Substrate translocation in SLC23 and SLC26 transporters

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt beim Fachbereich

Biochemie, Chemie und Pharmazie

der Johann Wolfgang Goethe-Universität

in Frankfurt am Main

von

Katharina Elisabeth Augusta Holzhüter

aus Köln

Frankfurt am Main, 2021

(D30)

Vom Fachbereich Biochemie, Chemie und Pharmazie (FB14)

der Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Clemens Glaubitz

Gutachter: 1. Prof. Dr. Klaas Martinus Pos

                2. Dr. Rupert Abele

Datum der Disputation:

Content

Summary

Specialized transporter proteins facilitate controlled uptake and extrusion of molecules across biological membranes that would otherwise be impermeable to them. The superfamily of solute carriers (SLC) comprises the second largest group of membrane proteins in humans, acting on a variety of small polar and non-polar molecules and ions. Because of their central role in metabolism, malfunctioning of these proteins often is pathogenic. The interest in SLC transporters as drug targets – as well as for drug delivery – has therefore increased in the past years. For many SLC subfamilies, however, structural and functional information remains scarce to date.

A number of structures showed that the transmembrane domains of the SLC4, SLC23 and SLC26 families share a common fold, the so-called seven transmembrane segments inverted repeat (7TMIR) fold. This is in stark contrast to their diverging substrate specificities: Functionally characterized SLC4 proteins transport bicarbonate and borate, SLC23 proteins are specific for nucleobases in all kingdoms of life with the exception of higher eucaryotes, in which they transport ascorbate instead, and members of the SLC26 family translocate a variety of different anions such as chloride, sulfate or fumarate. The substrate binding sites of the 7TMIR proteins are well conserved within each of the SLC4, SLC23 and SLC26 families, but differ – in line with the diverging substrate specificities – among each other. The exception is a conserved glutamate residue on TM8 located at the substrate binding site that is found in all three protein families and is therefore considered as crucial for providing the molecular basis for ion coupling.

Based on comparison of structures of 7TMIR proteins in different conformations, these proteins were suggested to facilitate alternating access to the two sites of the membrane by a so-called elevator mechanism. This mechanism is characterized by a lateral and vertical motion of one domain that is harboring the substrate binding site against another domain, which remains relatively static during the transport cycle. Although the number of available structures for the 7TMIR protein families has increased over the past years, important features of the substrate translocation mechanism are still not understood. The data presented here addresses some of these open questions, in particular concerning the molecular basis of substrate binding and ion coupling as well as the relevance of dimerization for substrate transport.

The SLC23 nucleobase transporters are widely reported to being proton-coupled, in contrast to the ones found in mammalian organisms that transport ascorbate in a sodium-dependent manner. Without exception, proton-coupling has been suggested on the basis of whole-cell based uptake assays using radioactively labeled nucleobases. For a number of SLC23 transporters, uptake was shown to be largely reduced in the presence of the protonophore CCCP. For the here presented work, similar assays

Summary

were not only performed for the *E. coli* uracil transporter UraA, the first 7TMIR protein for which a structure was solved, but also for two bacterial UraA homologs, 51ThXi and 55PyPi. For both of them, uracil was identified as the main substrate as well. Similar as for UraA, in *E. coli* overexpressing the homolog 51ThXi uracil uptake was diminished in the presence of CCCP, suggesting that also for this homolog nucleobase transport is coupled to the proton-motive force. However, a more detailed analysis of uracil uptake into *E. coli* revealed several shortcomings of this kind of *in vivo* uptake assay that has so far not been addressed in published work but precluded a further analysis of the transport mechanism in this setting. In order to circumvent the limitations of the *in vivo* activity assay, a proteoliposome-based uptake assay was developed for 51ThXi. This assay poses the first and to date only *in vitro* transport assay for the SLC23 family. Strikingly, uracil uptake of reconstituted 51ThXi was not affected by any additional outer gradients, suggesting that this protein works as a uniporter rather than a proton symporter.

In order to explain this contradiction, the role of uracil phosphoribosyltransferase (UPRT) for *in vivo* uracil uptake was explored. UPRT catalyzes the immediate downstream reaction of uracil uptake *in vivo,* the phosphoribosylation of uracil to UMP. The gene coding for UPRT, *upp*, is located in *E. coli* in one bicistronic operon with *uraA*. It has already been demonstrated before that deletion of *upp* largely reduces uptake of uracil in *E. coli*. Herein, it is demonstrated that uracil uptake in *Δupp* cells is not only reduced when overexpressing UraA, but any uracil transporter tested, indicating a more general role of UPRT activity for overall uracil uptake. Uracil uptake in a *Δupp* strain overexpressing UPRT was increased several-fold over uptake in cells expressing *upp* endogenously, suggesting that in a *ΔuraA* strain - that is commonly used to test UraA activity in literature – not the transport by UraA, but the activity of UPRT can be rate-limiting. We were able to show that the reduced uptake of uracil in presence of the protonophore is not based on proton-coupling in UraA, but due to an effect on the downstream reaction catalyzed by UPRT, caused by inactivation of this enzyme at low pH. For several UPRT homologs, a strict pH-dependency of activity has been reported. Establishing a novel activity assay for purified UPRT from *E. coli*, we here demonstrate that also this enzyme exhibits a narrow pH optimum with maximum activity around pH 8.5. Importantly, barely any activity was observed below pH 7.0. It has been shown that *E. coli* cells that are treated with protonophores such as CCCP fail to restore the luminal pH, which normally is rapidly adjusted to neutral when the cells are exposed to an acidic environment (Wilks and Slonczewski, 2007; Martinez *et al.*, 2012). We therefore conclude that UraA is not a proton-coupled symporter, but rather works as a uniporter. Modification of uracil to UMP, that is no longer a substrate for the transport protein, maintains the substate gradient across the membrane. Inactivation of UPRT at low pH leads to a reduction of overall uracil uptake, because uracil is only equilibrated and no longer accumulated in the modified form of UMP. The phenotype observed in a whole-cell based uracil uptake assay in the presence of CCCP therefore mimics the one

Summary

of a proton-coupled symporter. For other nucleobases similar downstream modifying enzymes were described, and for some of them, a crucial dependency for activity of pH was shown as well; uniporter function might therefore be a general feature of the SLC23 nucleobase transporters.

SLC23 transporters are said to be very specific for their substrates. The best functionally characterized homologs are the ones from *E. coli* as well as UapA from *Aspergillus nidulans*, for which also a structure is available. The molecular basis of substrate specificity and the implications for the transport mechanism yet are not well-understood. So far, only one study exists that investigates the binding of different nucleobase analogs to uracil-specific SLC23 transporters (Botou *et al.*, 2018). Therein, competition assays in *E. coli* were performed using a number of nucleobase analogs, most of them purine compounds. None of them appeared to bind to UraA, which is in line with the general finding that most SLC23 transporter recognize either purines or pyrimidines. Herein, binding as well as transport studies were therefore performed using pyrimidine analogs. Exchange assays in proteoliposomes revealed that the UraA homolog 51ThXi not only binds but also transports 5-fluorouracil, a known powerful cytotoxic, as well as cytosine that so far was not considered to be a substrate for the SLC23 family at all. The binding sites of SLC23 transporters contain a number of highly conserved residues, including three polar or charged residues found in hydrogen-bonding distance to the substrate. In UraA, these are two glutamates on TM8 and TM10, and one histidine next to the TM8 glutamate. For all three of these residues, it has been reported that replacement by alanine completely abrogates uracil binding. According to sequence alignments and Phyre2 modeling, the 51ThXi binding site harbors identical residues at corresponding positions, while the histidine on TM8 in UraA and 51ThXi is replaced by aspartate in 55PyPi. Also for the two homologs, uracil transport was severely reduced when any of the residues was mutated to alanine. Interestingly, in contrast to what has been reported, DSF experiments suggested that substrate binding *per se* was still possible. Rather, both E256A and H260A mutants of 51ThXi showed markedly reduced thermal stability in absence of substrate. The glutamate on TM8 was additionally found to being irreplaceable for transport, but not essential for binding, as exemplified by the finding that 51ThXi(E256A) was able to bind 3-methyluracil which is not recognized by the wildtype protein; the methyl group fits into the space that is created by removing the glutamate, indicating that the overall binding site is uncompromised. Although the glutamate and histidine on TM8 are located next to each other, charge compensation appears not to play a major role for protein stability, as 55PyPi harbors aspartate instead of histidine by nature. Moreover, the histidine could be replaced by aspartate in both UraA and 51ThXi without affecting transport activity. Importantly, the glutamate on TM8 is not only conserved in the SLC23, but also in the SLC4 and SLC26 families. The crucial role we find for protein stability might be the underlying cause why this residue was preserved throughout evolution, while the three protein families developed different substrate specificities and transport modes. The Na$^+$/H$^+$ antiporters of the SLC9 family that

Summary

are assumed to work by an elevator mechanism contain two pairs of discontinuous helices at their substrate binding sites. This was suggested to create a surplus of charges based on the helix dipoles that is counteracted by conserved oppositely charged residues. Removal of this charges caused severe protein instability, similar to what we observed for UraA and 51ThXi. The substrate binding sites of the 7TMIR proteins do not contain two pairs of disrupted helices, but two short helices each preceded by a short ß-strand and an extended, unwound stretch of amino acids. The orientation of the helices might result in a net positive charge that is balanced by the conserved glutamate in a similar fashion.

To date, all elevator proteins identified so far were found to form higher oligomers. As substrate binding is confined to the substrate binding domain that moves against the scaffold domain, each protomer harbors its own translocation pathway. As a consequence, the functional implication of oligomerization is not immediately obvious and has been addressed experimentally only in a few cases. Additionally, a limited number of studies indicate that in some elevator proteins functional crosstalk exists, while for other, the protomers appear to function independently. Using transport studies in cells and proteoliposomes, we investigated cooperativity in dimers of 51ThXi and the SLC26 homolog SLC26Si, a proton-oxalate symporter from *Sulfitobacter indolifex*. A concatemeric version of 51ThXi, in which two protomers were fused by a short amino acid linker, exhibited wild type like activity in *in vivo* uracil uptake assays. When one of the two protomers carried the E256A mutation, which is impaired in substrate transport, uptake was reduced to levels of the concatemer of two inactive subunits. This suggests that for 51ThXi both protomers need to be active for efficient substrate transport. A similar experiment performed in proteoliposomes, mixing wild type 51ThXi with the inactive mutant in different ratios, confirmed this finding. A similar negative dominance of one protomer over the other was also described before for the SLC23 transporter UapA, while for UraA, the two protomers were suggested to work largely independently from each other. It appears that within the same protein family, differences in functional crosstalk between the protomers might exist. This is in line with observations made for the SLC26 family: Functional studies published before on the homolog SLC26Dg suggested a positive cooperativity within the dimer. Here, a similar approach was used for the homolog SLC26Si, determining initial transport rates in proteoliposomes of wild type protein mixed in different ratios with a crosslinked, immobilized protein version. While for SLC26Dg the net activity was higher than what would be expected from independently working subunits, for SLC26Si the opposite effect was observed, with a lower activity of the wild type protomer when paired with the immobilized version. For SLC26Dg, immobilization of one protomer was suggested to extend the scaffold in such a way that the substrate binding domain of the active protomer can move more efficiently against it. In a similar fashion, immobilization of one protomer in SLC26Si might cause a less efficient movement, resulting in a lower activity. The molecular basis for this, however, is not clear at the moment.

Summary

SLC26 proteins vary from the SLC23 family, next to the differing substrate specificity, in the presence of an additional, C-terminal domain, the so-called STAS domain. Interestingly, removal of this domain reduces protein activity, although the substrate translocation is confined to the transmembrane domain. Soluble STAS-domain only proteins as well as other STAS-fusion proteins with different functions have been described, but the precise role of the STAS domain for the transport cycle in the SLC26 family is largely unknown. The first structure of a member of the SLC26 family, the one of SLC26Dg, was that of a monomer, although biochemical data indicated that the protein is a dimer in the lipidic environment. More recent structures of SLC26 proteins indeed showed structures of dimers, in which both protomers intersect each other to place the STAS domain of one protomer close to the transmembrane domain of the other protomer and *vice versa*. Herein, it is demonstrated by covalent crosslinking as well as EPR that SLC26Dg exhibits the same overall architecture. Interestingly, mutating the short linker that attaches the cytosolic STAS domain to the transmembrane domain to a GS-linker, thus making it more flexible, severely reduced the activity of reconstituted protein in both the cases of SLC26Dg and SLC26Si. EPR distance measurements, however, indicated that the distances between STAS domain and transmembrane domain were only slightly increased, suggesting that the positioning of the STAS domain with respect to the transmembrane domain was mostly preserved. Although the SLC4, SLC23 and SLC26 share the same architecture of the transmembrane part, the dimer interfaces vary considerably. The transmembrane domain in the SLC26 family was shown to contribute only with a small part of TM14 to the dimer interface, whereas in the other families, this interface is larger. Although the STAS-less SLC26 proteins appear to still be able to dimerize, the transport studies included here indicates that the STAS domain contributes to the dimer interface in a fashion that is crucial for correct protein function.

Overall, the here presented data provides important insights into different aspects of the transport mechanism of the SLC23 and SLC26 protein families. Importantly, we show that SLC23 nucleobase transporters, in contrast to what was been previously reported, work as uniporters rather than as proton-coupled symporters. In order to do so, we developed the first and only *in vitro* transport assay for the SLC23 family, which enables investigation of protein function in a defined environment. Moreover, we provide a hypothesis on the role of the extremely conserved negative charged substrate binding site residue found not only in the SLC23, but also SLC4 and SLC26 families. Based on a detailed analysis of binding and transport we conclude that this conserved negative charged has a relevance for protein stability rather than for substrate binding, which explains its conservation for all three protein families that otherwise differ in their substrate specificities and modes of transport. Lastly, we investigated the relevance of oligomerization for the SLC23 and SLC26 families, highlighting the importance of the STAS domain for forming active dimers in the SLC26 anion transporter family.

Zusammenfassung

Spezielle Proteine ermöglichen den kontrollierten Transport von Molekülen über biologische Membranen, die sonst undurchlässig für sie sind. Die zweitgrößte Familie von Membranproteinen im Menschen ist die Superfamilie der Solute Carrier (SLC), die eine Vielzahl verschiedener kleiner polarer und unpolarer Moleküle und Ionen transportieren. Aufgrund ihrer zentralen Rolle im Stoffwechsel ist eine Fehlfunktion dieser Proteine oft Ursache für Krankheiten. Das Interesse an SLC-Transportern als Zielstrukturen für Arzneimittel – und auch für deren Verabreichung - hat daher in den letzten Jahren zugenommen. Für viele SLC-Proteinfamilien jedoch mangelt es noch an Information über Struktur und Funktion.

Strukturen von Mitgliedern der SLC4-, SLC23- und SLC26-Familien haben gezeigt, dass deren Transmembrandomäne dieselbe Faltung aufweist, die *seven transmembrane segments inverted repeat* (7TMIR) genannt wird. Die Substratspezifitäten dieser Proteinfamilien unterscheiden sich jedoch deutlich: SLC4-Proteine sind spezifisch für Hydrogenkarbonat und Borat, während Mitglieder der SLC26-Familie eine Vielzahl verschiedener Anionen transportieren, z.B. Chlorid, Sulfat oder Fumarat. SLC23-Proteine dagegen transportieren Nukleobasen oder, aber nur in höheren Eukaryoten, Ascorbat. Innerhalb der einzelnen Familien sind die Substratbindestellen von 7TMIR-Proteinen stark konserviert, unterscheiden sich aber untereinander, analog zu den verschiedenen Substratspezifitäten. Eine Ausnahme bildet ein extrem konservierter Glutamatrest, der in allen drei Proteinfamilien zu finden ist und deshalb als bedeutsam für die Ionenkopplung vorgeschlagen wurde.

Ein Vergleich von Strukturen von 7TMIR-Proteinen in verschiedenen Konformationen führte zu der Annahme, dass diese Proteine einen sogenannten *elevator*-Mechanismus benutzen, um die Substratbindestelle abwechselnd von beiden Seiten der Membran zugänglich zu machen. Dieser Mechanismus ist gekennzeichnet durch eine laterale und vertikale Bewegung einer Domäne, die die Substratbindungsstelle beherbergt, gegen eine andere Domäne, die während des Transportzyklus relativ statisch bleibt. Obwohl in den letzten Jahren mehrere Strukturen von 7TMIR-Proteinen gelöst wurden, verbleiben viele Aspekte des Transportmechanismus unerforscht. Die Ergebnisse, die in dieser Arbeit präsentiert werden, sollen einige dieser Aspekte beleuchten. Dies sind vor allem die molekulare Grundlage für Substratspezifität und Ionenkopplung sowie die Notwendigkeit der Dimerisierung für den Transportzyklus.

Es wird allgemein berichtet, dass die Nukleobasen-spezifischen SLC23-Transporter protonengekoppelt sind, im Gegensatz zu denen in Säugetieren, in denen Ascorbat natriumabhängig transportiert wird. Als Nachweis für die Protonenkopplung wurde, ausnahmslos, ein *in vivo* Transportassay von radioaktiv markierten Substraten herangezogen. Für eine Reihe von SLC23-Transportern wurde gezeigt, dass die Aufnahme von Nukleobasen in Gegenwart von Protonophoren weitgehend reduziert ist. UraA, ein Uracil-Transporter aus *E. coli*, ist das erste SLC23-Protein, für das eine Struktur gelöst wurde. Im

Zusammenfassung

Rahmen der hier beschriebenen Arbeit wurden ähnliche Transportassays nicht nur für UraA durchgeführt, sondern auch für zwei bakterielle UraA-Homologe (51ThXi aus *Thalassospira xiamenensis* und 55PyPi aus *Pyramidobacter piscolens*). Für beide Homologe wurde Uracil als Substrat identifiziert. Auch für 51ThXi, ähnlich wie für UraA, wurde eine niedrigere Aktivität in der Gegenwart von CCCP gemessen, was auf eine Protonenkopplung hindeutet. Eine detaillierte Analyse dieser *in vivo* Transportassays offenbarte aber eine Reihe von Mängeln, die es unmöglich machten, zweifelsfreie Untersuchungen des Transportmechanismus in diesem Kontext durchzuführen. Aus diesem Grund wurde ein liposomenbasierter *in vitro*-Assay für 51ThXi entwickelt, der erste und bislang einzige *in vitro*-Transportassay für die SLC23-Familie. Überraschenderweise war in diesem Assay der Transport von Uracil von sekundären Gradienten nicht beeinflusst, was eher für eine Uniporter- als eine Symporterfunktion spricht.

Um diesen Widerspruch zwischen dem Transport in Zellen und dem in Liposomen zu erklären, wurde die Bedeutung der Uracilphosphoribosyltransferase (UPRT) für den Transport von Uracil *in vivo* untersucht. UPRT katalysiert die unmittelbare Folgereaktion der Uracil-Aufnahme *in vivo*, nämlich die Phosphoribosylierung von Uracil zu UMP. Das Gen, das für UPRT codiert (*upp*), befindet sich in *E. coli* in einem bicistronischen Operon mit *uraA*. Die Deletion von *upp* führt zu einer verminderten Aufnahme von Uracil in *E. coli*, wie schon vorher bekannt war. In der hier vorliegenden Arbeit wird gezeigt, dass eine Deletion von *upp* nicht nur zur verminderten Aufnahme von Uracil durch UraA führt , sondern auch durch andere Uracil-spezifische Transporter. Das spricht für eine grundlegende Bedeutung der UPRT-Aktivität für die Aufnahme von Uracil. Die Uracilaufnahme in einem *upp*-Knockoutstamm, der UPRT überexprimiert, war um ein Vielfaches erhöht gegenüber der in einem Stamm, der das Enzym endogen exprimiert. Als Folge dessen könnte für die Aufnahme in einem *ΔuraA*-Stamm – der standardmäßig für Transportassays verwendet wird, wie sie in der Literatur zu finden sind – nicht der Transport durch UraA, sondern die Aktivität von UPRT ratenbestimmend sein. Für eine Vielzahl von UPRT-Homologen wurde bereits eine pH-Abhängigkeit der Aktivität gezeigt. Anhand eines neu entwickelten Aktivitätsassays für *E. coli* UPRT konnte auch hier ein eng begrenztes Optimum um pH 8,5 für dieses Enzym gezeigt werden. Unterhalb von pH 7,0 war fast keine Aktivität mehr festzustellen. Es wurde schon früher gezeigt, dass *E. coli*, die mit Protonophoren wie CCCP behandelt werden, nicht in der Lage sind, den pH des Lumens wiederherzustellen. Dieser wird normalerweise schnell auf einen neutralen Wert zurückreguliert, wenn die Zellen einem sauren Medium ausgesetzt werden (Wilks and Slonczewski, 2007; Martinez *et al.*, 2012). Wir schlussfolgern daher, dass der Effekt von CCCP auf die Uracilaufnahme durch UraA oder andere SLC23-Transport nicht eine Folge von protonengekoppeltem Transport ist, sondern auf der Inaktivierung der UPRT bei niedrigem pH basiert. UraA ist damit kein sekundär aktiver Transporter, sondern ermöglicht die Uracilaufnahme mittels *facilitative diffusion*. Die Modifizierung von Uracil zu UMP, das kein Substrat für den Transporter ist, sichert die Aufrechterhaltung des Substratgradienten. Da für andere Nukleobasen ähnliche modifizierende

Zusammenfassung

Enzyme beschrieben wurde, ist es denkbar, dass andere SLC23-Transporter ebenfalls als Uniporter funktionieren.

