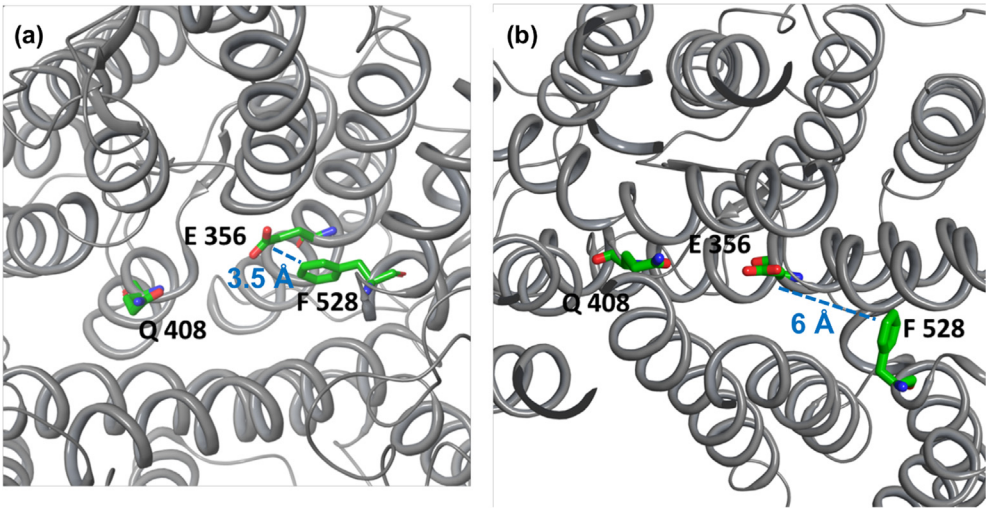


Figure 8. UapA outward conformation. (A) The model constructed using Band3 as template is shown mainly focusing on the interaction between F528 and E356. The E356 carboxylate group remains almost parallel to the F528 phenyl ring (3.5 Å) during the 100 ns MD simulation. (B) The corresponding distance in the inward conformation is much higher (pdb structure 5i6c).

mutational studies we still do not know how specificity is determined by residues outside the binding site. The recent determination of the crystal structures of UapA and UraA, corresponding to inward-facing and occluded topologies, respectively, has led to formal identification of a major substrate binding site, but provided little hints on the issue of specificity. MDs using UapA, UraA and a handful of NAT structural homologues (e.g. anion exchanger 1 AE1/Band3 protein and SLC26Dg) caught in distinct conformations (e.g. outward-open) led to proposing a *sliding-elevator* type of transport mechanism[44] and defined a putative substrate translocation pathway in UapA.[12] In addition, based on MDs, the most prominent specificity mutations concerned residues R481, T526 or F528, which are located along the proposed sliding trajectory of the core domain in the UapA dimer. Still however, how these mutations affect specificity is poorly understood.

Here we decided to re-address how UapA specificity is determined by trying to better understand the contribution of partially conserved residues of the NAT signature motif, which is known to be part of the substrate binding site located in sliding elevator core domain in all NATs. Thus, we performed directed mutational and functional analyses of residues located at positions 405–408 of the NAT signature motif of UapA (i.e. TFAQ). The rationale of the mutations designed was to introduce residues that are present in other NATs showing distinct specificities. An important novelty of the present study was the use of a strain (Δ7) that genetically lacks not only a wild-type UapA endogenous copy, but also all other major transporters related to purine, pyrimidine, nucleoside or allantoin transport. The Δ7 strain has practically no uptake transport activity for purine-related solutes and thus any UapA mutant introduced by genetic transformation in this strain could be analyzed for its ability to transport any purine-related substrate in a 'clean' background.