SLC23-Transporter werden als sehr spezifisch beschrieben. Die bestcharakterisierten Homologe sind die von *E. coli* sowie UapA aus *Aspergillus nidulans*, für das ebenfalls eine Struktur gelöst wurde. Die molekulare Grundlage für die Spezifität und deren Zusammenhang mit dem Transportmechanismus sind auf der anderen Seite weniger gut verstanden. Bislang existiert nur eine einzige detaillierte Untersuchung der Substratspezifität von Uraciltransportern (Botou *et al.*, 2018). In dieser wird die Kompetition verschiedener Analoga von Nukleobasen für den Transport von Uracil in *E. coli* untersucht. Dabei wurden größtenteils Derivate von Purinen verwendet. Von diesen scheint keines an UraA zu binden, was mit der generellen Annahme übereinstimmt, dass SLC23-Transporter entweder Purine oder Pyrimidine erkennen. In der vorliegenden Arbeit wurde daher Bindung und Transport vor allem von Pyrimidinderivaten untersucht. *Exchange assays* in Proteoliposomen zeigten, dass 51ThXi neben Uracil auch 5-fluorouracil, ein starkes Zytotoxikum, und Cytosin transportiert. Bislang wurde davon ausgegangen, dass Cytosin kein Substrat für die SLC23-Familie ist. Die Bindestelle von SLC23-Transportern enthält eine Reihe hochkonservierter Reste, unter anderem drei polare oder geladene Seitenketten, die sich in unmittelbarer Nähe zum Substrat befinden. In UraA handelt es sich dabei um zwei Glutamate in TM8 und TM10, sowie ein Histidin in TM8 benachbart zum Glutamat. Für alle drei Reste wurde ein vollständiger Aktivitätsverlust beschrieben, wenn sie zu Alanin mutiert werden. Die Substratbindestelle von 51ThXi enthält, einem Sequenz-Alignment und Phyre2-Model zufolge, dieselben Aminosäuren an den entsprechenden Positionen wie in UraA. Anstelle des Histidins in TM8 in UraA und 51ThXi ist in 55PyPi dagegen ein Aspartat zu finden. Auch für die beiden Homologe wurde ein starker Rückgang des Uraciltransports für die Alaninmutanten festgestellt. Im Gegensatz zu publizierten Ergebnissen zeigten DSF-Messungen jedoch, dass die Mutanten prinzipiell noch in der Lage waren, Substrat zu binden. Vielmehr zeigten die E256A- und H260A-Mutanten von 51ThXi in Abwesenheit von Substrat eine verringerte Temperaturstabilität. Außerdem wurde festgestellt, dass das Glutamat zwar für den Transport durch keine andere getestete Aminosäure ersetzt werden konnte, aber nicht ausschlaggebend für die Bindung des Substrats war. Dies spiegelt sich wider in der Tatsache, dass 51ThXi(E256A), im Gegensatz zum Wildtyp, 3-methyluracil binden kann: der Platz für die zusätzliche Methylgruppe wird im Wildtyp durch den Glutamatrest okkupiert. Obwohl sich in TM8 der konservierten Glutamat- und Histidinrest in unmittelbarer Nähe zueinander befinden, scheint eine Ladungskompensierung keine wesentliche Rolle zu spielen, da 55PyPi von Natur aus ein Aspartat anstelle von Histidin an dieser Stelle trägt. Darüber hinaus konnte sowohl für UraA als auch für 51ThXi das Histidin durch Aspartat ersetzt werden, ohne dass die Transportaktivität wesentlich beeinträchtigt wurde. Der konservierte Glutamatrest findet sich nicht nur in der SLC23-, sondern auch in den SLC4- und SLC26-Familien. Die Bedeutung für die Stabilität der Proteine könnte der Grund sein, warum dieser Rest evolutionär erhalten geblieben ist, während die drei Proteinfamilien sich voneinander weg hin zu verschiedenen Substratspezifitäten und Transportmodi entwickelten. Die Bindestelle der Na$^+$/H$^+$-

Zusammenfassung

Antiporter (SLC9), für die ein *elevator*-Mechanismus vorgeschlagen wurde, enthält zwei Paar unterbrochener Helices. Es wurde vorgeschlagen, dass diese unterbrochenen Helices, aufgrund der Helix-Dipolmomente, einen Überschuss an Ladung an der Substratbindestelle erzeugen, der durch entgegengesetzt geladene konservierte Reste ausgeglichen wird. Mutationen dieser Reste führte zu einer geringeren Proteinstabilität, ähnlich wie hier für UraA und 51ThXi beobachtet wurde. Die Substratbindestelle von 7TMIR-Proteinen enthält keine Paare unterbrochener Helices, sondern zwei kurze Helices, denen jeweils ein kurzer ß-Strang und eine ungefaltete Abfolge von Aminosäuren vorangeht. Die Orientierung der Helices könnte einen Überschuss positiver Ladung erzeugen, der durch den konservierten Glutamatrest ausgeglichen wird.

Alle bislang identifizierten *elevator*-Proteine bilden Oligomere. Substratbindung in diesen Proteinen ist auf eine Domäne beschränkt, die sich relativ zu einer anderen Domäne bewegt. Diese zweite Domäne ist dabei alleine verantwortlich für die Oligomerisierung. Folglich enthält jedes Protomer seinen eigenen Substratkanal, der unabhängig von dem in dem anderen Protomer ist. Es besteht also keine direkte strukturelle Notwendigkeit zur Oligomerisierung, und die Gründe dafür sind bislang wenig erforscht. Zusätzlich wurde, anhand einiger weniger wissenschaftlicher Studien, für einige *elevator*-Proteine ein funktioneller *crosstalk* zwischen den Protomeren vorgeschlagen, während in anderen die Protomere unabhängig voneinander zu transportieren scheinen. Anhand von Transportstudien in Zellen und Liposomen wurde hier die Kooperativität in Dimeren von 51ThXi und SLC26Si, einem SLC26-Homolog aus *Sulfitobacter indolifex*, untersucht. Ein 51ThXi-Konkatamer, in dem zwei Protomere durch einen kurzen Aminosäure-Linker verknüpft wurden, zeigte Wildtyp-Aktivität in *in vivo*-Transportassays. In Konkatameren, in denen eines der beiden Protomer die E256A-Mutation aufwies, war der Transport dagegen reduziert auf das Niveau der Fusion von zwei inaktiven Protomeren. Ein vergleichbares *in vitro*-Experiment, in dem die inaktive Mutante in verschiedenen Verhältnissen mit Wildtyp rekonstituiert wurde, zeigte ein ähnliches Ergebnis. Eine solche negative Dominanz wurde zuvor auch für den SLC23-Transporter UapA beschrieben. Für UraA dagegen wurde eine voneinander unabhängige Aktivität der Protomere vorgeschlagen. Es scheint, als könnten auch innerhalb einer Proteinfamilie Unterschiede existieren, was den funktionellen *crosstalk* zwischen den Protomeren anbelangt. Ähnliches wurde auch für die SLC26-Familie beobachtet. Funktionelle Studien an SLC26Dg, die zuvor veröffentlicht wurden, gaben Hinweise auf eine positive Kooperativität innerhalb des Dimers. Hier wurde ein ähnlicher Ansatz für das Homolog SLC26Si verwendet: Wildtyp wurde in verschiedenen Verhältnissen mit einer immobilisierten Variante des Proteins gemischt und rekonstituiert und die anfänglichen Transportraten gemessen. Während für SLC26Dg eine höhere Aktivität gesehen wurde, als von unabhängig funktionierenden Protomeren zu erwarten wäre, wurde für SLC26Si der gegenteilige Effekt beobachtet. Für SLC26Dg wurde vorgeschlagen, dass die immobilisierte Proteinvariante wie eine Vergrößerung des Membranankers wirkt, die eine effektivere Bewegung der Substratbindedomäne erlaubt. Die Immobilisierung eines Protomers könnte für SLC26Si, analog dazu,

zu einer weniger effizienten Bewegung des aktiven Protomers führen, sodass das Dimer insgesamt eine geringere Aktivität aufweist. Die molekularen Grundlagen dafür sind aber bislang unklar.

Die SLC26-Proteine unterscheiden sich von denen der SLC23-Familie neben den unterschiedlichen Substratspezifitäten außerdem durch eine zusätzliche, C-terminale lösliche Domäne, die sogenannte STAS-Domäne. Interessanterweise verursacht ein Fehlen dieser Domäne eine Inaktivierung des Proteins. Lösliche Proteine, die nur aus einer STAS-Domäne bestehen, und andere STAS-Fusionsproteine sind ebenfalls beschrieben worden. Aufgrund dessen wurden verschiedene Funktionen der STAS-Domäne für die SLC26-Familie vorgeschlagen. Über die tatsächliche Rolle für den Transportzyklus ist bislang aber wenig bekannt. Strukturen von Dimeren der SLC26-Familie zeigen, dass sich die zwei Protomere überschneiden, weshalb sich die STAS-Domäne eines Protomers nahe der Transmembrandomäne des anderen Protomers befindet. Kovalentes *Crosslinking* und EPR-Messungen zeigen hier, dass das Dimer von SLC26Dg einen ähnlichen Aufbau hat. Interessanterweise führt eine Mutation des kurzen Linkers zwischen der STAS-Domäne und der Transmembrandomäne zu einem flexibleren Glycin-Serin-Linker ebenfalls zu Aktivitätsverlust, sowohl für SLC26Dg, als auch SLC26Si. EPR-Messungen zeigten jedoch, dass die Position der STAS-Domänen hinsichtlich der Transmembrandomänen in den Linker-Mutanten größtenteils erhalten blieb. Zwar ist der Transmembranteil der SLC4-, SLC23- und SLC26-Familien ähnlich gefaltet, die Schnittstellen der Dimere unterscheiden sich jedoch stark. Innerhalb der Transmembrandomäne der SLC26-Proteine trägt nur ein kleiner Teil von TM14 zur Dimerschnittstelle bei. Obwohl die SLC26-Homologe, in denen die STAS-Domäne entfernt wurden, immer noch dimerisieren konnten, zeigen die hier vorgestellten Ergebnisse, dass die STAS-Domänen entscheidend zur Formung einer funktionellen Schnittstelle zwischen den Dimeren beitragen.

Insgesamt tragen die hier dargestellten Ergebnisse erheblich zu einem besseren Verständnis verschiedener Aspekte des *elevator*-Mechanismus bei. Vorrangig konnten wir zeigen, dass die SLC23-Transporter als Uniporter wirken, und nicht, wie allgemein angenommen, Protonen-gekoppelt. Des Weiteren haben wir den ersten und bislang einzigen *in vitro*-Transportassay für die SLC23-Familie entwickelt. Die detaillierte Analyse der Substratbindestelle der SLC23-Familie führte zur Formulierung einer Hypothese für die Relevanz der konservierten negativen Ladung für die Stabilität der 7TMIR-Proteine. Außerdem wurde die Relevanz der Oligomerisierung für den Transportzyklus in der SLC23-und SLC26-Familie untersucht, die die Bedeutung der STAS-Domäne für die Bildung aktiver Dimere in der SLC26-Familie aufzeigten.

1. Introduction

1.1.   The solute carrier (SLC) family of membrane transporters

Biological membranes are physical boundaries that prevent free diffusion of water-soluble biomolecules, thus confining the interior of the cell from the external environment and providing the basis for compartmentalization. In order to maintain essential cellular processes, membrane-embedded proteins mediate transport of substrates across the lipid bilayer. Up to 30% of the human genome is thought to code for membrane proteins, with the majority of those being transporters (Yin and Flynn, 2016; Gong *et al.*, 2019). Membrane proteins account for more than 60% of the targets of small-molecule drugs that are approved by the FDA (Santos *et al.*, 2019). The Transporter Classification Database (TCDB, https://www.tcdb.org/) classifies over 20,000 non-redundant protein sequences into more than 1000 transporter families (Saier *et al.*, 2006, 2016).

Among the human membrane proteins, the family of solute carriers (SLC) comprise the second-largest group, after the G-protein coupled receptors (GPCRs) (César-Razquin *et al.*, 2015). To date, over 400 human SLCs have been identified that were classified into over 65 sub-groups based on sequence homology and function (Hediger *et al.*, 2013; Perland and Fredriksson, 2017). They transport a vast variety of solutes, including ions, sugars, lipids, amino acids and neurotransmitters and are evolutionary ancient (Fredriksson *et al.*, 2008; Höglund *et al.*, 2011). Solute carriers include facilitative and coupled transporters and transporters with channel-like properties, but no primary active pumps (Bai *et al.*, 2017). For a vast number of SLCs, mutations have been linked to human disease, for example schizophrenia and Alzheimer's disease (SLC1 family), depression and autism (SLC6 family), chloride diarrhea or deafness (SLC26 family), and type 2 diabetes (several SLC families) (Hediger *et al.*, 2013; Hu *et al.*, 2020; Schumann *et al.*, 2020). Therefore, over the past years SLC transporters have gained increasing interest as both targets for drugs themselves, or for drug deliverance (Zhou *et al.*, 2017; Casiraghi *et al.*, 2021). Although more and more structures of SLC proteins are becoming available, structure-function relationships and their link to protein malfunctioning and pathophysiology remains poorly understood to date.

1.2.   Energetics of membrane transport

Lipid bilayers are virtually impermeable to large polar molecules or ions, but allow passage of small, non-polar molecules down their concentration gradient by simple diffusion. The transfer of other solutes, for which the diffusion rate over the lipid bilayer is slow, is mediated by specialized membrane transport proteins. These membrane transporters can be broadly divided into carriers and channels. Carriers bind their substrates with high specificity and undergo a series of conformational changes to

transport the solute across the membrane. Channels on the other hand exhibit, in general, less specific interactions and much higher rates of transport compared to carriers. Channels and many carriers mediate passive transport of solutes down their concentration gradients (facilitative diffusion). The transport of molecules that are charged is additionally influenced by the membrane potential, i.e., the electrical potential difference across the membrane caused by the uneven distribution of ions on both sides of the lipid bilayer. Typically, cellular membranes exhibit an electrical gradient that is negative on the inside. To accumulate solutes against their electrochemical gradients, active transporters use energy sources such as ATP or light (primary active transport), or couple transport to the concentration gradient of another, secondary substrate, typically protons or sodium ions. These secondary active transporters act in a bidirectional way, and the flow of substrates is governed by the net electrochemical gradients. The substrates can be transported in the same direction (symport) or in opposite directions (antiport). Generally, symporters have distinct binding sites for their substrates, while in antiporters, both substrates often compete for the same binding site. Antiporters cannot undergo the conformational change in absence of substrate, whereas for symporters, the empty transporter can switch between conformations (Bai *et al.*, 2017). Uniporters, that mediate uncoupled transport of solutes by facilitative diffusion, are considered to being closely related to antiporters; the difference lies in different energy barriers for the conformational transition of the unloaded carrier (Burtscher *et al.*, 2019). For secondary active transport, movement of one solute against its concentration gradient occurs as long as it is fueled by the energy stored in the electrochemical gradient of the second substrate. The ion gradients that are used for coupling are generated by primary active transporters. The major enzyme responsible for maintaining the sodium gradient across the membrane in human cells is the Na^+/K^+-ATPase, that converts the energy gained from ATP hydrolysis into the antiport of three Na^+ and two K^+. About 50% of the total ATP consumption in human brains originates from the activity of this protein (Pivovarov *et al.*, 2019; Gagnon and Delpire, 2021).

1.3.   *The elevator alternating-access mechanism*

In order to provide controlled translocation of substrates, secondary active transporters alternate between a minimum of two conformations, exposing the substrate binding site to either of the two sides of the membrane (Jardetzky, 1966). In this, currently three distinct mechanisms are distinguished: the rocker switch, rocking bundle, and elevator mechanisms (Drew and Boudker, 2016) (**Figure 1**).
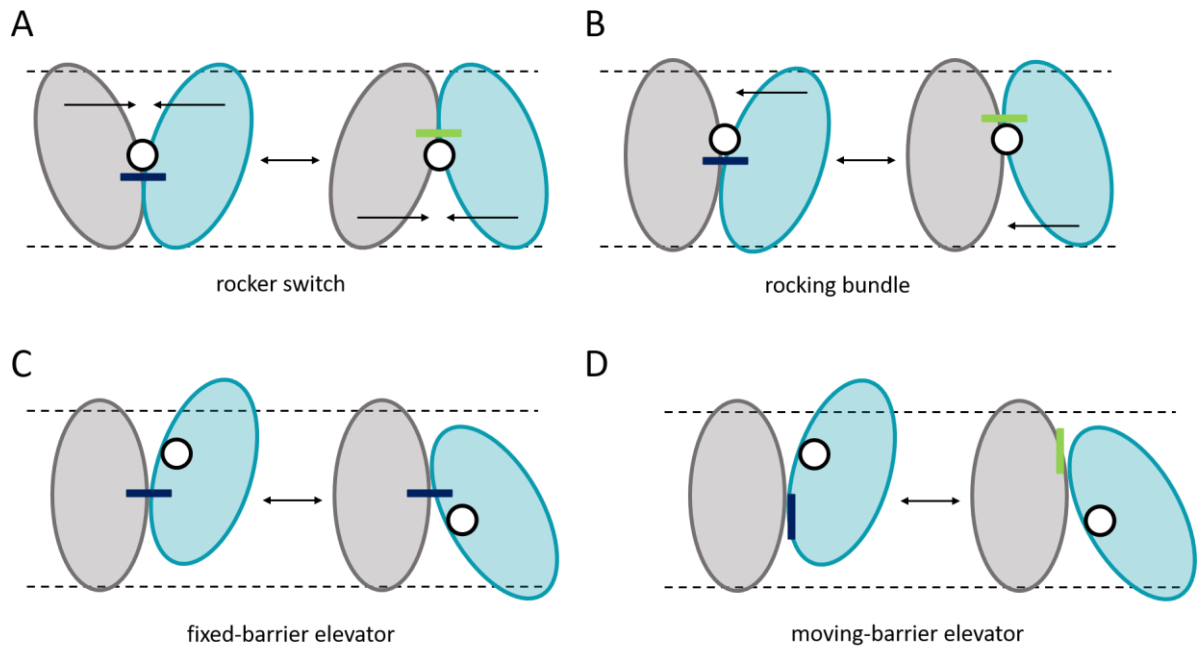
Figure 1: Different modes of alternating-access mechanisms. (A) In proteins working by a rocker switch mechanism, two protein domains move around the substrate binding site in order to allow access from either side of the membrane. The substrate is depicted as a circle. The conformational change is accompanied by gate opening and closing (blue and green) that prevents leakage. (B) The rocking bundle mechanism involves mostly movement of one domain against another one that remains relatively static. In both rocker-switch and rocking-bundle mechanisms, both domains contribute to substrate binding. (C) Fixed-barrier elevator mechanism. In contrast to the other two mechanisms, in elevator proteins substrate is bound by only one of the two domains. Substrate translocation is facilitated by an extensive movement of this domain against the other one that remains fixed. (D) Moving-barrier elevator mechanism. Initially, the fixed barrier was suggested as common to the elevator mechanism in general, but later proteins undergoing elevator-like movements with a moving barrier were described. Reviewed by Garaeva and Slotboom, 2020 (Garaeva and Slotboom, 2020).

In both rocker switch and rocking bundle mechanisms, the substrate is bound at the interface of two domains. External and internal barriers prevent substrate leakage and open and close during the transport cycle in order to allow access to the substrate binding site in the different conformations. The rocker switch mechanism is characterized by movements of both domains around the substrate binding site, whereas in the rocking bundle mechanism, mostly one domain moves while the other remains static. In the elevator mechanism, in contrast, substrate binding is more or less exclusively confined to one of the two domains that moves both vertical and lateral against the other domain to shuttle the substrate across the membrane. According to their function, the two domains are called substrate binding (or transport) domain and scaffold domain. More recently, more subtle differences between different types of elevator movements have been described: elevators with one or two gates, that can stay fixed or move during the conformational change (Garaeva and Slotboom, 2020). For many elevator proteins, structural elements such as helical hairpins have been defined that act as molecular gates, which upon closure prevent the substrate from leaving the transport domain. Open gates prevent the movement of the substrate binding against the scaffold domain. Rearrangement of the

helical hairpin due to substrate binding, referred to as gate closure, allows the elevator movement to take place. The role of the helical hairpin for the elevator movement has been well-investigated for the SLC1 family of glutamate and aspartate transporters as well as for CitS from *Klebsiella pneumoniae* (Boudker *et al.*, 2007; Verdon *et al.*, 2014; Kim *et al.*, 2017; Garaeva *et al.*, 2019; Arkhipova *et al.*, 2020). However, it should be noted that for many proteins that have been demonstrated or suggested to work as elevators, similar helical hairpin structures that act as gates are missing. For VcINDY and LaINDY, that are members of the SLC13 family of Na$^+$-coupled dicarboxylate symporters and exchangers, it has been recently suggested that sodium ion binding causes a net positive charge at the substrate binding site that impedes the movement of the substrate binding domain against the scaffold domain. Binding of the negatively charged substrate neutralizes this positive charge, allowing the elevator movement to take place (Sauer *et al.*, 2020). For many other proteins that have been suggested to work by an elevator mechanism, it is yet not well-understood what impedes or promotes the transition between the different conformations, or how coupling is achieved. One reason for this is that for many only a limited number of conformations was structurally resolved. Comparison of different proteins in different conformations can lead to contradictory conclusions, even within the same family. For example, for the SLC4 family, comparing the outward-facing structure of AE1 with an occluded state of *Arabidopsis thaliana* Bor1 indicates an elevator motion, while in comparison with the inward-facing structure of Bor1p from *Saccharomyces mikatae*, characteristics of a rocking-bundle mechanism are seen (Thurtle-Schmidt and Stroud, 2016; Coudray *et al.*, 2017). Seemingly, a clear distinction between the different mechanisms can be difficult. Comparison of the inward- and outward-facing structures of *Thermus thermophilus* NapA indicate both rotation and vertical movement of the substrate binding domain, consistent with an elevator mechanism. On the other hand, molecular dynamics simulations suggest that alternating access is mostly provided by the angular movement than the translation, which is more in line with a rocking bundle mechanism (Coincon *et al.*, 2016; Masrati *et al.*, 2020).

An overview of protein families that are currently suggested to work by an elevator mechanism is given in **Table 1**. For each family, only one representative protein is provided as an example, although for some of them structures in multiple conformations have been solved. Common characteristics and differences among the diverse elevator protein families will be discussed in the following.

Table 1: List of protein families that have been suggested to work by an elevator mechanism. For each protein family (according to the TCDB classification system), only one representative member is given as an example together with the PDB entry of the first related structure solved. A more detailed list of elevator proteins and PDB entries of different conformations can be found elsewhere (Garaeva and Slotboom, 2020). TCDB, Transporter Classification Database, HP, helical hairpin.

Family	TCDB	SLC	Transport mode	Inverted repeat	Example	PDB
Amino acid transporter	2.A.23	1	symport, antiport	3+3 and HP-1 + HP-1	Glt$_{Ph}$	1XFH
Bicarbonate transporter	2.A.31	4	symport, antiport	7+7	AE1	4YZF
Na$^+$/H$^+$ exchanger (NhaA family)	2.A.33	9	antiport	5+5	NhaA	4AU5
Cation-Proton Antiporter-1 (CPA1)	2.A.36	9	antiport	5+5	MjNhaP1	4CZB
Cation-Proton Antiporter-1 (CPA2)	2.A.37	9	antiport	5+5	NapA	4BWZ
Bile acid transporter	2.A.28	10	symport	5+5	nmASBT	3ZUX
Na$^+$-citrate cotransporter (NaCT)	2.A.47	13	symport	3-HP-2+3-HP-2	VcINDY	4F35
Nucleobase-ascorbate transporter	2.A.40	23	symport	7+7	UraA	3QE7
Anion transporter	2.A.53	26	symport, antiport	7+7	SLC26Dg	5DA0
Concentrative nucleoside transporter (CNT)	2.A.41	28	symport	1-HP-2+1-HP-2	vcCNT	4PD6
PTS (Glc) family	4.A.1		facilitated diffusion	5+5	bcMalT	5IWS
PTS (Lac) family	4.A.3		facilitated diffusion	5+5	bcChbC	3QNQ
PTS (L-asc)	4.A.7		facilitated diffusion	5+5	UlaA	4RP8
p-Aminobenzoyl glutamate transporter (AbgT)	2.A.68		symport	2-HP-2+2-HP-2	MtrF	4R1I
2-Hydroxycarboxylate Transporter (2-HCT)	2.A.24		symport, antiport	5+5	KpCitS	5XAS
Membrane oxidoreductase	5.A.1		-	3+3	CcdA	5VKV
Energy-coupling factor (ECF) transporter (S-and T-component)	diverse		-	-	RibU	3P5N

1.4. Overall characteristics of elevator proteins

1.4.1. Elevator proteins contain inverted repeats and unusual structural elements

As already mentioned above, proteins that facilitate alternating access by elevator-like motions fold into two distinct helix bundles that comprise the substrate binding and scaffold domains. Furthermore, for almost all of them, inverted-topology repeats can be identified, as indicated in **Table 1** and **Figure 2**. These inverted repeats were suggested to be the structural basis for asymmetric states that would enable alternating access to the substrate binding site (Forrest, 2015). The asymmetry in the inverted repeats creates the pathway for the substrate, opening either to the inward or the outward compartment. It can therefore be assumed that each structural repeat in a given protein state will adopt the conformation of the other one in the alternative state, and *vice versa*. This concept was used in a so-called repeat-swap homology modeling to predict so far structurally inaccessible conformations of elevator proteins, for example for vcCNT (a SLC28 family member from *Vibrio cholerae*), vcINDY (a sodium-coupled citrate transporter from the same organism), or AE1 (SLC4A1, the major bicarbonate transporter from human red blood cells) (Vergara-Jaque *et al.*, 2015; Mulligan *et al.*, 2016; Ficici *et al.*, 2017). The predicted conformations so far were in good agreement with existing structures and other experimental data.

Within the elevator proteins, the structural composition of these inverted repeats is diverse. In all cases, both structural repeats contribute to the substrate binding as well as the scaffold domain. In this light, the evolvement of this mechanism appears enigmatic, as structural repeats are thought to arise from duplication of subunits that are intrinsically functional (Forrest, 2015). At the moment, it is also unclear how these complicated, intertwined proteins adopt their correct fold in the membrane during their biogenesis. The by far best studied elevator protein is Glt$_{Ph}$ from *Pyrococcus horikoshii* that belongs to the SLC1 family of amino acid transporters. It is built from two structural repeats, the first composed of a three helix bundle ("3+3"), the second containing a hairpin structure and one helix ("HP-1 + HP-1") (**Figure 2A**). Structural repeats containing helical hairpins but involving differing numbers of helices can be also identified in VcINDY, MtrF, nwCNT and KpCitS (**Figure 2 B-E**). In all cases, the helical hairpins contribute to substrate binding and might, as already mentioned, act as gating elements that move during the transport cycle. NhaA, in contrast, does not contain helical hairpins but, unusually, a pair of disrupted helices (**Figure 2F**). These two helices, that are oriented in opposite directions, were shown to contain many key residues for substrate binding, but their precise role for substrate translocation is unknown at the moment (Maes *et al.*, 2012; Padan *et al.*, 2015; Winkelmann *et al.*, 2020). A similar fold was found for the ASBT bile acid transporters (SLC10 family), although the sequence similarity to Na$^+$/H$^+$-antiporters is low (Hu *et al.*, 2012; Zhou *et al.*, 2014).

A Glt_{Ph}: 3+3 and HP-1 + HP-1
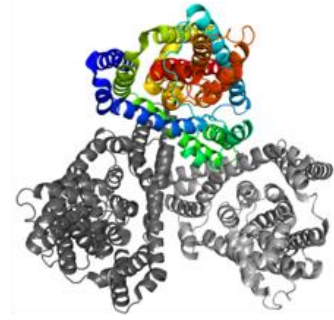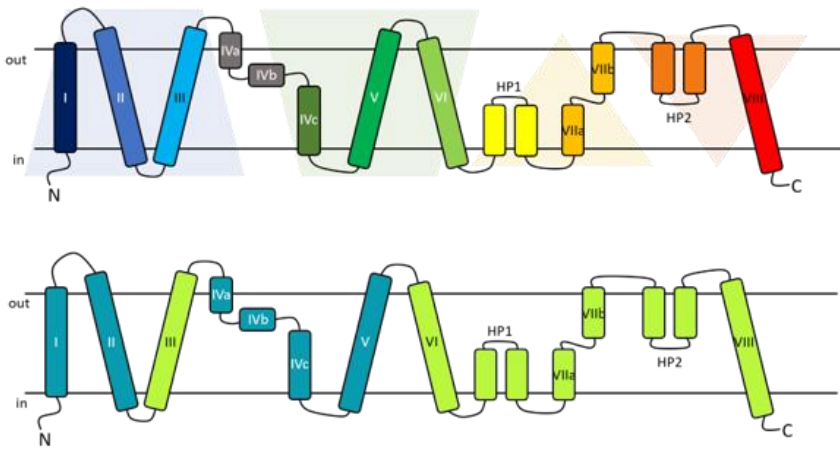
B   VcINDY: 3-HP-2 + 3-HP-2
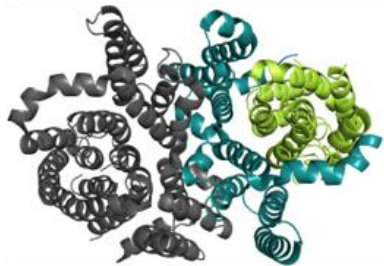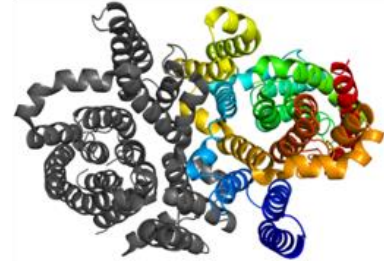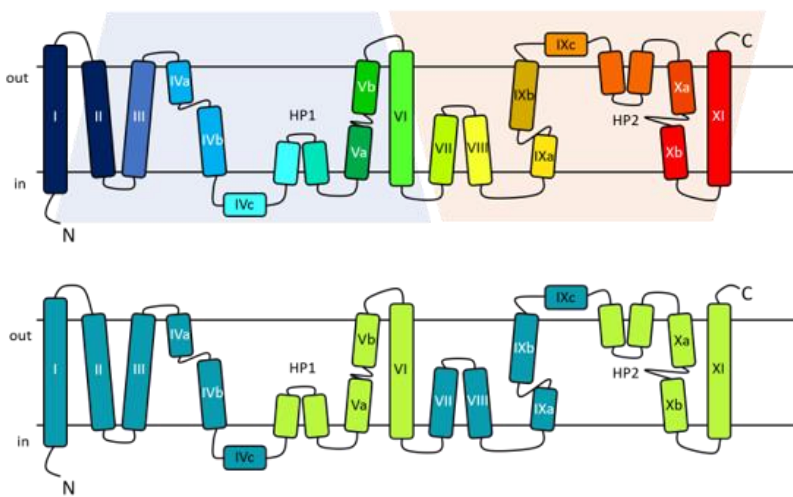
C   MtrF: 2-HP-2 + 2-HP-2
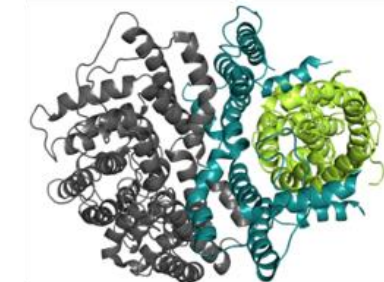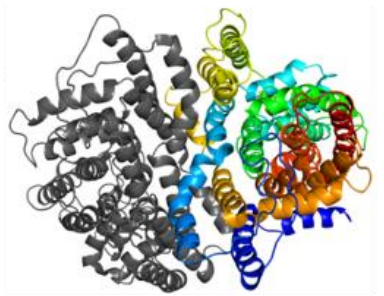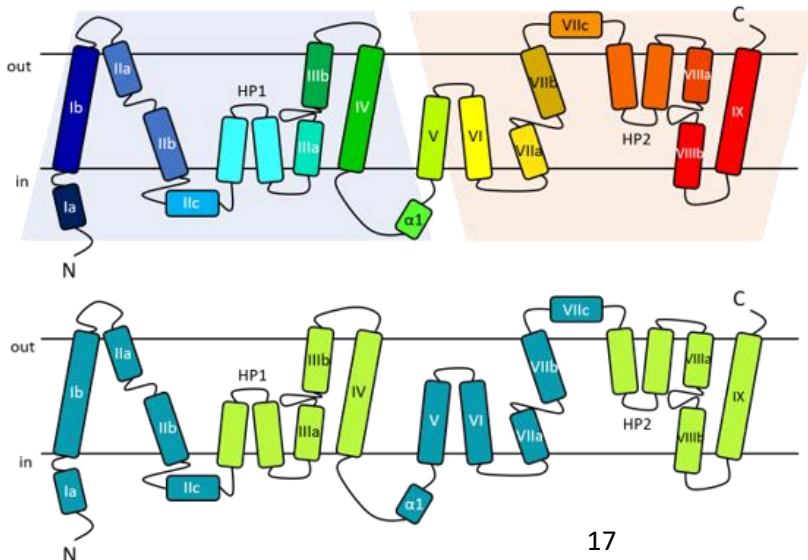
D   nwCNT: 1-HP-2 + 1-HP-2
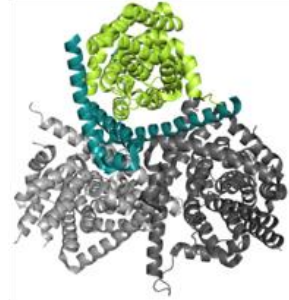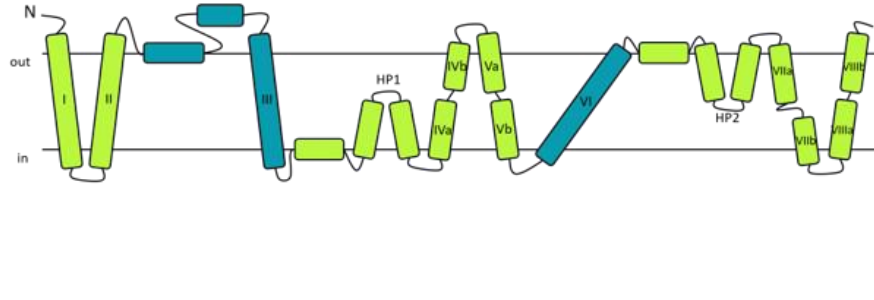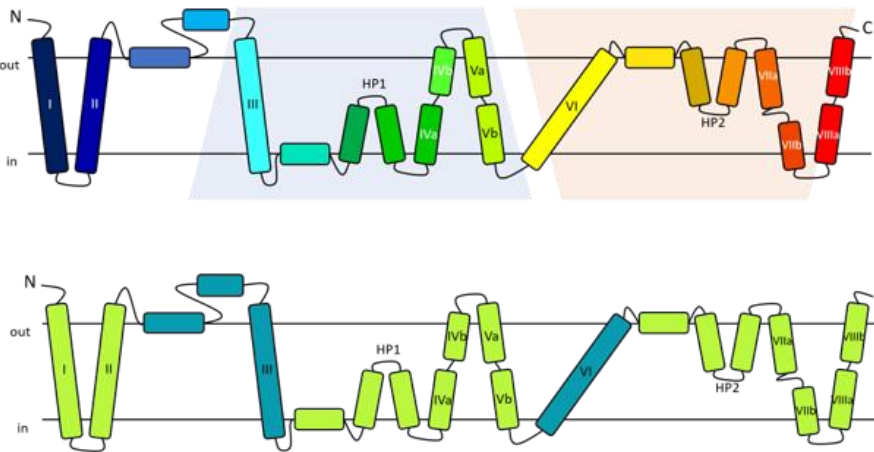
E   KpCitS: 5+5
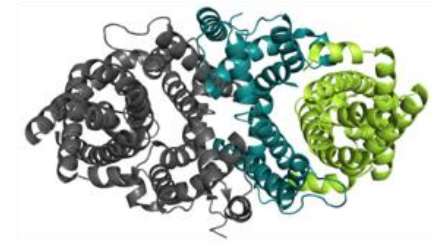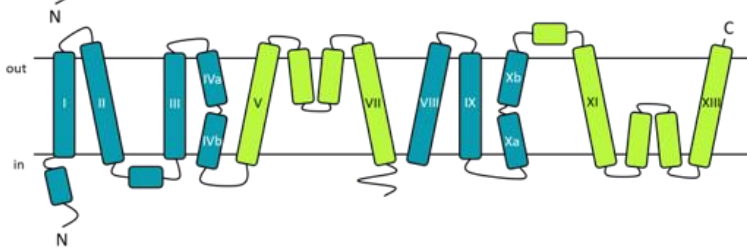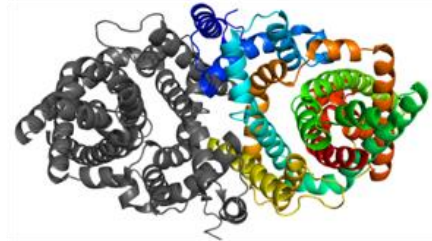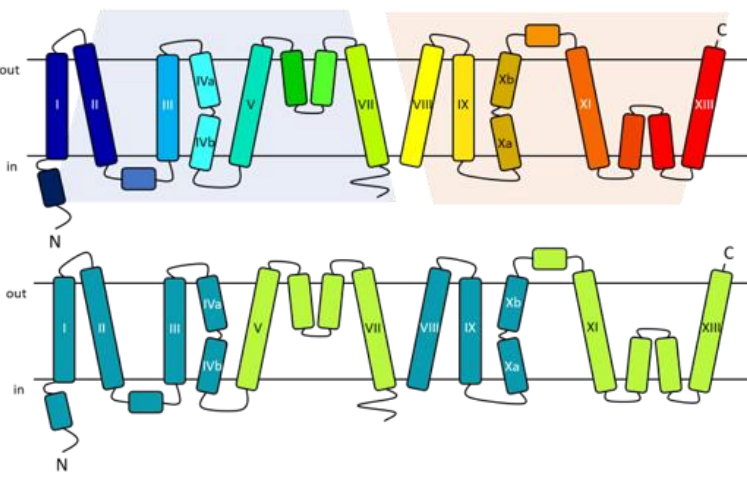
F   NhaA: 5+5
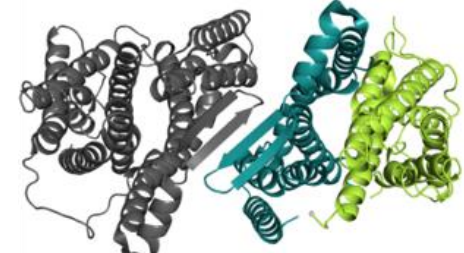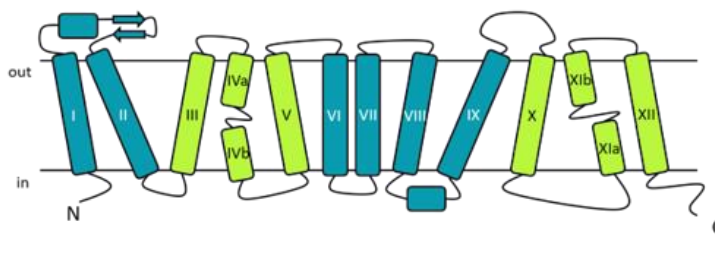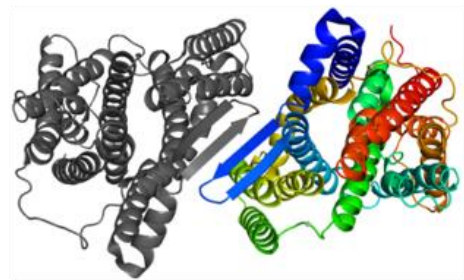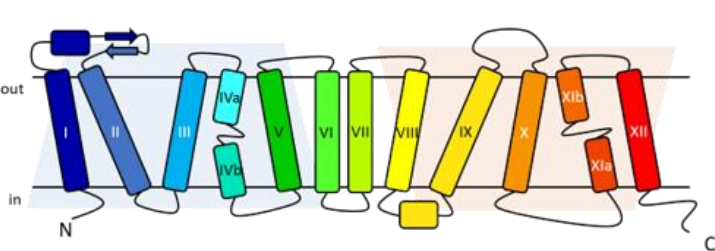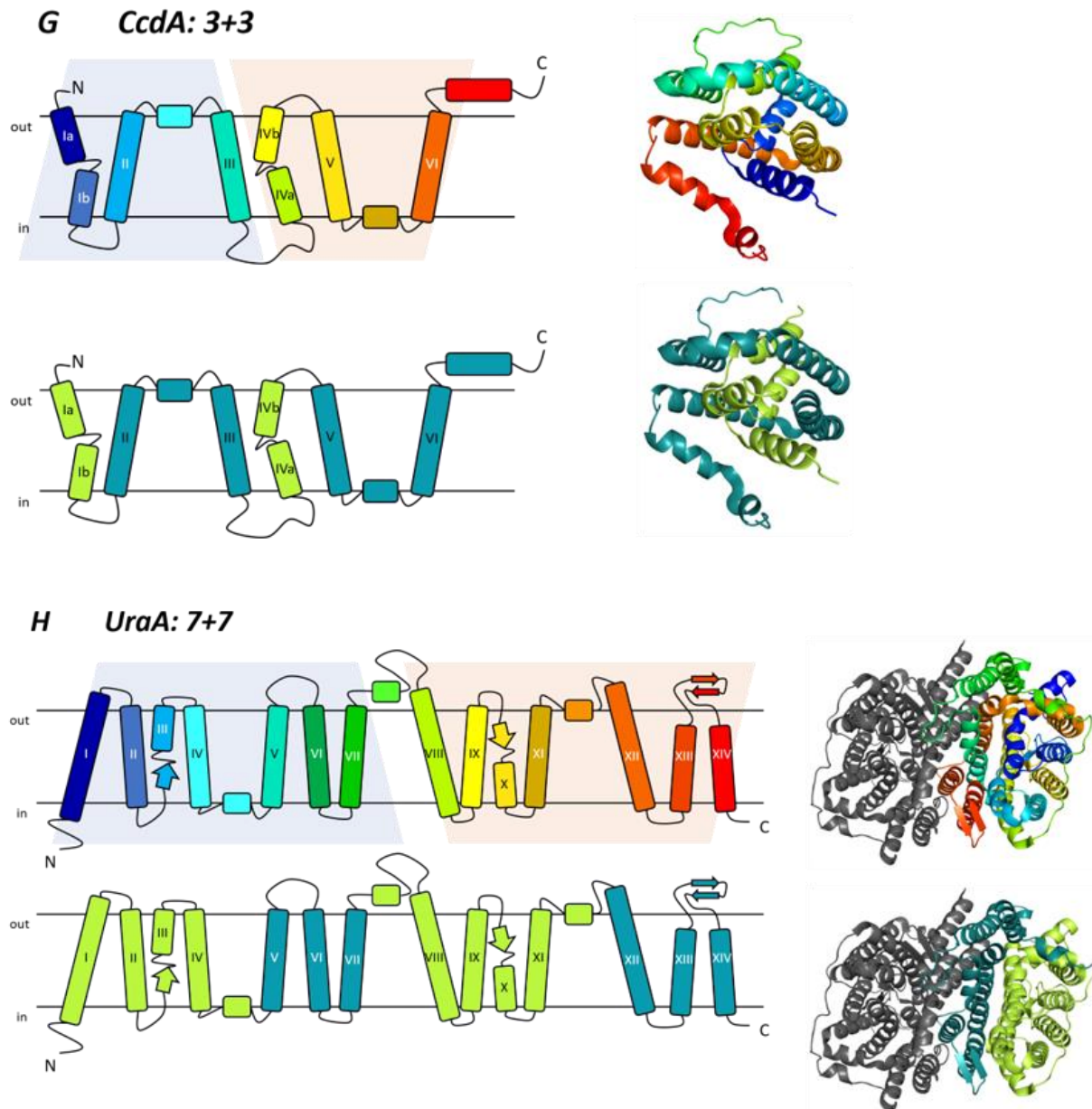
Figure 2: Topology and domain organization of elevator proteins. Proteins that have been suggested to work by an elevator mechanism are diverse in their architecture. For each example, membrane topology in rainbow-coloring is given in the upper panel, coloring according to domain architecture in the lower panel, with elements contributing to the scaffold domain in dark teal and the ones comprising the transport domain in light green. On the right, proteins are shown in top view following the same color coding. The following PDB entries were used to generate the figures: (A) GltPh - 1XFH; (B) VcINDY - 4F35; (C) MtrF - 4R1I; (D) nwCNT: - 5L26; (E) KpCitS - 5XAR; (F) NhaA - 4AU5; (G) CcdA - 5VKV; (H) UraA - 5XLS. Inverted repeats are indicated in the upper panel. Transmembrane helices are numbered, N- and C-termini are indicated. HP, helical hairpin.