Based on the above system we revealed context-dependent cryptic roles of specific partially conserved residues located in the first part of the NAT signature motif. Most impressively, introducing the **SYSE** sequence in UapA leads to a transporter that lost its capacity to transport its physiological substrates or any other purine, but gained a capacity for toxic 5FU accumulation. In addition, combination of **SYSE** with F528 mutations led to a transporter able to translocate uracil or 5FU, in addition to purines. These findings are in line with the observation that functionally characterized bacterial NATs known to transport uracil and 5FU (e.g. UraA or RutG)[22,45] conserve the Tyr and Glu residues (e.g. possess the sequence T**Y**G/A**E**), while the plant and metazoan NATs that recognize purines *and* uracil,[6,10,11] also conserve the Glu residue (e.g. XXX**E**). In other words, the presence of a Glu residue seems to dramatically increase uracil transport capability, while the Tyr residue also contributes to that. Our relative MDs have showed why this might be so. In UapA-**SYSE** uracil is well stabilized between E408 and E356 forming also pi-pi stacking interactions with F155 and Y406 and one more H-bond with S407 side chain. This docking position of uracil is not seen with the wild-type UapA. The important context-dependent role of E408 in UapA specificity is corroborated by comparing UapA-**SYSE** and UapA-**SYS**Q, the former being rather specific for

5FU accumulation and the latter promiscuous (transport of purines-uracil-5FU). Relative to these observations, our MD analysis also revealed possible interactions that rationalize why UapA-TFA**E** binds, but does not transport hypoxanthine, or why UapA-**SYSE** with UapA-**SYS**Q, have very distinct transport activities and specificities compared to wild-type UapA.

Among the UapA mutations functionally analyzed here two were specifically designed to introduce amino acids found in the NAT signature motif of mammalian L-ascorbate transporters, namely **SSSP** and **SSS**Q (see Figure 1(A)). The first one introduces the full set of 4 residues found in L-ascorbate transporters (SVCT1/2), and the second conserves the essential for purine binding Q408. Thus, we tested whether UapA could be converted to a transporter capable of L-ascorbate transport, or at least improve its capacity for binding L-ascorbate. We have recently shown that wild-type UapA might recognize with extremely low affinity (Ki ~ 20 mM) L-ascorbic acid,[9] but we were unable to convincingly show any relative transport activity. Mutant UapA-**SSSP**, as probably expected due to replacement of residues critical for purine recognition (F406 and Q408), lost any detectable transport activity for xanthine and uric acid or purines in general. UapA-**SSS**Q also showed little, if any, purine transport activity, despite conserving Q408. This strongly suggests that the introduction of a series of polar Ser residues upstream of Q408 is detrimental for purine transport, at least within the context of an otherwise wild-type UapA sequence. Surprisingly, unlike wild-type UapA, both UapA-**SSSP** and UapA-**SSS**Q seem to be capable of accumulating low subtoxic levels of 5FU, judged from the reduction in conidiospore production when compared to wild-type UapA (see Figure 2(A)). However, we failed to obtain any evidence that UapA-**SSSP** and UapA-**SSS**Q recognize better or transport L-ascorbate. Given however the limitations of our system for detecting low-affinity L-ascorbate transport, we cannot exclude the possibility that UapA-**SSS**P and UapA-**SSS**Q might have an ameliorated capacity for recognizing this metabolite. It seems that engineering L-ascorbate transport capacity in UapA may not be feasible by simply introducing residues that mimic the presumed binding site of L-ascorbate transporter homologues (e.g. SVCT1/2).

The present work also provides novel findings concerning how F528, a residue outside the substrate binding site, might function as a key amino acid in determining UapA specificity. The great majority of suppressor mutations that restored purine (e.g. uric acid) transport in UapA-**SYSE** replace F528 with aliphatic or polar residues. These suppressors in general enlarge UapA specificity. In previous studies, similar F528 replacements also enlarged UapA specificity, so that it can transport, albeit with low affinity, all purines and uracil. The only known substitution in this residue that conserves a wild-type functional profile is F528Y,[41] strongly suggesting that an aromatic ring at residue 528 is responsible for conferring specificity for uric acid or xanthine. How removing the aromaticity of residue 528 allows UapA to transport, in addition to its physiological substrates, all purines and uracil is not clear. However, our present MD analysis tried to shed some light into this question. By constructing the outward topology of UapA we were able to see that F528 approaches very close and probably interacts via pi-pi stacking with E356, which is a major substrate-binding residue. This interaction is apparently dynamic, as it is not present in the inward-facing UapA crystal. One might thus hypothesize that in the apo state (i.e. no substrate bound) the F528-E356 interaction 'locks' UapA in its outward topology. Upon substrate (and probably H^+) loading and stabilization between E356 and Q408, the F528-E356 interaction is lost, and UapA is allowed to shift to its inward topology via sliding of the elevator domain. Replacing F528 by small aliphatic amino acids seems to genetically unlock the sliding mechanism in a way that becomes less dependent on substrate/H^+ binding, and thus convert UapA into a more promiscuous transporter for low-affinity substrates. Given that mutations other than those concerning F528, as shown previously,[13] but also here (e.g. L234 and A519), can similarly convert UapA into a more promiscuous transporter for weakly-binding substrates, these might also modify ('loosen') the sliding mechanism. A more rigorous understanding of how specificity is determined in UapA and other NATs, but also in elevator-type transporters in general,[46] will require more structural data and relative MDs involving several distinct topologies.