More recently, an elevator mechanism was also proposed for CcdA from *Thermus thermophilus*, based on a structure in an outward-open state and a model of the inward-facing state obtained from an repeat-swap homology approach (Zhou and Bushweller, 2018). This protein is confined to a 3+3 inverted repeat, thus representing the smallest elevator protein identified so far (**Figure 2G**). One of the helices of each repeat is discontinuous. Interestingly, the scaffold domain in this case was

suggested to have a novel, "O" shaped architecture that surrounds the discontinuous helices, which were assigned as "transport helices" as they solely appear to move during substrate translocation. However, so far, only a monomeric structure is available for this transporter family although biochemical data indicates that these proteins are dimeric (Bocian-Ostrzycka *et al.*, 2017). It is difficult to clearly identify the structural elements of the scaffold domain without the structure of a higher oligomer. Lastly, the elevator mechanism was also suggested for the so-called seven transmembrane segment inverted repeat family (7TMIR), of which the first structure was solved for the *E. coli* uracil transporter UraA (Lu *et al.*, 2011). These proteins do not contain a pair of discontinuous helices, but two short helices that each are preceded by a short ß-strand and an extended, unwound fragment (**Figure 2H**). Due to the inverted structural repeat, these two transmembrane segments cross each other in the mid of the membrane plane. This structural feature resembles the discontinuous helices found in other elevator proteins, but in this, one of the half helices is replaced by a ß-strand. The ß-strands are centered at the substrate binding site and contain key residues for substrate binding (see below).

1.4.2. Elevator proteins are oligomers

Strikingly, all proteins that have been suggested to work by an elevator mechanism so far hold higher oligomeric states. Members of the SLC1 and SLC28 families form trimers, all others are dimeric (**Figure 2**). In all cases, oligomerization is mediated by the scaffold domain, which is therefore also referred to as oligomerization domain. As defining feature of elevator proteins, substrate binding is, more or less exclusively, confined to the transport domain, which does not contribute to the oligomerization interface. As a consequence, every protomer harbors its own substrate binding and translocation pathway. From this perspective, oligomerization appears not to be a prerequisite for the elevator motion. Whether the monomers of elevator proteins are functional has been addressed only in a few cases experimentally. Monomerized mutants of UraA were demonstrated to bind uracil, but were incapable of transportation (Yu *et al.*, 2017). In contrast, monomers of the Na^+/H^+ antiporter NhaA showed wild type activity under most conditions tested. Interestingly, however, the NhaA dimer was shown to transport more efficiently than the monomeric version under stress conditions (high salt and high pH) (Rimon *et al.*, 2007).

Whether oligomerization has a general, unified relevance for the elevator mechanism is not clear to date. For the SLC26 family, dimerization has been proposed to provide a more rigid scaffold that enables an efficient movement of the substrate binding against the oligomerization domain (Chang *et al.*, 2019). For the SLC23 transporter UapA from *Aspergillus nidulans*, it was shown that dimerization is necessary for correct trafficking to the membrane (Martzoukou *et al.*, 2015a). Recently, it has been

suggested that oligomerization of elevator proteins might be the structural basis for functional cooperativity between the protomers (Holzhüter and Geertsma, 2020). Also in this regard, differences between the various elevator protein families appear to exist. For example, for Glt$_{Ph}$, independent movement of the single subunits in the trimer has been demonstrated using high-speed AFM and smFRET (Erkens *et al.*, 2013; Ruan *et al.*, 2017). In contrast, studies on the SLC23 transporters UraA and UapA suggested cooperativity, albeit of different types (positive for UraA, negative for UapA) (Alguel *et al.*, 2016; Yu *et al.*, 2017). To achieve functional crosstalk in oligomers, information on substrate binding and/or transport has to be transferred from one protomer to the other. For soluble proteins, the mechanistic basis for cooperativity lies in substrate binding-induced conformational changes at the interface of the protomers within the oligomer. These changes are either directly transmitted to adjacent protomers (**Figure 3A**) or indirectly by a change in the oligomeric state (**Figure 3B**). For elevator proteins, these mechanisms for cooperativity are unlikely to occur. The oligomerization interface in elevator proteins is formed exclusively by the scaffold domains of each protomer. As one characteristic of the elevator mechanism, the scaffold domains remain relatively static during the transport cycle. This suggests that cooperativity between protomers is not conferred *via* pronounced changes at the oligomerization interface. Additionally, there are so far no indications that their oligomeric state would change during the transport cycle. We proposed the bilayer deformation caused by the large movement of the substrate binding domain with respect to the scaffold domain as a common basis for cooperativity in elevator proteins (**Figure 3C**) (Holzhüter and Geertsma, 2020). The elevator movement would place a considerable part of the hydrophobic surface of the protein outside the membrane, which is energetically unfavorable. This led to the hypothesis that the movement of the substrate binding domain would rather be accompanied by membrane deformations that keep the protein in the hydrophobic environment. Recent MD simulations on Glt$_{Ph}$ and VcINDY and multiple structures of the Glt$_{Ph}$ homolog Glt$_{Tk}$ in nanodiscs support this idea (Zhou *et al.*, 2019; Arkhipova *et al.*, 2020). More experimental data is necessary in order to verify that these membrane deformations contribute, or cause, the observed cooperative behavior in many elevator proteins.
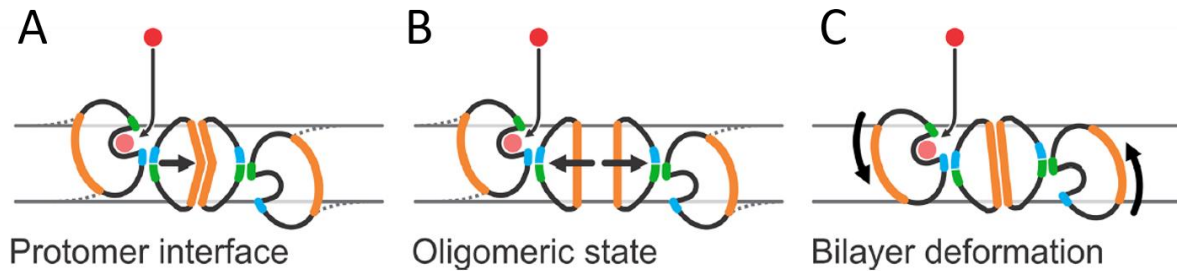
Figure 3: Options for achieving cooperative behavior in oligomeric elevator proteins. To transfer information between protomers, conformational changes caused by substrate binding to one protomer lead to conformational changes in the other protomer that influence its transport activity. Based on observations made on soluble proteins, different possible mechanisms might provide the basis for cooperativity within membrane proteins. (A) Substrate binding to one protomer leads to conformational changes in the interface of the protomers. (B) Substrate binding induces a change in the oligomeric state. (C) Substrate translocation in one protomer causes a deformation of the membrane bilayer that influences that other protomer. Adapted from Holzhüter and Geertsma, 2020, with minor changes.

1.4.3. Elevator proteins employ different transport modes

For several proposed elevator proteins, their transport modes and its molecular basis have been studied in detail. For example, the EIIC membrane components of the bacterial phosphotransferase (PTS) system mediate substrate translocation by facilitative diffusion. These proteins transport various sugars and other compounds across the membrane that are immediately phosphorylated after translocation. Substrate transport thus is fueled by the concentration gradient only, which is maintained as the phosphorylated sugars are no substrates for the transporters. This mechanism is also referred to as group translocation (Deutscher *et al.*, 2006; Vastermark *et al.*, 2016; Luo *et al.*, 2018). For other elevator proteins, proton- and sodium-coupled symport as well as exchange or uniport have been described (**Table 1**). Different modes of transport also exist within one protein family, for example for the SLC26 transporters: most of the ten members encoded in the human genome are exchangers for a variety of anions (chloride, iodide, sulfate, bicarbonate, oxalate, malate), while three of them (SLC26A7, SLC26A9 and SLC26A11) additionally show chloride channel characteristics (Rapp *et al.*, 2017). More recently, it has been demonstrated that murine SLC26a9 operates as a fast uniporter rather than as a true ion channel (Walter *et al.*, 2019). SLC26A5, also known as prestin, is an absolute exception as it works as a motor protein in the outer hair cell. Chloride binding leads to a conformational change in prestin following elongation or contraction of the whole cell, depending on the membrane potential (Dallos and Fakler, 2002). The SLC4 bicarbonate transporter family contains both electrogenic and electroneutral Na^+-dependent and -independent Cl^-/HCO_3^- exchangers, as well as borate transporters in plants (Choi, 2012). Interestingly, the human Na^+-coupled bicarbonate transporter NBCe1 could be converted into an uncoupled Cl^-/HCO_3^- exchanger by mutating a small set of amino acids into the ones found in the exchanger AE1 (Huynh *et al.*, 2018). In addition,

the E699Q mutant at the substrate binding site of the murine homolog of AE1 abolished proton transport, resulting in electrogenic SO_4^{2-}/Cl^- exchange instead of proton-coupled electroneutral exchange of H^+/SO_4^{2-} for Cl^- (Chernova *et al.*, 1997). It appears that only small changes are sufficient to alter the transport mode for these proteins.

1.5. The structural class of 7TMIR carriers

The first structure of UraA, a SLC23 uracil transporter from *E .coli*, discovered a novel fold built from two inverted repeats of seven transmembrane helices each (**Figure 2H**) (Lu *et al.*, 2011). Later on, structures of a prokaryotic SLC26 homolog (SLC26Dg) and of human anion exchanger 1 (AE1/Band 3/ SLC4A1) revealed that their transmembrane domains share the same architecture (Arakawa *et al.*, 2015; Geertsma *et al.*, 2015). The SLC4, SLC23 and SLC26 families were therefore grouped into a new structural class called seven transmembrane segment inverted repeat (7TMIR) carriers. Based on comparison of the inward-facing conformation of SLC26Dg and the outward-open structure of AE1, they were suggested to employ an elevator mechanism (Chang and Geertsma, 2017).
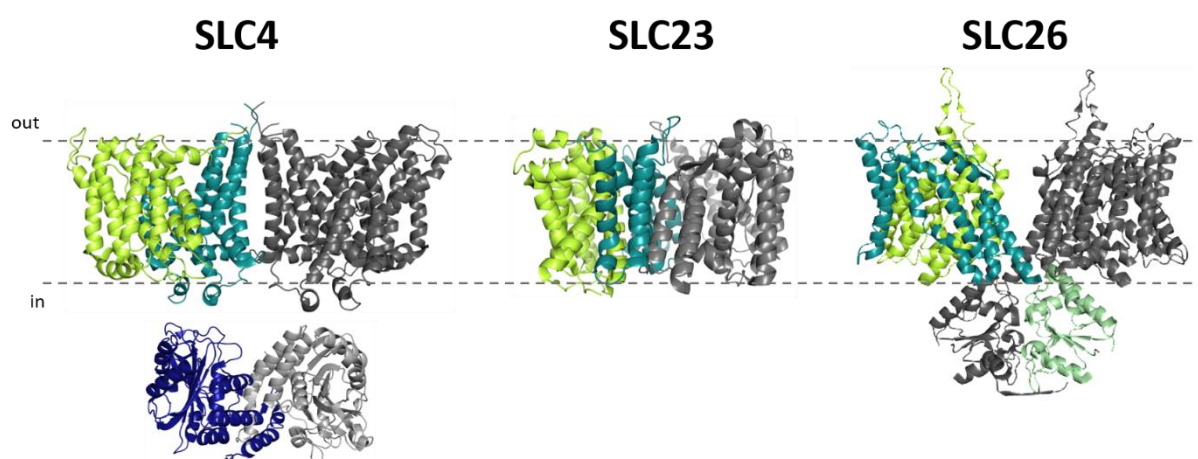
Figure 4: Domain architecture of 7TMIR carriers. The transmembrane domains of the SLC4 (left, AE1 – PDB: 4YZF), SLC23 (middle, UraA – PDB: 5XLS) and SLC26 (right, SLC26A9 – PDB: 6RTC) proteins share the same architecture of two inverted repeats of seven helices each, which fold into the scaffold and transport domains typical for elevator proteins (shown in dark teal and green, respectively). In the SLC4 family, an additional soluble N-terminal domain is found, which has been crystallized isolated as a dimer. One protomer is shown in dark blue, one in light grey (PDB: 1HYN). The SLC26 family holds an additional C-terminal soluble domain, the so-called STAS domain, shown in light green.

Whereas AE1 crystallized as a homodimer, both UraA and SLC26Dg structures were obtained from monomers. More recent structures as well as biochemical data confirmed that also the SLC23 and SLC26 proteins form dimers in lipidic environment (Alguel *et al.*, 2016; Yu *et al.*, 2017; Chang *et al.*, 2019; Walter *et al.*, 2019; Chi *et al.*, 2020). In contrast to the SLC23 family, that is confined to the 7TMIR transmembrane domain, the SLC4 and SLC26 families contain additional, soluble domains (**Figure 4**).