Materials and Methods

Media, strains and growth conditions

Standard complete (CM) and minimal media (MM) for *A. nidulans* growth were used. Media and supplemented auxotrophies were used at the concentrations given in http://www.fgsc.net.[47] Glucose 1 % (w/v) was used as carbon source. 10 mM sodium nitrate (NO_3) was used as a standard nitrogen source. Nucleobases, nucleosides and toxic analogues were used at the following final concentrations: 5FU, 5FC or 5FUd at 10–100 μM, uric acid, adenine, hypoxanthine, guanine, cytosine, thymine, inosine, adenosine, guanosine, allantoin at 0.5–2.0 mM, and xanthine at 0.3 mM. L-ascorbic acid was added at 0.1–1% in the presence of sodium nitrate as N source. All media and chemical reagents were obtained from Sigma-Aldrich (Life Science Chemilab SA, Hellas) or AppliChem (Bioline Scientific SA, Hellas). 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 *uapA* alleles based on complementation of the pantothenic acid auxotrophy *pantoB100*.[26] *pabaA1* is a paraminobenzoic acid auxotrophy. *A. nidulans* protoplast isolation and transformation was performed as previously described.[48] Growth tests were performed at 37 ° C for 48 h, at pH 6.8. All strains used in this work are described in Table S1.

Standard molecular biology manipulations and plasmid construction

Genomic DNA extraction from *A. nidulans* was performed as described in FGSC (http://www.fgsc.net). Plasmids, prepared in *E. coli*, and DNA restriction or PCR fragments were purified from agarose 1% gels with the Nucleospin Plasmid Kit or Nucleospin ExtractII kit, according to the manufacturer's instructions (Macherey- Nagel, Lab Supplies Scientific SA, Hellas). Standard PCR reactions were performed using KAPATaq DNA polymerase (Kapa Biosystems). PCR products used for cloning, sequencing and re-introduction by transformation in *A. nidulans* were amplified by a high fidelity KAPA HiFi HotStart Ready Mix (Kapa Biosystems) polymerase. DNA sequences were determined by VBC-Genomics (Vienna, Austria). Site directed mutagenesis was carried out according to the instructions accompanying the Quik-Change® Site-Directed Mutagenesis Kit (Agilent Technologies, Stratagene). The principal vector used for most *A. nidulans* mutants is a modified pGEM-T-easy vector carrying a version of the *gpdA* promoter, the *trpC* 3′ termination region and the panB selection marker.[31] Mutations were constructed by oligonucleotide-directed mutagenesis or appropriate forward and reverse primers (Table S2). Transformants arising from single copy integration events with intact *uapA* ORFs were identified by Southern and PCR analysis

Correlated sequence conservation algorithm

The BIS2Analyzer (athttp://www.lcqb.upmc.fr/BIS2Analyzer/) was used for detected residues co-conserved with specific residues of the NAT motif. It is an openly accessible server providing online analysis of co-evolving amino-acid pairs in protein alignments, especially designed protein families which typically display a small number of highly similar sequences.[49]

Uptake assays

Kinetic analysis of wild-type and mutant UapA was measured by estimating uptake rates of [^{3}H]-xanthine uptake (40 Ci mmol^{-1}, Moravek Biochemicals, CA, USA), as previously described in.[26] In brief, [^{3}H]-xanthine uptake was assayed in *A. nidulans* conidiospores germinating for 4 h at 37° C, at 140 rpm, in liquid MM, pH 6.8. Initial velocities were measured on 10^7 conidiospores/100 μL by incubation with concentrations of 0.2–2.0 μM of [^{3}H]-xanthine at 37° C. For the competition experiments, initial uptake rates of [^{3}H]- xanthine were measured in the simultaneous presence increasing concentrations (5 μM − 1 mM) of various putative nucleobase-related inhibitors or L-ascorbic acid at 1%, as indicated. [^{3}H]-uracil uptake was also measured similarly. The time of incubation (1 or 2 min) was defined through time–course experiments. All transport assays were carried out at least in two independent experiments and the measurements in triplicate. Standard deviation was <20%. Results were analyzed in GraphPad Prism software.