Introduction

In SLC4 family members, a large N-terminal domain is found that was removed in AE1 in order to obtain the crystal structure (residues 1-360 of total 911) (Arakawa *et al.*, 2015). In AE1, both soluble and transmembrane domain on their own are functional. The N-terminal cytosolic domain primarily acts as an anchor to the cytoskeleton and interacts with a variety of other proteins involved in different cellular processes, such as aldolase, phosphofructokinase, glyceraldehyde 3-phosphate dehydrogenase, and hemoglobin (Zhang *et al.*, 2000). The SLC26 family contains a C-terminal, cytosolic STAS domain, which is described in more detail below (section 1.7).
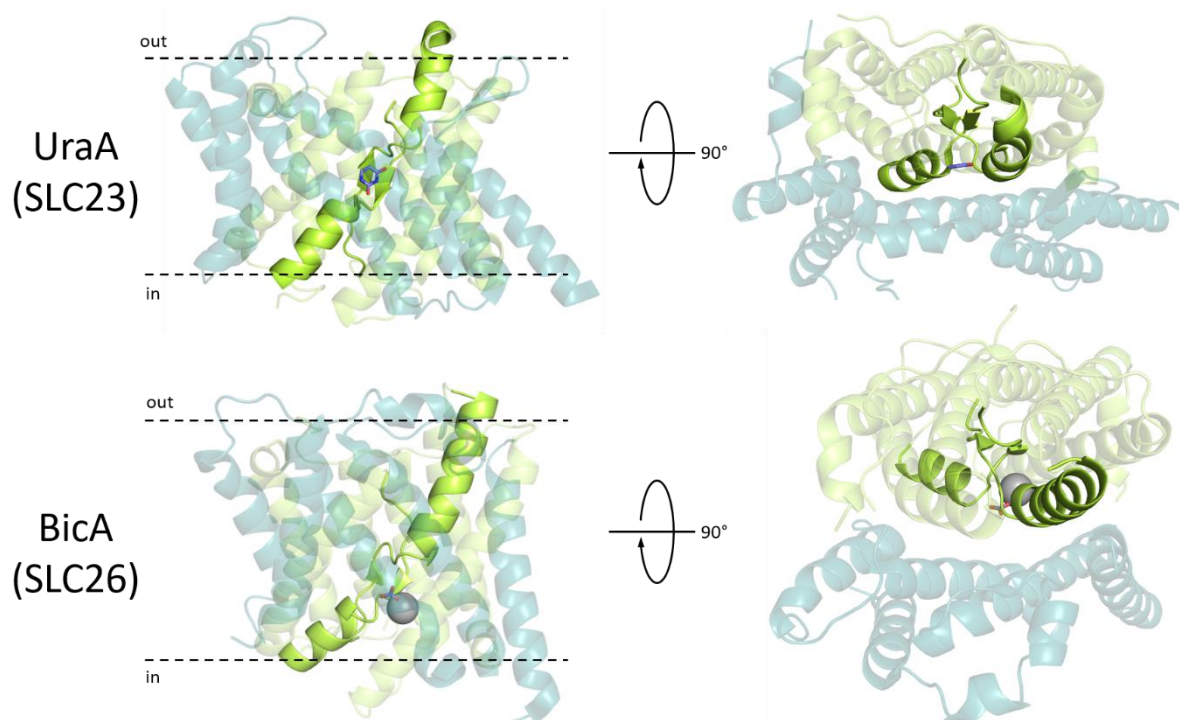
Figure 5: Substrate binding in the SLC23 and SLC26 families. Substrate bound structures for 7TMIR carriers. As an example for the SLC23 family, UraA is shown, and BicA for the SLC26 family. The proteins are shown in side view on the left and in top view on the right, with the scaffold and transport domains colored in dark teal and green in transparent, with the exception of TM3 and TM10 and the ß-sheet at the substrate binding site, which are shown in opaque. Ligands are shown in sticks (uracil for UraA, bicarbonate for BicA) and the sodium in the BicA structure as a grey sphere.

The transmembrane parts of the 7TMIR families holds all characteristics commonly found in elevator proteins that are described in section 1.4: they contain an inverted structural repeat, they fold into two domains of which one is mediating dimerization and the other binding the substrate, and the substrate binding site is located at an unusual structural element, in this case a ß-strand followed by a short helix that only halfway spans the membrane. These observations are based on crystal structures of 7TMIR carriers that have been solved in presence of substrate (UraA and UapA for the SLC23 family, SLC26A9 and BicA for the SLC26 family) (**Figure 5**) (Lu *et al.*, 2011; Alguel *et al.*, 2016; Walter *et al.*, 2019; Wang *et al.*, 2019). For the SLC4 family, to date no structure of a substrate-bound state is

available, but a large amount of biochemical data indicates that the substrate binding site is located at a similar position as in the other two 7TMIR families (Jennings, 2005; Alper, 2009; Arakawa *et al.*, 2015; Huynh *et al.*, 2018).

While for the SLC23 and SLC26 families only structures in inward-open or inward-occluded states are available, for the SLC4 both inward-open (from *Saccharomyces mikatae* Bor1p) and outward-open (from *Homo sapiens* AE1 and NBCe1) conformations were structurally resolved. Additionally, an outward-occluded state (from *Arabidopsis thaliana* Bor1) was reported (Arakawa *et al.*, 2015; Thurtle-Schmidt and Stroud, 2016; Coudray *et al.*, 2017; Huynh *et al.*, 2018). However, as already mentioned, comparisons of the different conformations led to different conclusions concerning the underlying transport mechanism. Based on the structure of Bor1 in an occluded state, an elevator mechanism was proposed, as for both comparisons to the outward-open conformation of AE1 as well as inward-open structures of the SLC23 and SLC26 families a vertical displacement of the substrate binding domain by 5 Å was seen, meaning 10 Å in total for the movement of inward- to outward-open or *vice versa* (Thurtle-Schmidt and Stroud, 2016). In contrast, Coudray *et al.* compared the inward-open structure of Bor1 with the outward-open structure of AE1 and additionally modelled the outward-open conformation of Bor1. In both cases, only small movements of both domains with respect to the membrane were observed, which is more in line with a rocking-bundle mechanism (Coudray *et al.*, 2017). It should also be noted that AE1 was crystallized in presence of an inhibitor; it was suggested that the resulting open-outward conformation might be more occluded than an unbound outward-open state (Ficici *et al.*, 2017). In summary, although the members of the SLC4, SLC23 and SLC26 families show all characteristics that are found in other elevator proteins, formal proof for elevator-like motions as basis for achieving alternating access is still missing to date.

1.6. The SLC23 family of nucleobase and ascorbate transporters

The members of the SLC23 family are transporters of nucleobases in all kingdoms of life with the exception of higher eukaryotes, in which they transport ascorbate instead. Based on phylogenetic analysis, they are subdivided into two COGs (Cluster of Orthologous Genes) (Galperin *et al.*, 2018). The different clusters have been suggested to contain transporters with distinct substrate specificities: One containing transporters for adenine or guanine (COG2252) and one with subgroups specific for uracil, ascorbate, or xanthine and uric acid (COG2233) (**Figure 6**) (Frillingos, 2012).
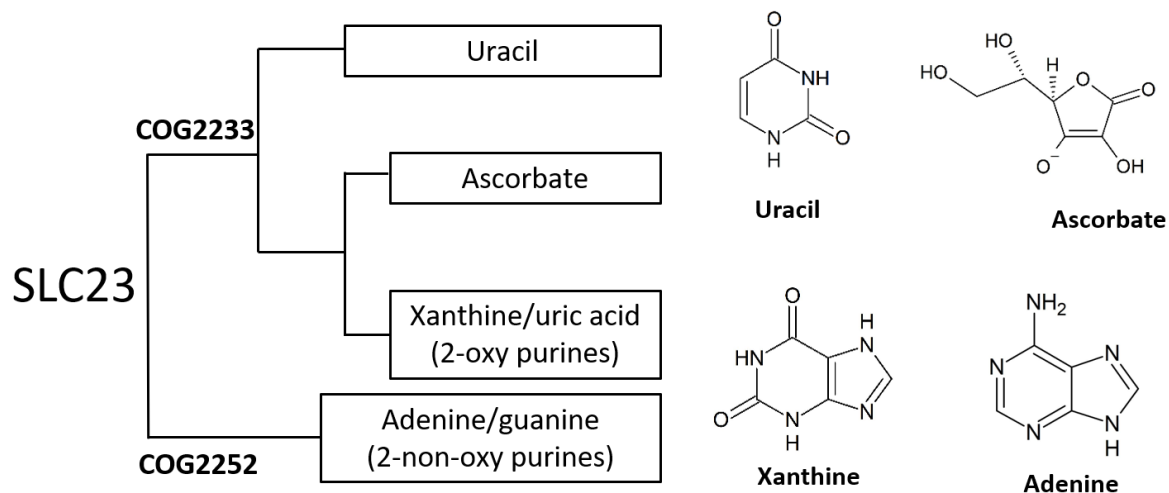
Figure 6: Substrate specificities of SLC23 transporters. SLC23 transporters phylogenetically subdivide into two COGs (Cluster of Orthologous Genes), one (COG2252) specific for adenine and guanine, in the other (COG2233) transporters specific for uracil, ascorbate, or xanthine/uric acid are found.

The assignment of substrate specificities to the different SLC23 subgroups that are transporting nucleobases is based mostly on the functionally well-characterized homologs from *E. coli.* In *E. coli*, ten SLC23 homologs are found from which most are reported as very specific for their respective substrates. The functionally best characterized SLC23 transporters are XanQ, one of two xanthine transporters in *E. coli*, and UapA from *Aspergillus nidulans* that transports both xanthine and uric acid (Diallinas *et al.*, 1995; Papakostas and Frillingos, 2012). In general, SLC23 transporters exhibit high affinities for their substrates. The dissociation constant (K_D) of uracil binding to UraA was determined as $0.41 \pm 0.07\,\mu M$ using scintillation proximity assays (SPA) and the K_m for uracil transport as $0.49 \pm 0.07\,\mu M$ with a v_{max} of 59 ± 3 nmol min^{-1} mg^{-1} in cell-based uptake assays using radioactively labeled substrate (Lu *et al.*, 2011). The SLC23 nucleobase transporters are thought to being proton-coupled symporters, based on comparison of transport rates in absence and presence of a protonophore and MD simulations (Goudela *et al.*, 2005; Karatza and Frillingos, 2005; Kosti *et al.*, 2012; Papakostas and Frillingos, 2012; Niopek-Witz *et al.*, 2014; Yu *et al.*, 2017; Botou *et al.*, 2018; Yang *et al.*, 2018).
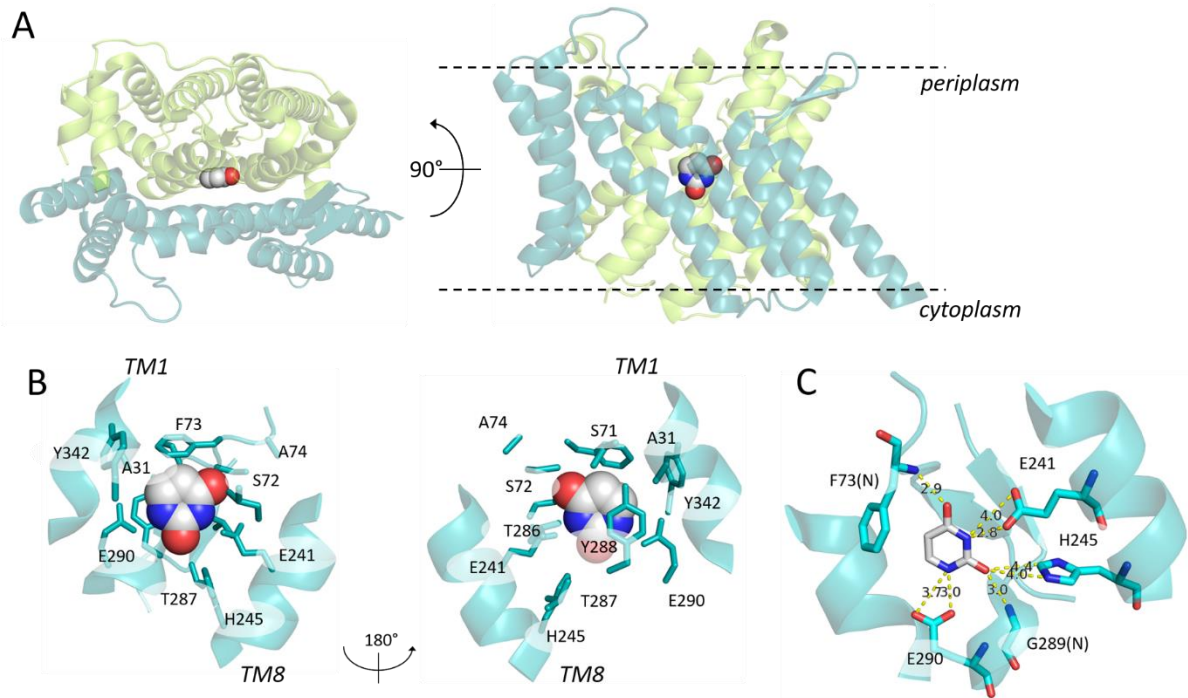
Figure 7: The substrate binding site of UraA. (A) Uracil (depicted in spheres) is located at the interface between the substrate binding and scaffold domains, shown in light green and teal, respectively. (B) Front and back view of the UraA substrate binding site. Residues found within 5 Å around the ligand are shown in sticks and annotated. (C) Distances between opposite charges close to the ligand. Distances were measured using the built-in PyMOL function. Distances ≤3.5 Å might hint to hydrogen bond formation (Bondar and White, 2013). All figures are based on the dimer structure of UraA (5XLS).

For SLC23 transporters, two proteins have been structurally characterized: *E. coli* UraA (in inward-open, monomeric, as well later on in a dimeric state, inward-occluded) and UapA from *Aspergillus nidulans* (Lu *et al.*, 2011; Alguel *et al.*, 2016; Yu *et al.*, 2017). All three structures were solved on substrate-bound protein. As typical for elevator proteins, the substrate is located at the interface between the substrate binding domain and the scaffold domain (**Figure 7A**). Uracil is coordinated by a variety of polar and non-polar interactions at two ß-strands each followed by a short helix (TM3 and TM10). The residues that are found near the substrate are, with the exception of I132 on TM5 and Y342 on TM12, exclusively located on the substrate binding/transport domain. In particular, these are A31 (TM1), S71, S72, F73 and A74 that are part of the ß-strand before TM3, E241 and H245 of TM8, as well as T286, T287, Y288, G289 and E290 of the second ß-strand before TM10 (**Figure 7B**). Regarding the two available conformations, only subtle differences exist regarding which residues are found close to uracil; a shift of helix 5, moving from the inward-open to the occluded state, places I132 instead of L136 in 5 Å distance to the substrate. Three chargeable residues in the vicinity of the ligand (E241, H245 and E290) suggest polar or electrostatic interactions to be involved in substrate coordination (**Figure 7C**). It was shown that replacing either of those with alanine severely impairs the transporter's function (Lu *et al.*, 2011; Yu *et al.*, 2017). These residues have been suggested to contribute to

27

proton-coupling, though their precise role has remained unclear so far.

In humans, four genes of SLC23 homologs are found, of which two are pseudogenes (Yamamoto *et al.*, 2010a; Bürzle *et al.*, 2013). Interestingly, the SLC23 homologs from higher eukaryotes do not transport nucleobases, but ascorbate. SVCT1 (SLC23A1) is mostly located in epithelial cells in distinct organs such as kidney, intestine, liver, lung and skin (Wohlrab *et al.*, 2017), whereas SVCT2 (SLC23A2) is expressed in almost all tissues throughout the body (May, 2011). SVCT2 knockout mice die shortly after birth (Sotiriou *et al.*, 2002). Loss of SVCT1 function in mice is less severe and can be compensated by increased intake of ascorbate with the diet (Corpe *et al.*, 2010). The oxidized form of ascorbate, dehydroascorbic acid, is also transported in the intestine by GLUT2 and GLUT8, albeit with a low affinity (Corpe *et al.*, 2013). In contrast to the proposed proton-coupling in the SLC23 nucleobase transporters, ascorbate uptake by SLC23 homologs was shown to depend on a sodium gradient instead (Tsukaguchi *et al.*, 1999b; Godoy *et al.*, 2007). No high-resolution structure is available to date for an ascorbate transporter of the SLC23 family, and the location of the sodium binding site(s) is unknown. For SVCT2, two sodium ions are translocated per ascorbate. Sodium binding seems to convert the transporter from a low-affinity to a high-affinity form and accelerates the transport rate cooperatively at low concentrations of ascorbate (Godoy *et al.*, 2007). SVCT1 in contrast is reported to mediate sodium-dependent electrogenic transport, also with a two-to-one stoichiometry of sodium to ascorbate (Wang *et al.*, 2000).

1.7. The SLC26 family of multifunctional anion transporters

The SLC26 family contains multifunctional exchangers and channels for a number of various anions such as iodide, chloride, bicarbonate, sulfate or different dicarboxylates (Alper and Sharma, 2013). In humans, ten homologs are found that are expressed in tissues throughout the whole body. Mutations in SLC26 transporters have been linked to a number of diseases, for example congenital chloride diarrhea (SLC26A3/DRA) or deafness (SCL26A4,/pendrin, SLC26A5/prestin) (Rapp *et al.*, 2017). Functionally well-characterized SLC26 transporters are *E. coli* DauA that transports a number of C_4-dicarboxylates with a millimolar affinity, several sulfate transporters (Sultr1;1, Sultr1;2 and Sultr1;3) from *Arabidopsis thaliana*, Sul1p and Sul2p from *Saccharomyces cerevisiae*, which also transport sulfate, as well as the fumarate/proton symporter SLC26Dg from *Deinococcus geothermalis* (Cherest *et al.*, 1997; Hawkesford, 2003; Karinou *et al.*, 2013; Geertsma *et al.*, 2015). As described above, SLC26 transporters exhibit different modes of transport such as exchange, symport and channel activities (section 1.4.3).

The transmembrane domain of SLC26 transporters holds the 7TMIR fold as does the SLC23 family, although their sequence identity is low (<15%). The first crystal structure of a SLC26 transporter was

solved on a prokaryotic fumarate-proton symporter from *Deinococcus geothermalis*, termed SLC26Dg (Geertsma *et al.*, 2015). Since then, three other structures became available: one from BicA, a low-affinity bicarbonate transporter from cyanobacteria, and one structure each from the murine and human SLC26A9 homologs that both show chloride channel-like characteristics but were also suggested to being fast uniporters instead (Walter *et al.*, 2019; Wang *et al.*, 2019; Chi *et al.*, 2020). SLC26Dg crystallized as a monomer, but a later dimer model based on EPR constraints as well as the more recent structures revealed the overall dimeric state of the SLC26 family (Chang *et al.*, 2019). For some SLC26 transporters, however, also tetramers have been observed (Zheng *et al.*, 2005; Hallworth *et al.*, 2013). BicA was co-crystallized with bound bicarbonate and sodium (**Figure 5**), suggesting that the substrate-binding site is located at a similar position as in the SLC23 family, that is, at the ß-strand formed by the transmembrane segments TM3 and TM10. The molecular basis for substrate binding and coupling has not been investigated in detail so far.

In contrast to the SLC23 family, members of the SLC26 family contain additionally a cytosolic domain that is called STAS (Sulfate Transporter and Anti-Sigma factor antagonist) domain. STAS domains exhibit a conserved fold of four ß-strands mixed with a variable number of α-helices (Sharma *et al.*, 2011) (**Figure 8**).
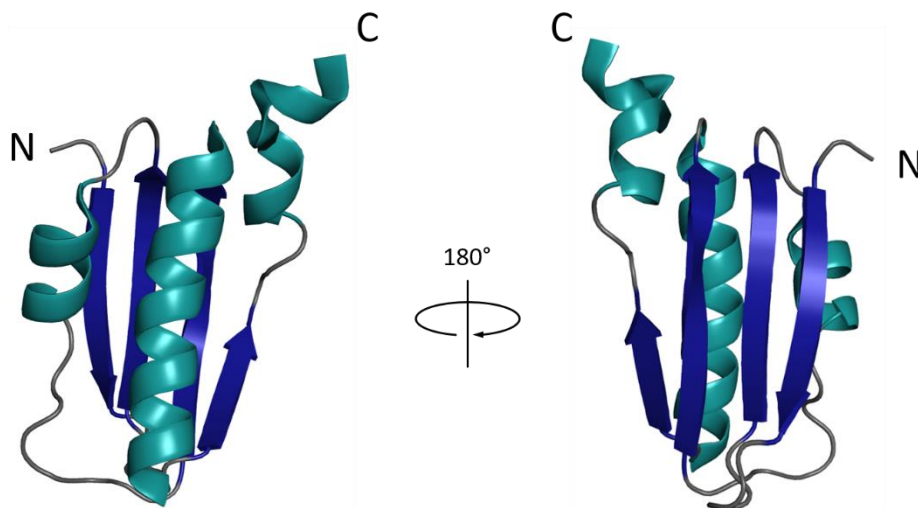
Figure 8: Structure of the STAS domain of SLC26Dg. STAS domains contain a conserved fold of four ß-strands (shown in dark blue), interspersed with α-helices (shown in teal). The N-terminus (that is connected to the transmembrane part) and C-terminus are indicated.