Isolation and characterization of suppressor mutations

Suppressor mutations of 10^9 conidiospores of the strain UapA-**SYSE** were obtained after 3 min 45 sec exposure at a standard distance of 20 cm from an Osram HNS30 UV-B/C lamp and subsequent selection of colonies capable of growing on MM containing uric acid as sole nitrogen source, at 37 °C. Spores from positive colonies were collected after 4–8 days and further isolated on the same selective medium that was used to obtain the original colonies. Genomic DNA from 34 purified colonies was isolated and the ORF of UapA was amplified and sequenced. In all cases the amplified fragments contained a new mutation.

Epifluorescence microscopy

Samples for standard epifluorescence microscopy were prepared as previously described.[50,51] In brief, sterile 35 mm l-dishes with a glass bottom (Ibidi, Germany) containing liquid minimal media supplemented with $NaNO_3$ and 1% glucose were inoculated from a spore solution and incubated for 16 h at 25 °C. The samples 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 and contrast adjustment were made using the ZEN 2012 software while further processing of the TIFF files was made using Adobe Photoshop CS3 software for brightness adjustment, rotation and alignment.

Molecular simulations

Protein Model Construction. Models of UapA-TFAE, UapA-SYSE and UapA-SYSQ were generated by mutating the specific residues on Maestro platform (Maestro, version 2018-4, Schrödinger, LLC, New York, NY, 2018), on the basis of the structure of UapA crystallized in the inward-open conformation (PDB ID: 5i6c). In order to construct the model of UapA outward conformation Band3 transporter (4yzf) was used

as template.[43] Band3 is an anion exchanger of 14 transmembrane helices, crystallized in the outward conformation. The model was constructed using Prime Homology in Prime 2018-4 (Schrödinger, LLC, New York, NY, 2018), by superimposing helices one by one and aligning all the important residues.

Protein Preparation. The protein was prepared using the Protein Preparation Protocol implemented in Schödinger suite (Schrödinger Suite 2018, Protein Preparation Wizard) and accessible within the Maestro program (Maestro, version 2018-4, Schrödinger, LLC, New York, NY, 2018). Specifically, at first hydrogen atoms were added followed by an optimization of the orientation of hydroxyl groups of Asn, Gln, and of the protonation state of His in order to maximize hydrogen bonding. The final step was that of minimization of the protein, using the OLS3 force field.[52]

Ligand preparation Ligand preparation for docking was performed with LigPrep application (LigPrep, version 2018-4, Schrödinger, LLC, New York, NY, 2018), which consists of a series of steps that perform conversions, apply corrections to the structure, generate ionization states and tautomers, and optimize the geometries. The force field chosen was OPLS3.[52]

Induced Fit Docking Schrödinger Suite protocol was used (Schrödinger Suite 2018-4 Induced Fit Docking protocol; Glide, Schrödinger, LLC, New York, NY, 2016; Prime, Schrödinger, LLC, New York, NY, 2018), taking into account the side chain or backbone movements, or both, upon ligand binding, thus circumventing an inflexible binding site. In the first softened-potential docking step, of the protocol, 20 poses per ligand were retained. In the second step, for each docking pose, a full cycle of protein refinement was performed, with Prime 2018-4 (Prime, version 3.0, Schrödinger, LLC, New York, NY, 2018) on all residues within 5 Å of any out of the 20 ligand poses. The Prime refinement starts by performing conformational search and by minimizing the side chains of the selected residues. After convergence to a low-energy result, an additional minimization of all selected residues (side chain and backbone) is performed with the Truncated-Newton algorithm using the OPLS3 parameter set[52] and a surface Generalized Born implicit solvent model. The ranking of the obtained complexes is implemented according to Prime calculated energy (molecular mechanics and solvation), and the complexes within 30 kcal/mol of the minimum energy structure are used in the last step of the process, redocking with Glide (Schrödinger, LLC, New York, NY, 2016) using standard precision, and scoring. Finally, the ligands used in the first docking step are re-docked into each of the receptor structures retained from the Prime refinement step. The final ranking of the complexes is performed by a composite score which takes into account the receptor–ligand interaction energy (GlideScore) and receptor strain and solvation energies (Prime energy).

Molecular Dynamics (MD). For the construction of the protein–ligand complex CHARMM-GUI platform was used. Each model was inserted into a heterogeneous fully hydrated bilayer 120 Å × 120 Å × 120 Å, consisting of DPPC lipids and ergosterol. The membrane embedded system was solvated with TIP3P water molecules, neutralizing with counter ions, and adding 150 mM Na^+ and Cl^-. CHARMM36m[53] force field was used for protein and lipids, while the ligand was prepared using Antechamber[54] and the general Amber force field.[55] The protein orientation into the membrane was calculated using the PPM server.[56] The assembled simulation system consisted of ~130,000 atoms. The systems were simulated using GROMACS software.[57] The models were minimized and equilibrated to obtain stable structures. Minimization was carried out for 2,000 steps with a step size of 0.001 kJ/mol applying a steepest descent followed by a conjugate gradient algorithm, and the system was equilibrated for 20 ns by gradually heating and releasing the restraints to expedite stabilization. Finally, the system was further simulated free of restraints at a constant temperature of 300 K for 100 ns, using Nose-Hoover thermostat[58] and Parrinello-Rahman semi-isotropic pressure coupling[59] and compressibility at 4.5e-5 bar-1. The van der Waals and electrostatic interactions were smoothly switched off at 1.2 nm by switching function, while long-range electrostatic interactions were calculated using the particle mesh Ewald method.[60]

CRediT authorship contribution statement

Anezia Kourkoulou: Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing - original draft, Writing - review & editing. **Iliana Zantza:** Data curation, Formal analysis, Investigation, Methodology. **Konstantina Foti:** Investigation. **Emmanuel Mikros:** Formal analysis, Funding acquisition, Resources, Supervision, Validation, Software, Writing - original draft, Writing - review & editing. **George Diallinas:** Conceptualization, Formal analysis, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing - original draft, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by a "Stavros S. Niarchos Foundation" grant and by computational time granted from the Greek Research & Technology Network (GRNET) in the National HPC facility -ARIS- under project NCS1_Mechanism (pr006040).

Author contributions

A.K. performed genetic, molecular, biochemical cell biology experiments, analyzed and discussed results. K.F. assisted in genetic experiments. I.Z. and E.M performed the *in silico* analysis, the Molecular Dynamics, and analyzed *in silico* results. E.M. wrote the *in silico* part of the article. G.D conceived and planned experiments analyzed results and wrote the article.

Appendix A. Supplementary Data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2021.166814.

Received 26 October 2020;
Accepted 4 January 2021;
Available online xxxx

Keywords:
Aspergillus nidulans;
Fungi;
UapA;
Genetics;
Nucleobase

References

1. Algüel, Y., Cameron, A.D., Diallinas, G., Byrne, B., (2016). Transporter oligomerization: form and function. *Biochem. Soc. Trans.*, **44**, 1737–1744.
2. Diallinas, G., Martzoukou, O., (2019). Transporter membrane traffic and function: lessons from a mould. *FEBS J.*, **286**, 4861–4875.
3. Diallinas, G., Gournas, C., (2008). Structure-function relationships in the nucleobase- ascorbate transporter (NAT) family: lessons from model microbial genetic systems. *Channels (Austin).*, **2** (5), 363–372.
4. 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.
5. 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.
6. Niopek-Witz, S., Deppe, J., Lemieux, M.J., Möhlmann, T., (2014). Biochemical characterization and structure-function relationship of two plant NCS2 proteins, the nucleobase transporters NAT3 and NAT12 from Arabidopsis thaliana. *BBA*, **1838**, 3025–3035.
7. 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.
8. Bürzle, M., Suzuki, Y., Ackermann, D., Miyazaki, H., Maeda, N., Clémençon, B., Burrier, R., Hediger, M.A., (2013). The sodium-dependent ascorbic acid transporter family SLC23. *Mol. Aspects Med.*, **34**, 436–454.
9. Kourkoulou, A., Pittis, A.A., Diallinas, G., (2018). Evolution of substrate specificity in the Nucleobase-Ascorbate Transporter (NAT) protein family. *Microb. Cell.*, **5**, 280–292.
10. Maurino, V.G., Grube, E., Zielinski, J., Schild, A., Fischer, K., Flügge, U.I., (2006). Identification and expression analysis of twelve members of the nucleobase- ascorbate transporter (NAT) gene family in Arabidopsis thaliana. *Plant Cell Physiol.*, **47**, 1381–1393.
11. Chai, W., Peng, X., Liu, B., Wang, J., Zhu, Z., Liu, Y., Zhao, K., Cheng, B., Si, W., Jiang, H., (2018). Comparative Genomics, Whole-Genome Re-sequencing and Expression Profile Analysis of Nucleobase: Cation Symporter 2 (NCS2) Genes in Maize. *Front. Plant Sci.*, **9**, 856.
12. Algüel, Y., Amillis, S., Leung, J., Lambrinidis, G., Capaldi, S., Scull, N.J., Craven, G., Iwata, S., Armstrong, A., Mikros, E., Diallinas, G., Cameron, A.D., Byrne, B., (2016). Structure of eukaryotic purine/H(+) symporter UapA suggests a role for homodimerization in transport activity. *Nature Commun.*, **7**, 11336.
13. Diallinas, G., (2016). Dissection of Transporter Function: From Genetics to Structure. *Trends Genet.*, **32**, 576–590.
14. 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.
15. 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.
16. 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.
17. 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.
18. 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.
19. 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.