STAS domains from eukaryotic SLC26 transporters in general are larger, containing additional α-helices as well as an unstructured "intervening sequence" (IVS) composed of up to 100 amino acids following the first α-helix. The STAS domains have a hydrophilic nature and can be purified on their own without the use of detergent (Geertsma *et al.*, 2015; Wang *et al.*, 2019).

STAS domains also exist as isolated, soluble proteins. The best characterized one is the *Bacillus subtilis* anti-sigma factor antagonist SpoIIAA that is involved in regulation of σ^F, which induces sporulation as biological stress response. SpoIIAA in turn is regulated by phosphorylation, and was additionally shown to exhibit GTPase activity (Najafi *et al.*, 1996; Seavers *et al.*, 2001). The STAS domain of YtvA, which is fused to a LOV (light-oxygen-voltage sensing) domain was suggested to be responsible for transducing the photosignal by conformational changes of the linker between the two domains (Tang *et al.*, 2010). The isolated STAS domain of *E. coli* DauA co-crystallized with acyl carrier protein (ACP), which implies a role in linking the bacterial bicarbonate transport to the fatty acid metabolism (Babu *et al.*, 2010). For the human SLC26 transporters, a number of pathogenic mutations are known that map to the STAS domain (Karniski, 2004; Bai *et al.*, 2016; Rapp *et al.*, 2017). SLC26A3, SLC26A4 and SLC26A6 STAS domains were shown to interact with CFTR (Cystic Fibrosis Transmembrane Conductance Regulator) (Ko *et al.*, 2004). It appears that STAS domain fusions can serve in extending the functional diversity by interacting with a variety of other cytosolic proteins. For SLC26Dg, removal of the STAS domain caused inactivity of the reconstituted transporter, but the reason for this is unclear at the moment (Geertsma *et al.*, 2015).

1.8. Aim of the thesis

The here presented work aimed at providing a deeper understanding of distinct aspects of the substrate transport mechanism in the SLC23 and SLC26 families. In particular the molecular basis for substrate binding and proton-coupling in SLC23 nucleobase transporters and the relevance of the dimeric state for the transport mechanism in both SLC23 and SLC26 proteins were investigated.

Based on whole-cell uptake assays, SLC23 transporter are reported to being proton-coupled throughout literature. In order to investigate uracil uptake under defined conditions, a proteoliposome-based uracil transport assay was developed. This *in vitro* assay revealed, unexpectedly, a uniporter rather than a symporter function. The first section of the here presented work therefore addresses the basis for the apparent pH-dependency of uracil uptake in *in vivo* assays in contrast to the observed uncoupled uptake *in vitro*. Furthermore, the role of the conserved residues in the substrate binding site of SLC23 homologs for substrate specificity and the translocation mechanism were investigated in detail using a combination of transport and binding studies. Lastly, the relevance of dimer formation for transport in both SLC23 and SLC26 families was examined using *in vivo* and *in vitro* transport assays as well as PELDOR EPR spectroscopy in collaboration with Dr. Benesh Joseph. In this, especially the role of the STAS domain in the SLC26 family was investigated. The data presented here elucidates important and so far not well-investigated aspects of the transport mechanism in these two protein families.

2. Materials

2.1. Chemicals

Chemicals and suppliers used for this study are listed in **Table 2**.

Table 2: List of chemicals and suppliers.

Chemical	Supplier
4-Acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMdiS)	Fisher Scientific
Acetic acid	Roth
Aceton	Roth
Agar powder, food grade	Applichem
Agarose	Sigma-Aldrich
Ammonium persulfate (APS)	Sigma-Aldrich
Ampicillin sodium salt	Roth
Anhydrotetracycline	Roth
L-Arabinose	Roth
ATP	Roth
Benzamide oxime	Sigma-Aldrich
Bromophenol blue	Roth
Calcium chloride dihydrate	Roth
Carbonylcyanid-3-chlorphenylhydrazone (CCCP)	Sigma-Aldrich
Charcaol activated	Roth
Chloramphenicol	Roth
Coomassie® Brillantblau R250 (CBB)	Roth
Copper sulfate anhydrous	Sigma-Aldrich
Cytosine	Sigma-Aldrich
7-Diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM)	Fisher Scientific
Dimethyl sulfoxide (DMSO)	Roth
DL-Dithiothreitol (DTT)	Roth
DNAse I	Applichem
EDTA disodium salt dihydrate	Roth
Ethanol ≥99.8%, p. a.	Roth
Fluorescein-5-maleimide	Fisher Scientific
5-fluorouracil	Sigma-Aldrich
Formaldehyde	Roth
Glutaraldehyde 25%	Roth
Glycerol 86%	Roth
Glycerol ≥99.7%, p.a.	Roth
Glycine PUFFERAN® ≥99%, p.a.	Roth
GTP	Sigma-Aldrich
Hepes	Roth
Hydrochloric acid solution 6N	Roth
Imidazole	Roth
Isopropanol	Roth
$K_3[Fe(CN)_6]$	Roth
Kanamycin sulfate	Roth

Materials

Chemical	Supplier
Lysozyme BioChemica	Applichem
Magnesium sulfate heptahydrate	Roth
Manganese (II) chloride tetrahydrate	Roth
ß-Mercaptoethanol	Roth
Methanol	Roth
1-methyluracil	Sigma-Aldrich
3-methyluracil	Sigma-Aldrich
6-methyluracil	Sigma-Aldrich
MOPS sodium salt	Roth
N-ethylmaleimide (NEM)	Sigma-Aldrich
PEG-maleimide 5000 (PEG-5K)	Fluka
1,10 - Phenanthroline monohydrate	VWR
Phenylmethanesulfonyl fluoride (PMSF)	Roth
Phosphoribosyl pyrophosphate (PRPP)	Sigma-Aldrich
Polypropylene glycol P 2,000	Sigma-Aldrich
Potassium acetate	Roth
Potassium chloride	Roth
Potassium hydroxide pellets	Sigma-Aldrich
Potassium phosphate dibasic	Applichem
Potassium phosphate monobasic	Applichem
Pyrimidine	Sigma-Aldrich
Rapilait milk powder	Migros
Rotiphorese Gel 30 (37.5:1)	Roth
SDS pellet	Roth
Serva DNA Stain G	SERVA
Silver nitrate ($AgNO_3$)	Sigma-Aldrich
Sodium carbonate (Na_2CO_3)	Roth
Sodium chloride	Roth
Sodium fumarate	Sigma-Aldrich
Sodium hydroxide pellets	Roth
Sodium oxalate	Roth
Sodium phosphate dibasic	Roth
Sodium phosphate monobasic	Roth
Sodium sulfate	Roth
Sodium thiosulfate ($Na_2S_2O_3$)	Sigma-Aldrich
N,N,N',N'-Tetramethylethylenediamine (TEMED)	Roth
Thymine	Sigma-Aldrich
Trichloroacetic acid	Sigma-Aldrich
TRIS	Roth
Tryptone BioChemica	Applichem
TWEEN® 20	Roth
UMP	Sigma-Aldrich
Uracil	Sigma-Aldrich
Valinomycin	Sigma-Aldrich
Xanthine	Sigma-Aldrich
Yeast extract BioChemica	Applichem

2.2. Radioactive substrates

Radioactive substrates used for transport assays are listed in **Table 3**.

Table 3: List of radioactive substrates used for transport assays.

Substrate	Supplier
[5,6-^{3}H] - Uracil	American Radiolabeled Chemicals
[8-^{3}H] - Xanthine	American Radiolabeled Chemicals
[1,4-^{14}C] - Fumarate	Movarek
[^{14}C$_2$] - Oxalate	American Radiolabeled Chemicals

2.3. Detergents and lipids

Detergents and lipids used for protein purification and reconstitution are summarized in **Table 4**.

Table 4: List of detergents and lipids used.

Item	Supplier
n-Decyl-ß-maltoside (DM)	Glycon
n-Dodecyl-ß-maltoside (DDM)	Glycon
n-Octyl-ß-D-glucoside (OG)	Glycon
Triton X-100	Roth
L-α-phosphatidylcholine from soy bean, Type II-S, 14-23% choline basis (soy PC)	Sigma-Aldrich

2.4. Commercially available kits

Commercially available kits that were used are listed in **Table 5**.

Table 5: List of commercially available kits used for this study.

Kit	Supplier
Zippy^{TM} Plasmid Miniprep kit	ZymoResearch
GenElute^{TM} Bacterial Genomic DNA	Sigma-Aldrich
Zymoclean^{TM} Gel DNA Recovery kit	ZymoResearch
DNA Clean&Concentrator kit	ZymoResearch
Immobilon Western Chemiluminescent HRP Substrate	Millipore

2.5.   Enzymes

Table 6 shows a list of enzymes that were used for this study.

Table 6: Summary of enzymes used.

Enzyme	Supplier
Phusion HF polymerase	NEB
SapI	NEB
XhoI	NEB
T4 DNA ligase	Thermo Scientific
DNAse I	Applichem
HRV 3C protease	in-house
***anti*-6xHis peroxidase-conjugate**	Roche

2.6.   Media and buffers

The composition of media of buffers are summarized in **Table 7** and **Table 8**.

Table 7: Media composition.

Medium	Composition
LB agar	10 g tryptone, 5 g yeast extract, 5 g NaCl, 15 g agar per liter
LB	10 g tryptone, 5 g yeast extract, 5 g NaCl per liter
TB	12 g tryptone, 24 g yeast extract, 4.7 mL glycerol per liter phosphates added after autoclaving (17 mM KH_2PO_4, 72 mM K_2HPO_4)
TYM	20 g tryptone, 5 g yeast extract, 5.8 g NaCl, 1.2 g $MgSO_4$ per liter
TB stock for 9 L fermenter	108 g tryptone, 216 g yeast extract, 42.3 mL glycerol
20x phosphates for 9 L fermenter	112.9 g K_2HPO_4, 20.9 g KH_2PO_4 in 450 mL
Antifoam	10% Polypropylene glycol P 2,000

Table 8: Buffer composition.

Buffer	Composition
10x DNA loading dye	50% (w/v) glycerol, 0.05% (w/v) bromophenol blue, 100 mM EDTA
5x reducing sample buffer	120 mM Tris pH 6.8, 50% (w/v) glycerol, 100 mM DTT, 2% (w/v) SDS, 0.1 % (w/v) bromophenol blue
5x non-reducing sample buffer	120 mM Tris pH 6.8, 50% (w/v) glycerol, 2% (w/v) SDS, 0.1 % (w/v) bromophenol blue
50x TAE buffer	2 M Tris, 1 M acetic acid, 50 mM EDTA pH 8.0
10x SDS running buffer	256 mM Tris, 1.92 M glycine, 1% (w/v) SDS
Gel stain solution	40% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) CBB
Gel destain solution	40% (v/v) methanol, 10% (v/v) acetic acid
10x PBS	1.37 M NaCl, 27 mM KCl, 100 mM KH_2PO_4, 18 mM Na_2HPO_4
1x PBST	1x PBS, 0.1% (v/v) Tween® 20
Blotting buffer	48 mM Tris, 39 mM glycine, 20% (v/v) methanol

2.7. Bacterial strains

Table 9 contains the genotypes of bacterial strains that were used during the course of this study.

Table 9: Genotypes of bacterial strains used within this study.

Strain	Genotype	Source
E. coli **MC1061**	F⁻ [araD139]$_{B/r}$ Δ(araA-leu)7697 Δ(codB-lacI)3 galK16 galE15(GalS) λ- e14- mcrA0 relA1 rpsL150(strR) spoT1 mcrB1 hsdR2	Lab collection (Casadaban and Cohen, 1980)
E. coli **DB3.1**	F⁻ gyrA462 endA1 Δ(sr1-recA) mcrB mrr hsdS20 glnV44 (=supE44) ara14 galK2 lacY1 proA2 rpsL20 xyl5 leuB6 mtl1	Lab collection (Yanisch-Perron et al., 1985)
E. coli **BW25113**	F⁻ Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) λ⁻ rph-1 Δ(rhaD-rhaB)568 hsdR514	Lab collection (Datsenko and Wanner, 2000)
E. coli **BW25113(ΔuraA)**	F⁻ Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) λ⁻ ΔuraA745::kan rph-1 Δ(rhaD-rhaB)568 hsdR514	Keio collection (Baba et al., 2006)
E. coli **BW25113(ΔxanP)**	F⁻ Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) λ⁻ ΔyicE748::kan rph-1 Δ(rhaD-rhaB)568 hsdR514	Keio collection (Baba et al., 2006)
E. coli **BW25113(ΔuppΔuraA)**	F⁻ Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) λ⁻ ΔuppΔuraA::kan rph-1 Δ(rhaD-rhaB)568 hsdR514	this study
E. coli **BW25113(ΔuraAΔrutG)**	F⁻ Δ(araD-araB)567 ΔlacZ4787(::rrnB-3) λ⁻ ΔuraA Δrutg:: kan rph-1 Δ(rhaD-rhaB)568 hsdR514	this study[1]

---

[1] The knockout of *rutG* in *E. coli* BW25113(*ΔuraA*) was prepared by Leonard Präve as part of his master thesis under my supervision.

2.8.   Vectors and plasmids

The vectors that were used for gene manipulation and expression throughout this study are listed in **Table 10**. A list of SLC23 and SLC26 homologs that were used for functional studies is given in **Table 11**.

Table 10: List of vectors used for cloning in this study.

Vector	Description	Source
pINITcat	Sequencing vector for FX sub-cloning to expression vectors, contains a ColE1 origin and a cam resistance cassette	Lab collection
pBXC3GH	FX cloning-compatible expression vector. Target genes are expressed from arabinose-inducible promoter as C-terminal fusion to GFP with 10x His-tag that is preceded by a 3C protease cleavage site. Contains a ColE1 origin and an amp resistance cassette.	Lab collection
p7XC3RH	FX cloning-compatible expression vector. Target genes are expressed from IPTG-inducible promoter as C-terminal fusion to RFP with 10x His-tag that is preceded by a 3C protease cleavage site. Contains a ColE1 origin and a kan resistance cassette.	Dr. Alina Ornik-Cha
pTX-TH	FX cloning-compatible expression vector. Target genes are expressed from anhydrotetracycline-inducible promoter with C-terminal 10x His-tag that is preceded by a 3C protease cleavage site. Contains a p15A origin and a cam resistance cassette.	Lab collection

Table 11: List of SLC23 and SLC26 homologs used within this study. All proteins were expressed from pBXC3GH.

Homolog	Source organism	NCBI Reference Sequence
UraA	*Escherichia coli*	NP_416992.1
RutG	*Escherichia coli*	NP_415526.4
UacT	*Escherichia coli*	NP_417364.2
XanQ	*Escherichia coli*	NP_417358.2
XanP	*Escherichia coli*	NP_418111.1
YbbY	*Escherichia coli*	NP_415046.4
GhxP	*Escherichia coli*	NP_418488.1
GhxQ	*Escherichia coli*	YP_026186.1
51ThXi	*Thalassospira xiamenensis*	WP_007091992.1
55PyPi	*Pyramidobacter piscolens*	WP_040550746.1
DeOr01	*Desulfosporosinus orientis*	WP_014183842.1
DeOr02	*Desulfosporosinus orientis*	WP_014186949.1
DeOr03	*Desulfosporosinus orientis*	AET67949.1
IlPo01	*Ilyobacter polytropus*	WP_013387787.1

Homolog	Source organism	NCBI Reference Sequence
IlPo02	*Ilyobacter polytropus*	WP_013387549.1
PrMi01	*Proteus mirabilis*	WP_036908698.1
PrMi02	*Proteus mirabilis*	WP_036895180.1
ShBl01	*Shimwellia blattae*	WP_002442531.1
ShBl02	*Shimwellia blattae*	WP_002441299.1
YeEn01	*Yersinia enterolitica*	WP_011815817.1
YeEn02	*Yersinia enterolitica*	WP_005176025.1
YeEn03	*Yersinia enterolitica*	WP_005169999.1
DiIn01	*Dialister invisus*	EEW97258.1
DiIn02	*Dialister invisus*	EEW97316.1
ClAs01	*Clostridium asparagiforme*	EEG55749.1
ClAs02	*Clostridium asparagiforme*	WP_007706169.1
MiMu01	*Mitsuokella multacida*	WP_005839890.1
MiMu02	*Mitsuokella multacida*	WP_005839347.1
LaLa01	*Lactococcus lactis*	ADJ60341.1
LaLa02	*Lactococcus lactis*	WP_011834850.1
SLC26Dg	*Deinococcus geothermalis*	WP_011525799.1
SLC26Si	*Sulfitobacter indolifex*	WP_007117611.1

Some properties of the proteins that were purified are summarized in Table 12. For the extinction coefficient ε_0 that was used to determine protein concentrations spectrophotometrically, the first value given refers to oxidizing (assuming all cysteines form cystines), the second to reducing conditions.

Table 12: **Properties of proteins that were purified within the course of this study**.

Protein	Length (aa)	Molecular weight (kDa)	ε_0 (M^{-1} cm^{-1})
UraA	429	45.1	35660/35410
RutG	442	45.6	36565/36440
51ThXi	431	44.4	40910/40910
55PyPi	454	47.5	59400/58900
SLC26Dg	499	53.5	37930/37930
SLC26Si	573	60.7	42400/42400

Protein	Length (aa)	Molecular weight (kDa)	ε_0 (M^{-1} cm^{-1})
UPRT	208	22.5	12950/12950

2.9. Primers

In **Table 13**, the primers used for routine sequencing in pINITcat or pBXC3GH are specified. All sequencing was done by Microsynth (Göttingen).

Table 13: Sequencing primers used to verify gene sequences in pINITcat or pBXC3GH.

Name	Sequence (5'-3')
pINIT_for	GAGTAGGACAAATCCGC
pINIT_rev2	TGGCAGTTTATGGCGGGCGT
pBAD24_rev	GCTGAAAATCTTCTCTCATCCG

For mobilizing target ORFs from genomes, FX-cloning compatible primers were used generated using *fxcloning.org* (extensive mode). Their sequences are shown in **Table 14**. Several SLC23 homologs were already available in pINITcat and pBXC3GH from my own master thesis and are not included.

Table 14: FX cloning-compatible primers to amplify target genes from genome.

Name	Sequence (5'-3')
UraA_FX_for	ATATATGCTCTTCTAGTACGCGTCGTGCTATCGGGGTGAGTGAA
UraA_FX_rev	TATATAGCTCTTCATGCTTTGTCTGTTATATCCGCGTCTTCTGC
RutG_FX_for	ATATATGCTCTTCTAGTGCAATGTTCGGTTTTCCTCACTGGCAG
RutG_FX_rev	TATATAGCTCTTCCTGCTGGCTCCTGATGAACAACTTCTGGCGG
51ThXi_FX_for	ATATATGCTCTTCTAGTCGATATTTTCCTTCCTGGTCTCTGACC
51ThXi_FX_rev	TATATAGCTCTTCATGCGTCTGCGGGGGCCTCGGATGATTTGCG
55PyPi_FX_for	ATATATGCTCTTCTAGTGCTCGAAAAGAACTGGTGTATGGAGTG
55PyPi_FX_rev	TATATAGCTCTTCCTGCGTCCCTGAAGCCTCGTTCCTCGCGC
ShBl02_FX_for	ATATATGCTCTTCTAGTACACGTCGGGCCATCGGGGTAAGCG
ShBl02_FX_rev	TATATAGCTCTTCATGCAGCACGCTTTTCGTCGCTCTCCTGTAT
YeEn01_FX_for	ATATATGCTCTTCTAGTAGTCGTCGTACCATTGGCGTCAGCG
YeEn01_FX_rev	TATATAGCTCTTCATGCTTTAAGCGCAGGCTTGTCTTCCGTCGC
YeEn02_FX_for	ATATATGCTCTTCTAGTTTATTTTTCTTGTGGGAAGATGAGGTT
YeEn02_FX_rev	TATATAGCTCTTCATGCATGTCTACGCAATCCGCCGAACTTAAT

Name	Sequence (5'-3')
YeEn03_FX_for	ATATATGCTCTTCTAGTGCGAATAACTGGTTTCCAACATGGCGT
YeEn03_FX_rev	TATATAGCTCTTCATGCACCATCTTGCGGCAAGGGCTTACCATC
MiMu02_FX_for	ATATATGCTCTTCTAGTCAAAAACGTCCTATCGGTGTCGAAGAG
MiMu02_FX_rev	TATATAGCTCTTCATGCTTGATGCGTGTTGTCTTCTGCCTTCTG
AnHy02_FX_for	ATATATGCTCTTCTAGTGTGAAAATAATGGCGTATGATCCTGAA
AnHy02_FX_rev	TATATAGCTCTTCATGCTTGATTTTTATTATAAAACACGTAATT
UPRT_FX_for	ATATATGCTCTTCTAGTAAGATCGTGGAAGTCAAACACCCA
UPRT_FX_rev	TATATAGCTCTTCATGCTTTCGTACCAAAGATTTTGTCACCGGC

Table 15 shows the sequences of primers used to generate the *Δupp ΔuraA* and *ΔuraA ΔrutG* knockouts in *E. coli* BW25113 as well as the primers used for verification of the knockouts by PCR and sequencing.