20. 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
21. 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.*, **2000** (17), 47–57.
22. 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.
23. Yu, X., Yang, G., Yan, C., Baylon, J.L., Jiang, J., Fan, H., Lu, G., Hasegawa, K., Okumura, H., Wang, T., Tajkhorshid, E., Li, S., Yan, N., (2017). Dimeric structure of the uracil:proton symporter UraA provides mechanistic insights into the SLC4/23/26 transporters. *Cell Res.*, **27**, 1020–1033.
24. Karatza, P., Panos, P., Georgopoulou, E., Frillingos, S., (2006). Cysteine-scanning analysis of the nucleobase-ascorbate transporter signature motif in YgfO permease of Escherichia coli: Gln-324 and Asn-325 are essential, and Ile-329-Val-339 form an alpha-helix. *J. Biol. Chem.*, **281**, 39881–39890.
25. 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.
26. 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.
27. Krypotou, E., Lambrinidis, G., Evangelidis, T., Mikros, E., 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.
28. Krypotou, E., Kosti, V., Amillis, S., Myrianthopoulos, V., Mikros, E., 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.
29. Sioupouli, G., Lambrinidis, G., Mikros, E., Amillis, S., Diallinas, G., (2017). Cryptic purine transporters in Aspergillus nidulans reveal the role of specific residues in the evolution of specificity in the NCS1 family. *Mol. Microbiol.*, **103**, 319–332.
30. 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.
31. Krypotou, E., Evangelidis, T., Bobonis, J., Pittis, A.A., Gabaldón, T., Scazzocchio, C., Mikros, E., Diallinas, G., (2015). Origin, diversification and substrate specificity in the family of NCS1/FUR transporters. *Mol. Microbiol.*, **96**, 927–950.
32. 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.
33. Papadaki, G.F., Amillis, S., Diallinas, G., (2017). Substrate Specificity of the FurE Transporter Is Determined by Cytoplasmic Terminal Domain Interactions. *Genetics*, **207**, 1387–1400.
34. Smirnoff, N., (2018). Ascorbic acid metabolism and functions: A comparison of plants and mammals. *Free Radic. Biol. Med.*, **122**, 116–129.
35. Sahin-Tóth, M., Akhoon, K.M., Runner, J., Kaback, H.R., (2000). Ligand recognition by the lactose permease of Escherichia coli: Specificity and affinity are defined by distinct structural elements of galactopyranosides. *Biochemistry*, **39**, 5097–5103.
36. Sahin-Tóth, M., Gunawan, P., Lawrence, M.C., Toyokuni, T., Kaback, H.R., (2002). Binding of hydrophobic D-galactopyranosides to the lactose permease of Escherichia coli. *Biochemistry*, **41**, 13039–13045.
37. Papakostas, K., Frillingos, S., (2012). Substrate selectivity of YgfU, a uric acid transporter from Escherichia coli. *J. Biol. Chem.*, **287**, 15684–15695.
38. Botou, M., Yalelis, V., Lazou, P., Zantza, I., Papakostas, K., Charalambous, V., Mikros, E., Flemetakis, E., Frillingos, S., (2020). Specificity profile of NAT/NCS2 purine transporters in Sinorhizobium (Ensifer) meliloti. *Mol. Microbiol.*, **114**, 151–171.
39. Amillis, S., Koukaki, M., 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.
40. Kourkoulou, A., Grevias, P., Lambrinidis, G., Pyle, E., Dionysopoulou, M., Politis, A., Mikros, E., Byrne, B., Diallinas, G., (2019). Specific Residues in a Purine Transporter Are Critical for Dimerization, ER Exit, and Function. *Genetics*, **213**, 1357–1372.
41. 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.
42. Emel'yanenko, V.N., Zaitsau, D.H., Verevkin, S.P., (2017). Thermochemical Properties of Xanthine and Hypoxanthine Revisited. *J. Chem. Eng. Data*, **62**, 2606–2609.
43. Arakawa, T., Kobayashi-Yurugi, T., Alguel, Y., Iwanari, H., Hatae, H., Iwata, M., Abe, Y., Hino, T., Ikeda-Suno, C., Kuma, H., Kang, D., Murata, T., Hamakubo, T., Cameron, A.D., Kobayashi, T., Hamasaki, N., Iwata, S., (2015). Crystal structure of the anion exchanger domain of human erythrocyte band 3. *Science*, **350**, 680–684.
44. Boudker, O., Ryan, R.M., Yernool, D., Shimamoto, K., Gouaux, E., (2007). Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. *Nature*, **445**, 387–393.
45. Botou, M., Lazou, P., Papakostas, K., Lambrinidis, G., Evangelidis, T., Mikros, E., Frillingos, S., (2018). Insight on specificity of uracil permeases of the NAT/NCS2 family from analysis of the transporter encoded in the pyrimidine utilization operon of Escherichia coli. *Mol. Microbiol.*, **2**, 204–219.