Table 15: Primers used to generate and verify knockout strains of *E. coli* BW25113.

Name	Sequence (5'-3')
upp_uraA_KO_for	GGTATAATCCGTCGATTTTTTTTGTGGCTGCCCCTCAAAGGAGAAAGAGTGTGTAGGCTGGAGCTGCTTC
upp_uraA_KO_rev	CAGAATCCCATGATGTTTGAACCGGGCAGCAACACTGCCCGGTCGGTACAATGGGAATTAGCCATGGTCC
rutG_KO_for	TGATGCGCCCTGCTTGTTAACCTCCTTAAGGAGACAGCTGGTGTAGGCTGGAGCTGCTTC
rutG_KO_rev	ACTGCGACCATAATCAAATTACCGTTCTGGCTTAAATCGAATGGGAATTAGCCATGGTCC
kan_check_for	ACGGGCGTTCCTTGCGCAGC
kan_check_rev	GGAGCGGCGATACCGTAGAG
upp_check_for	AGGGTAAAGGTTAGTTTTCGGATGGAATAATCTTCTTTCA
uraA_check_rev	GAGGAGTTATCCCCCGGCCAGAAACTTGCAAAGGTTTCGT
rutG_check_for	GGATGCCCGCAACTGGAAG
rutG_check_rev	CGCCGGTTACCCGGTGGTTG

Materials

Table 16 is a list of primers used to generate substrate binding site mutants in the SLC23 homologs UraA, 51ThXi and 55PyPi. All mutagenesis primers were designed using *PrimerX* (https://www.bioinformatics.org/primerx/cgi-bin/protein_1.cgi).

Table 16: List of primers used to generate substrate binding site mutants of SLC23 transporters.

Name	Sequence (5'-3')
UraA_E241A_for	GTTAGTGGTTATTGCCGCGCACGTAGGGCACCTGG
UraA_E241A_rev	CCAGGTGCCCTACGTGCGCGGCAATAACCACTAAC
UraA_H245A_for	GTTATTGCCGAACACGTAGGGGCACTGGTAGTAACGGCTAATATC
UraA_H245A_rev	GATATTAGCCGTTACTACCAGTGCCCCTACGTGTTCGGCAATAAC
UraA_H245D_for	CGAACACGTAGGGGACCTGGTAGTAAC
UraA_H245D_rev	GTTACTACCAGGTCCCCTACGTGTTCG
UraA_E290A_for	CAAATACTACTTACGGAGCAAACATTGGCGTGATGG
UraA_E290A_rev	CCATCACGCCAATGTTTGCTCCGTAAGTAGTATTTG
51ThXi_E256A_for	CATTCTGGTGGCGGCCAACCTCGGTCATATC
51ThXi_E256A_rev	GATATGACCGAGGTTGGCCGCCACCAGAATG
51ThXi_E256D_for	CATTCTGGTGGCGGATAACCTCGGTCATATC
51ThXi_E256D_rev	GATATGACCGAGGTTATCCGCCACCAGAATG
51ThXi_E256S_for	CCTTCATTCTGGTGGCGAGCAACCTCGGTCATATCAAG
51ThXi_E256S_rev	CTTGATATGACCGAGGTTGCTCGCCACCAGAATGAAGG
51ThXi_E256Q_for	CCTTCATTCTGGTGGCGCAGAACCTCGGTCATATCAAG
51ThXi_E256Q_rev	CTTGATATGACCGAGGTTCTGCGCCACCAGAATGAAGG
51ThXi_H260A_for	CGGAAAACCTCGGTGCTATCAAGGCGATTGG
51ThXi_H260A_rev	CCAATCGCCTTGATAGCACCGAGGTTTTCCG
51ThXi_H260D_for	CGGAAAACCTCGGTGATATCAAGGCGATTG
51ThXi_H260D_rev	CAATCGCCTTGATATCACCGAGGTTTTCCG
51ThXi_E303A_for	GACCACCTATGTTGCCAACATGGGTGTCATG
51ThXi_E303A_rev	CATGACACCCATGTTGGCAACATAGGTGGTC
51ThXi_E303D_for	GACCACCTATGTTGATAACATGGGTGTCATG
51ThXi_E303D_rev	CATGACACCCATGTTATCAACATAGGTGGTC
51ThXi_E303Q_for	GGGTGACCACCTATGTTCAGAACATGGGTGTCATGG
51ThXi_E303Q_rev	CCATGACACCCATGTTCTGAACATAGGTGGTCACCC
51ThXi_E256A_H260A_for	GCCTTCATTCTGGTGGCGGCAAACC
51ThXi_E256A_H260A_rev	GGTTTGCCGCCACCAGAATGAAGGC

42

Name	Sequence (5'-3')
51ThXi_E256H_H260E_for	GTGGCCTTCATTCTGGTGGCGCATAATCTGGGCGAAATCAAGGCGATTGGTGGCATG
51ThXi_E256H_H260E_rev	CATGCCACCAATCGCCTTGATTTCGCCCAGATTATGCGCCACCAGAATGAAGGCCAC
51ThXi_F82A_for	CTATCTGGGATCTTCCGCGGCCTTTATCGCCGTTG
51ThXi_F82A_rev	CAACGGCGATAAAGGCCGCGGAAGATCCCAGATAG
55PyPi_E264A_for	CTCGGTCATTCTCGCTAGCATCGGCGACTAC
55PyPi_E264A_rev	GTAGTCGCCGATGCTAGCGAGAATGACCGAG
55PyPi_D268A_for	GAAAGCATCGGCGCCTACTTCAACGTC
55PyPi_D268A_rev	GACGTTGAAGTAGGCGCCGATGCTTTC
55PyPi_E313A_for	CACGAGCTATACCGCAAACATCGGCCTGATC
55PyPi_E313A_rev	GATCAGGCCGATGTTTGCGGTATAGCTCGTG

To generate an RFP-fusion version of UraA in pTX-XH, UraA was subcloned into p7XC3RH and together with RFP into the desired expression vector. For this, a reverse primer directed against RFP (**Table 17**) was used together with the UraA_FX_for primer (**Table 14**).

Table 17: FX-cloning- compatible primer used to mobilize UraA as RFP-fusion from p7XC3RH.

Name	Sequence (5'-3')
RFP_FX_rev	TATATAGCTCTTCATGCTTATTAATGATGGTGATGATGATGGTG

Primers used to generate cysteine mutants of UraA and 51ThXi for site-directed crosslinking or labeling were as well designed using *PrimerX* and are listed in **Table 18**. Cysteine-less UraA (C61S/C97S/C102S/C318S) was generated by myself in the work preceding this study.

Table 18: Primers used to generate cysteine mutants of UraA and 51ThXi.

Name	Sequence (5'-3')
UraA_G343C_for	GTTTCGCTGCTGCTTTATTGCGTCATCGGTGCTTCCGG
UraA_G343C_rev	CCGGAAGCACCGATGACGCAATAAAGCAGCAGCGAAAC
UraA_S370C_for	GAACCTGATCCTGACTTGCGTGATTTTGATCATC
UraA_S370C_rev	GATGATCAAAATCACGCAAGTCAGGATCAGGTTC
UraA_I374C_for	CTTCCGTGATTTTGTGCATCGGCGTCAGTGG
UraA_I374C_rev	CCACTGACGCCGATGCACAAAATCACGGAAG

Name	Sequence (5'-3')
51ThXi_C69S_for	GATCTTCTTTATCAGCACGGCGGGGC
51ThXi_C69S_rev	GCCCCGCCGTGCTGATAAAGAAGATC
51ThXi_G356C_for	CTTGGTGTTGCGGTATTTTGCCTTATTGCGGCGGCAATG
51ThXi_G356C_rev	CATTGCCGCCGCAATAAGGCAAAATACCGCAACACCAAG
51ThXi_G383C_for	CGCAATCTGTTCACCGTTTGCATTACCCTTATTTTCGGG
51ThXi_G383C_rev	CCCGAAAATAAGGGTAATGCAAACGGTGAACAGATTGCG
51ThXi_I387C_for	CACCGTTGGTATTACCCTTTGCTTCGGGGGCGGCGACTTC
51ThXi_I387C_rev	GAAGTCGCCGCCCCCGAAGCAAAGGGTAATACCAACGGTG
51ThXi_S97C_for	GACCGGCTATGCCGGATGCGGCCCCAACCCGGATATC
51ThXi_S97C_rev	GATATCCGGGTTGGGGCCGCATCCGGCATAGCCGGTC
G383C_I387W_for	CCGCGCAATCTGTTCACCGTTTGCATTACCCTGTGGTTCGGGGGCGG
G383C_I387W_rev	CCGCCCCCGAACCACAGGGTAATGCAAACGGTGAACAGATTGCGCGG
51ThXi_L379W_G383C_I387W_for	TTCACCGTTTGCATTACCCTGTGGTTCGGGGGCGG
51ThXi_L379W_G383C_I387W_rev	CCGCCCCCGAACCACAGGGTAATGCAAACGGTGAA

Tryptophans were introduced at the dimer interface of 51ThXi by site-directed mutagenesis using the primers listed below (**Table 19**).

Table 19: Primers used to generate tryptophan dimer interface mutants of 51ThXi.

Name	Sequence (5'-3')
51ThXi_A153W_for	CATCGGTCTTAATCTGTGGCCCGTCGGGATCAAC
51ThXi_A153W_rev	GTTGATCCCGACGGGCCACAGATTAAGACCGATG
51ThXi_L379W_for	GATTTTTCCGATCCGCGCAATTGGTTCACCGTTGGTATTACCC
51ThXi_L379W_rev	GGGTAATACCAACGGTGAACCAATTGCGCGGATCGGAAAAATC
51ThXi_I387W_for	CACCGTTGGTATTACCCTTTGGTTCGGGGGCGGCGACTTC
51ThXi_I387W_rev	GAAGTCGCCGCCCCCGAACCAAAGGGTAATACCAACGGTG

Concatemers of 51ThXi were constructed as described in section 3.2.7 using the primers shown in **Table 20**.

Table 20: Primers used to generate concatemers of 51ThXi.

Name	Sequence (5'-3')
51ThXi_9GS_for	ATGCTCTTCTAGTGGCTCGAGCGGTAGTAGCGGGTCGTCACGATATTTTCCTTCCTGGTC
51ThXi_3GS_rev	TATATAGCTCTTCATGCGCTCGAGCCGTCTGCGGGGGCCTCGGATGATTTGCGACC

Primers used for mutating SLC26Dg and SLC26Si are listed in **Table 21** and **Table 22**, respectively. Both genes in pBXC3GH as well as the cysteine-free versions (SLC26Dg(C282S) and SLC26Si(C578S)) were kindly provided by Dr. Yung-Ning Chang.

Table 21: Primers used to generate mutants of SLC26Dg.

Name	Sequence (5'-3')
SLC26Dg_FX_for	ATATATGCTCTTCTAGTACTGTACATTCACCGAGGTTCGATCTC
SLC26Dg_FX_rev	TATATAGCTCTTCATGCATGCCCACCCATCCGATCCAGCGCATC
SLC26Dg_K121C_for	CGCGCTATCTCTGCTTCGTGCCGCG
SLC26Dg_K121C_rev	CGCGGCACGAAGCAGAGATAGCGCG
SLC26Dg_V129C_for	CGCAGCGTCATGTGCGGCTTCGTGAATG
SLC26Dg_V129C_rev	CATTCACGAAGCCGCACATGACGCTGCG
SLC26Dg_Q247C_for	GAAAGTCTGCTGACCGCGTGCCTGATCGACGAACGAAC
SLC26Dg_Q247C_rev	GTTCGTTCGTCGATCAGGCACGCGGTCAGCAGACTTTC
SLC26Dg_L248C_for	CTGCTGACCGCGCAGTGTATCGACGAACGAAC
SLC26Dg_L248C_rev	GTTCGTTCGTCGATACACTGCGCGGTCAGCAG
SLC26Dg_K390C_for	GTTGTTCTTCGCTCGGTGCGTGTCGCAACTCTCG
SLC26Dg_K390C_rev	CGAGAGTTGCGACACGCACCGAGCGAAGAACAAC
SLC26Dg_R412C_for	GGACGTACCGGGTTTGCGGACAGCTGTTCTTC
SLC26Dg_R412C_rev	GAAGAACAGCTGTCCGCAAACCCGGTACGTCC
SLC26Dg_S419C_for	CAGCTGTTCTTCGTCTGTACGCATGATTTC
SLC26Dg_S419C_rev	GAAATCATGCGTACAGACGAAGAACAGCTG
SLC26Dg_H421C_for	CTGTTCTTCGTCAGTACGTGCGATTTCCTGCACCAGTTTG
SLC26Dg_H421C_rev	CAAACTGGTGCAGGAAATCGCACGTACTGACGAAGAACAG
SLC26Dg_L459C_for	CTGGACAAGGTGATGTGCAAGTTCATGCGCCAGG

45

Name	Sequence (5'-3')
SLC26Dg_L459C_rev	CCTGGCGCATGAACTTGCACATCACCTTGTCCAG
SLC26Dg_A475C_for	GGGGACTGAATTGCGCGTCGGCGAC
SLC26Dg_A475C_rev	GTCGCCGACGCGCAATTCAGTCCCC
SLC26Dg_V481C_for	CGGCGACCCTGTGCGAGCGGCTGGC
SLC26Dg_V481C_rev	GCCAGCCGCTCGCACAGGGTCGCCG
SLC26Dg_GS-L_for	GTTGTTCTTCGCTCGGAAGGGCAGCGGCGGCTCGCAGGTGACGCCGGTG
SLC26Dg_GS-L_rev	CACCGGCGTCACCTGCGAGCCGCCGCTGCCCTTCCGAGCGAAGAACAAC

Table 22: Primers used to generate mutants of SLC26Si.

Name	Sequence (5'-3')
SLC26Si_L49C_for	CAATCGCTGGCCTATGCGCTGTGCGCGGGGATGCC
SLC26Si_L49C_rev	GGCATCCCCGCGCACAGCGCATAGGCCAGCGATTG
SLC26Si_S153C_for	CATCCTGATTGCGGCATGCCAGTTGCGGCATATTC
SLC26Si_S153C_rev	GAATATGCCGCAACTGGCATGCCGCAATCAGGATG
SLC26Si_GS-L_for	CATCTTTATAAGGGCAGCGGCGGCAGCATGGCTGTGGTG
SLC26Si_GS-L_rev	CACCACAGCCATGCTGCCGCCGCTGCCCTTATAAGATG

3. Methods

3.1. Microbiology

3.1.1. Preparation of chemically competent E. coli for transformation

In order to prepare chemically competent *E. coli* MC1061 for routine calcium chloride transformation (Dagert and Ehrlich, 1979), 5 mL LB medium without antibiotics were inoculated from glycerol stock or plate and incubated overnight at 37°C, shaking. From this overnight culture, 200 mL TYM medium (20 g/L tryptone, 5 g/L yeast extract, 0.1 M NaCl, 10 mM $MgSO_4$) in 1 L baffled flasks were inoculated to 1% (v/v) and cells were incubated at 37°C, shaking, until OD_{600} reached 0.5 - 0.9. The cells were placed on ice for 15 min prior to harvesting by centrifugation at 5000 rpm (SLC-6000 rotor) for 15 min at 4°C. The sample was kept on ice for all following steps. Cells were resuspended in 100 mL ice-cold TfB I buffer (30 mM KAc, 50 mM $MnCl_2$, 0.1 M KCl, 10 mM $CaCl_2$, 15% (w/v) glycerol) and pelleted again. The pellet was resuspended in 10 mL ice-cold TfB II buffer (10 mM Na-MOPS pH 7.0, 75 mM $CaCl_2$, 10 mM KCl, 15% (w/v) glycerol) and 100 μL aliquots were frozen in liquid N_2 and stored at -80°C until use.

For fast small-scale preparation of chemically competent *E. coli* BW25113, 5 mL LB without antibiotics were inoculated from glycerol stock or plate and incubated overnight at 37°C, shaking. From this, small amounts of LB medium (5-10 mL) were inoculated to 1% (v/v) and incubated at 37°C, shaking, until OD_{600} reached ~0.5. 1 mL aliquots were pelleted by centrifugation at 9000 g for 30 sec at room temperature and resuspended in 500 μL of ice-cold 0.1 M $CaCl_2$. After incubation on ice for 30 min, cells were pelleted again and resuspended in 50 μL of ice-cold 0.1 M $CaCl_2$. These cells were immediately used for standard heat shock transformation (section 3.1.2).

3.1.2. Transformation of chemically competent E. coli

Aliquots of competent cells prepared as described in section 3.1.1 were incubated with plasmid DNA for 15 min on ice. Cells were subjected to heat shock at 42°C for 1 min, followed by 1 min incubation on ice and recovery in 500 μL LB medium at 37°C for 1 hour with gentle agitation. Typically, 100 μL of transformed cells were incubated overnight at 37°C on LB agar plates containing the appropriate antibiotics as selection markers.

3.1.3. Preparation of electrocompetent E. coli

In order to prepare electrocompetent *E. coli* cells, 5 mL LB medium were inoculated with 1% (v/v) of an overnight culture and incubated at 37°C, shaking, until OD_{600} ~0.5. Aliquots of 1 mL were incubated 15 min on ice. Samples were pelleted by centrifugation at 2200 g for 7 min at 4°C and resuspended in

1 mL of ice-cold 1 mM Hepes pH 7.5. Samples were pelleted again and resuspended in 500 µL 10% (v/v) glycerol, followed by another centrifugation step and a final resuspension to 50 µL using 10% (v/v) glycerol. Cells were immediately used for electroporation.

3.1.4. *Transformation of electrocompetent E. coli*

For electroporation, either 10 ng plasmid DNA or 100 ng PCR product were provided in pre-cooled electroporation cuvettes (gap width 0.1 cm) and electrocompetent cells prepared as described in section 3.1.3 were added on top. 2 kV pulses were applied using a Biorad *MicroPulser*. Cells were immediately transferred to 1 mL LB + 1% (w/v) glucose and incubated for 1 hour at 37°C with gentle agitation for recovery. Typically, 100 µL of transformed cells were incubated overnight at 37°C on LB agar plates containing the appropriate antibiotics as selection markers.

3.1.5. *Cultivation and storage of E. coli*

Single colonies were picked from plate and placed in 5 mL LB medium containing the proper antibiotic. These cultures were incubated overnight at 37°C, shaking, and either used for plasmid isolation, for subsequent experiments or for preparation of glycerol stocks for long-term storage. For this, 750 µL of bacterial culture were mixed with 250 µL sterile 86% (w/v) glycerol, frozen in liquid N_2 and stored at -80°C.