46. Sauer, D.B., Trebesch, N., Marden, J.J., Cocco, N., Song, J., Koide, A., Koide, S., Tajkhorshid, E., Wang, D.N., (2020). Structural basis for the reaction cycle of DASS dicarboxylate transporters. *Elife*., **9**, e61350
47. McCluskey, K., Wiest, A., Plamann, M., (2010). The Fungal Genetics Stock Center: repository for 50 years of fungal genetics research. *J. Biosci*., **35**, 119–126.
48. Koukaki, M., Giannoutsou, E., Karagouni, A., Diallinas, G., (2003). A novel improved method for Aspergillus nidulans transformation. *J. Microbiol. Methods*, **55**, 687–695.
49. Oteri, F., Nadalin, F., Champeimont, R., Carbone, A., (2017). BIS2Analyzer: a server for co-evolution analysis of conserved protein families. *Nucleic Acids Res*., **W1**, W307–W314.
50. 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.
51. 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.
52. Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J.Y., Wang, L., Lupyan, D., Dahlgren, M.K., Knight, J.L., Kaus, J.W., Cerutti, D.S., Krilov, G., Jorgensen, W.L., Abel, R., Friesner, R.A., (2016). OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput*., **12**, 281–296.
53. Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B.L., Grubmüller, H., MacKerell Jr., A.D., (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods*, **14**, 71–73.
54. Wang, J., Wang, W., Kollman, P.A., Case, D.A., (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model*., **25**, 247–260.
55. Wang, J., Wol, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A., (2004). Development and testing of a general amber force field. *J. Comput. Chem*., **25**, 1157–1174.
56. Lomize, M.A., Pogozheva, I.D., Joo, H., Mosberg, H.I., Lomize, A.L., (2012). OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res*., **40**, D370–D376.
57. Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., Lindah, E., (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, **1–2**, 19–25.
58. Evans, D.J., Holian, B.L., (1985). The Nose-Hoover thermostat. *J. Chem. Phys*., **83**, 4069–4074.
59. Parrinello, M., Rahman, A., (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys*., **52**, 7182–7190.
60. Darden, T., York, D., Pedersen, L., (1993). Particle mesh Ewald: An Nṡlog(N) method for Ewald sums in large systems. *J. Chem. Phys*., **98**, 10089–10092.