3.1.6. *Genomic knockouts in E. coli BW25113*

Genetic knockouts in *E. coli* BW25113 were essentially generated as described in standard protocols (Datsenko and Wanner, 2000). For the generation of *E. coli* BW25113(*ΔuppΔuraA*), both *upp* and *uraA* were exchanged with a kanamycin resistance cassette in one single step, as both genes are located in one operon. For generating the *E. coli* BW25113(*ΔuraAΔrutG*) double knockout, *E. coli* BW25113(*ΔuraA*) (Baba *et al.*, 2006) was cured from kanamycin resistance using FRT recombinase and *rutG* was subsequently exchanged by a kanamycin resistance cassette[2]. All steps were monitored by PCR and sequencing of isolated genomic DNA.

---

[2] The *E. coli* BW25113(*ΔuraAΔrutG*) knockout was generated by Leonard Präve under my supervision as part of his master thesis.

3.2. *Molecular biology*

3.2.1. *Plasmid isolation*

Typically, 3 mL of bacterial culture were pelleted by centrifugation, resuspended in 600 µL Milli-Q and used for plasmid isolation employing the Zymo Research *Zippy^TM Plasmid Miniprep Kit* according to the manufacturer's recommendations. Plasmid DNA was eluted from the column with 42 µl of Milli-Q and DNA content was determined spectrophotometrically. Plasmids were, if necessary, sequence-verified and used for subcloning, as templates for PCR or for new transformations. All sequencing was performed by Microsynth (Göttingen).

3.2.2. *Isolation of genomic DNA*

For analysis of knockout mutants that were generated within this study (section 3.1.6), genomic DNA was isolated from 5 mL bacterial cultures using the GenElute^TM Bacterial Genomic DNA kit (Sigma-Aldrich) according to the manufacturer's recommendation. DNA was eluted from the column using Milli-Q and yield was assessed by A_{260} measurement at the Nanodrop. Isolated genomic DNA was used for PCR, subsequent agarose gel electrophoresis and sequence analysis of PCR products in order to determine the success of the genomic knockout.

3.2.3. *PCR amplification of open reading frames*

To obtain DNA sequences of a set of UraA homologs as well as the one of *E. coli* uracil phosphoribosyltransferase (UPRT), open reading frames were amplified from the genomes of a variety of bacterial and archaeal organisms, which were purchased from the *German Collection of Microorganisms and Cell Cultures GmbH (DSMZ).* Primers were designed containing SapI restriction sites suitable for FX cloning using *FXcloning.org* and synthesized by Microsynth. ORFs were amplified using Phusion HF polymerase in a touchdown PCR approach according to the following specifications:

Table 23: Composition of PCR reaction to amplify ORFs from genomic DNA.

Table 24: PCR protocol for amplification of ORFs from genomic DNA.

Stock	Amount
5x Phusion HF buffer	10 µL
DMSO	2 µL
10 mM dNTPs	1 µL
5 µM forward primer	5 µL
5 µM reverse primer	5 µL
1-100 ng/µL genomic DNA	1 µL
Phusion HF polymerase	0.5 µL
Milli-Q	to 50 µL

Step	Temperature	Time	
Initial denaturation	98°C	2 min	
Denaturation	98°C	10 sec	
Annealing	65°C – 0.5°C/cycle	15 sec	14 x
Elongation	72°C	30 sec/kb	
Denaturation	98°C	10 sec	
Annealing	53°C	15 sec	14 x
Elongation	72°C	30 sec/kb	
Final elongation	72°C	5 min	

3.2.4. Site-directed mutagenesis

For this study, mutants were generated by site-directed mutagenesis using a two-step megaprimer approach (Forloni *et al.*, 2019). With pINIT constructs as templates, in the first step the target gene was only partially amplified using either the forward or reverse pINIT sequencing primers in combination with a suitable mutagenesis primer. The two PCR products were mixed in the second PCR step to serve as a megaprimer and the full-length, mutated gene was amplified using both pINIT sequencing primers. This method also allows to combine more than two gene fragments in the second PCR step in order to introduce multiple mutations at the same time. In **Figure 9**, as an example a PCR procedure is illustrated in which two mutations are introduced into the target gene, meaning in the first step three different PCR reactions are performed that are combined in step 2 to amplify the full-length, mutagenized gene.
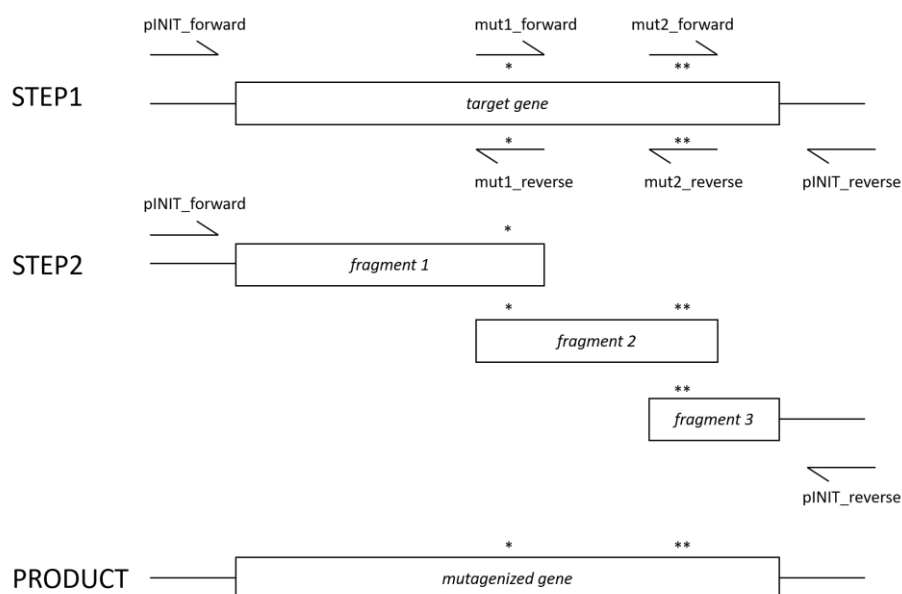
Figure 9: The megaprimer PCR approach for site-directed mutagenesis. As an example, a PCR reaction is illustrated that introduces two mutations into the target gene, indicated by one (*) and two asterisks (**), respectively. In step 1, 3 different reactions are carried out in different tubes, amplifying parts of the gene that are in step 2 combined to a megaprimer as template to amplify the full-length gene. mut1 and mut2 refer to the two mutagenesis primers.

For amplification of fragments and the full-length gene, a touchdown PCR approach was used.

Table 25: Composition of samples for the first step of the megaprimer PCR approach.

Stock	Amount
5x Phusion HF buffer	10 µL
DMSO	2 µL
10 mM dNTPs	1 µL
5 µM forward primer	5 µL
5 µM reverse primer	5 µL
1-100 ng/µL template	1 µL
Phusion HF polymerase	0.5 µL
Milli-Q	to 50 µL

Table 26: Composition of samples for the second step of the megaprimer PCR approach.

Stock	Amount
5x Phusion HF buffer	10 µL
DMSO	2 µL
10 mM dNTPs	1 µL
5 µM pINIT_forward	5 µL
5 µM pINIT_reverse	5 µL
Template 1	1 µL
Template 2	1 µL
(Template 3)	(1 µL)
Phusion HF polymerase	0.5 µL
Milli-Q	to 50 µL

Table 27: Temperature protocol for the megaprimer PCR approach.

Step	Temperature	Time	
Initial denaturation	98°C	2 min	
Denaturation	98°C	15 sec	
Annealing	65°C – 0.5°C/cycle	45 sec	20 x
Elongation	72°C	30 sec/kb	
Final elongation	72°C	5 min	

3.2.5. Agarose gel electrophoresis and DNA extraction

Generally, PCR products supplemented with appropriate amounts of 10x DNA loading dye were separated on 1% (w/v) or 1.5% (w/v) TAE-agarose gels applying 10 V/cm for 45 min. For DNA visualization, the Serva DNA stain G was used that emits a green fluorescence when binding to DNA or RNA. DNA fragments of correct size were cut from gel and purified using the ZymoResearch Gel DNA Recovery kit according to the manufacturer's recommendations. DNA was eluted using 10 µL Milli-Q and concentration and quality were monitored spectrophotometrically using the NanoDrop 1000 (Thermo Scientific).

3.2.6. FX cloning

Typically, PCR products were cloned to pINIT and subcloned into expression vectors using FX cloning as described previously (Geertsma and Dutzler, 2011; Geertsma, 2013). 50 ng of vector and 5x molar excess of insert were incubated in CutSmart buffer with 5U SapI in a total volume of 10 µL for 1 hour at 37°C, followed by inactivation of the restriction enzyme at 65°C for 20 min. 1.25 µL 10 mM ATP and 5U T4 DNA ligase were added and samples were incubated for 1 hour at room temperature followed by 20 min at 65°C. 5 µL of this were used for heat shock transformation to *E. coli* MC1061. All sequences were verified (Microsynth, Göttingen) before subcloning into pBXC3GH, an FX-cloning compatible vector allowing expression of C-terminal GFP fusion proteins with 10x His-tag from an arabinose-inducible promoter. If not stated otherwise, this vector was used for overexpression for all experiments described herein.

3.2.7. Construction of 51ThXi concatemers

In order to construct a concatemeric version of 51ThXi, in which the two protomers are linked with a GS-linker of variable length, the 51ThXi gene was extended either C-terminally (for the first protomer within the concatemer) or N-terminally (for the second protomer) with codons for glycine and serine residues and a XhoI restriction site (C^{V}TCGAG) which can be incorporated into a GSS-linker sequence,

52

as both TCG and AGC are triplets for serine (**Figure 10**). The genes for the single protomers were separately sequence-verified in pINIT. The concatemers were generated by subcloning of the two pINIT constructs into pBXC3GH while simultaneously digesting with SapI and XhoI. This strategy allows for fast screening of concatemers with different linker lengths (by combination of pINIT constructs containing a different number of triplets coding for GSS) and also for rapidly generating mutant concatemers (by site-directed mutagenesis of already extended pINIT constructs).
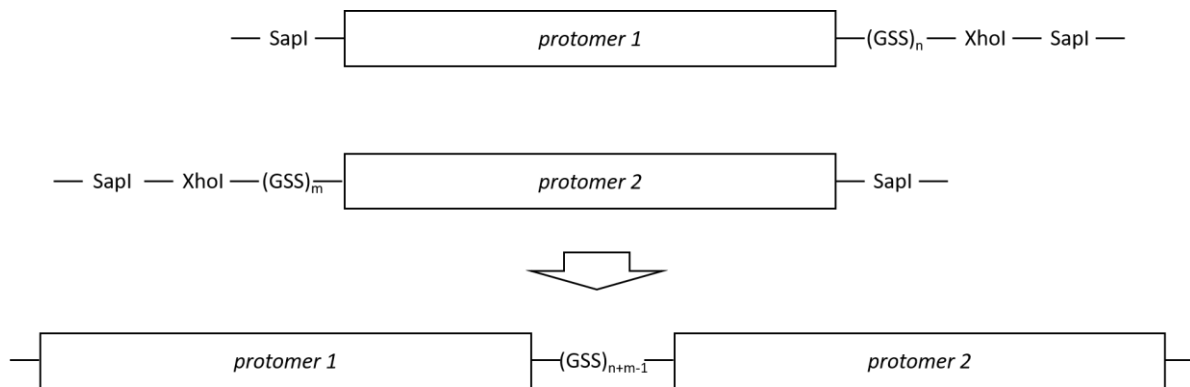
Figure 10: Cloning strategy to construct concatemeric protein versions.

3.2.8. Construction of vectors to express RFP fusions from pTet promoter

For co-expression of UraA and UPRT and unambigiuous identification of both proteins, a vector was constructed for UraA-RFP fusion protein expression on the basis of the pTX-XH vector (generated by Dr. Eric Geertsma). The pTX-TH vector is a tetracycline-inducible vector for expression of 10x His-tagged protein, harboring a chloramphenicol resistance cassette and a p15A origin of replication. As no RFP-fusion version of this vector was available, UraA was first subcloned into p7XC3RH (kindly gifted by Dr. Alina Ornik-Cha) and from this cloned together with RFP into pTX-XH.

3.3. Protein biochemistry

3.3.1. Small-scale expression in 96-well plate format

Small-scale expression for analytical purposes, for example to determine expression levels or for crosslinking experiments in membrane vesicles, was routinely performed in a 96-well plate format (Marino *et al.*, 2017). For this, 700 µL TB medium supplemented with 1 mM $MgSO_4$ and 100 µg/mL ampicillin were inoculated in a 96 deep-well block to 1% (v/v) using an overnight preculture prepared in LB medium containing 100 µg/mL ampicillin. Samples were incubated at 37°C, 950 rpm (Infors shaker) for 1.5 hours. The temperature was set to 25°C and cells were incubated for another hour before they were induced using 10^{-3}% (w/v) L-arabinose. After overnight incubation at 25°C, 950 rpm, cells were harvested by centrifugation at 2500 g for 10 min at 4°C. Pellets were resuspended in 300 µL 50 mM KPi pH 7.5, 150 mM NaCl, 10% (w/v) glycerol. 75 µL were transferred to a transparent 96-well plate, 150 µL to a black 96-well plate to measure OD_{600} and GFP fluorescence at the Tecan Infinite 200 plate reader. For O_{D600}, a correction factor was determined using the *Ultrospec 10* cell density meter (Amersham) or comparable. Assuming an OD_{600} of 1 corresponds to 0.3 mg/mL of total protein, 1 or 2 mg total protein were pelleted in 2 mL Fastprep tubes (Sarstedt). The supernatant was discarded and samples were either stored at -20°C or immediately used for membrane vesicle preparation. For this, pellets were resuspended in 400 µL 50 mM KPi pH 7.5, 150 mM NaCl, 10% (w/v) glycerol, 1 mM $MgSO_4$, 1 mM PMSF and trace amounts of DNAse I. Roughly 300 mg glassbeads diameter 0.1 mm were added and cells were disrupted by bead beating using the *Fastprep-24* device (MP Biomedicals) by shaking twice for 20 sec at force 4 with 5 min incubation on ice in between. Samples were either used to immediately analyze expression levels on gel, for fluorescent-detection size-exclusion chromatography (FSEC) or for analytical crosslinking (section 3.3.10).

3.3.2. Large-scale expression of membrane proteins for preparative purification

E. coli MC1061 transformed with pBXC3GH constructs were cultivated either in 9 L TB/ampicillin in a fermenter (Bioengineering) or in 5 L baffled flasks. In either case, the medium was inoculated to 1% (v/v) using a preculture prepared in TB medium, 1 mM $MgSO_4$, 100 µg/mL ampicillin. Cells were grown at 37°C until OD_{600} = 1-2 before the temperature was reduced to 25°C. Protein expression was induced using 5 x 10^{-3}% (w/v) L-arabinose and continued for 16 hours at 25°C. Cell pellets were resuspended in 50 mM KPi pH 7.5, 150 mM NaCl, 10% (w/v) glycerol, 1 mM $MgSO_4$ and incubated in the presence of 1 mg/mL lysozyme and trace amounts of DNAse I for 1 h at 4°C. Cells were disrupted using either an APV Manton/Gaulin homogenizer or a pressure cell homogenizer from Stanstedt. PMSF was added to 1 mM, the lysate was cleared by low-spin centrifugation (15,000 g, 20 min) and membranes were collected by ultracentrifugation (140,000 g, 1 h). Membranes were resuspended to

0.5 g/mL in 50 mM KPi pH 7.5, 150 mM NaCl, 10% (w/v) glycerol, flash-frozen and stored at -80°C until use. In case the proteins were intended for oxidative crosslinking or site-directed spin-labeling, this buffer additionally contained 10 mM DTT.

3.3.3. Preparative purification of membrane proteins

Membrane proteins were extracted for 1 h at 4°C at a concentration of 0.1 g/mL in 50 mM KPi pH 7.5, 150 mM NaCl, 10% (w/v) glycerol supplemented with 1-1.5% (w/v) n-decyl-ß-maltoside (DM, Glycon). Solubilized target proteins were purified by immobilized metal affinity chromatography (IMAC). Protein was either eluted from the Ni-NTA resin by incubating with HRV 3C protease for 1 h or overnight, or using 300 mM imidazole in appropriate buffer. In the latter case, protein-containing fractions were combined and dialyzed with a cutoff of 8 kDa (Spectra/Por® dialysis membranes, Roth) overnight in 20 volumes 50 mM KPi pH 7.5, 150 mM NaCl, 10% (w/v) glycerol, 0.2 % (w/v) DM while simultaneously cutting with 3C protease. The protease was subsequently removed by reverse NiNTA IMAC. Cleaved target protein was concentrated using 50 kDa MWCO Amicon Ultra devices (Sigma-Aldrich) and subjected to size-exclusion chromatography on Superdex 200 10/300 GL Increase column using 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM as running buffer. All steps of protein purification were carried out at 4°C. For proteins intended for cysteine-labeling in proteoliposomes, purification buffers additionally contained 5 mM ß-mercaptoethanol. Protein-containing fractions were pooled and used immediately for reconstitution or DSF (sections 3.3.14 and 3.4.10).

3.3.4. Large-scale expression and purification of E. coli uracil phosphoribosyltransferase

For expression of *E. coli* uracil phosphoribosyltransferase (UPRT), *E. coli* MC1061 transformed with pBXC3GH_UPRT was cultivated as described for expression of membrane proteins above. After overnight cultivation, cell pellets were stored at -20°C until use.

For purification, cell pellets from 1 L culture were resuspended to 100 mL using 20 mM Hepes pH 8.0, 500 mM NaCl, 2 mM $MgSO_4$ and incubated with 1 mg/mL lysozyme and trace amounts of DNAse I for 1 h at 4°C, stirring. All steps of UPRT purification were performed at 4°C. Cells were lysed using a pressure cell homogenizer (Stanstedt). PMSF was added to 1 mM and lysate was cleared by ultracentrifugation. Supernatant was incubated with 5 mL solid NiNTA for 1 h at 4°C and resin was washed with 20 CV of 20 mM Hepes pH 8.0, 500 mM NaCl, 2 mM $MgSO_4$ supplemented with 10% (w/v) glycerol and 50 mM imidazole. His-tagged protein was eluted using 20 mM Hepes pH 8.0, 500 mM NaCl, 2 mM $MgSO_4$, 10% (w/v) glycerol, 300 mM imidazole. Protein was cleaved using HRV 3C protease during dialysis against buffer without imidazole overnight. His-tagged GFP and protease were removed by IMAC, the cleaved protein was concentrated using 10 kDa MWCO Amicon Ultra centrifugation

devices (Sigma-Aldrich) and subjected to SEC on Superdex 200 30/100 GL Increase column using 20 mM Hepes pH 8.0, 500 mM NaCl, 2 mM MgSO$_4$ as running buffer. Protein-containing fractions were pooled and supplemented with 20% (w/v) ethylene glycol (Rasmussen *et al.*, 1986). 50 µM aliquots were stored at -20°C until use.

3.3.5. *SDS-Polyacrylamide gel electrophoresis*

SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to standard protocols. Samples were supplemented with 5x reducing or non-reducing sample buffer and separated applying 100 -160 V for appropriate amounts of time. For analyzing samples obtained from purifications or expression checks, 12% gels were used, for evaluation of crosslinking and dimerization 8% or 10% gels. Proteins were either visualized immediately using *in gel* fluorescence (section 3.3.6), gels were used for Coomassie or silver staining (section 3.3.7 and 3.3.8) or further processed by *anti*-His immunoblotting (section 3.3.9).

3.3.6. In gel *fluorescence for protein visualization*

In gel fluorescence was recorded at the *ImageQuant LAS 4000* using the blue channel (460 nm/Filter510DF10) for visualization of GFP fluorescence or the green one (520 nm/Filter575DF20) for RFP. For quantification, signal intensities were analyzed using ImageJ (Schindelin *et al.*, 2013).

3.3.7. *Coomassie staining*

For Coomassie staining of proteins, gels from SDS-PAGE were incubated in staining solution (10% (v/v) acetic acid, 40% (v/v) methanol, 50% (v/v) H$_2$O, 0.25% (w/v) Coomassie brilliant blue R250) for at least 1 hour at room temperature with gentle agitation. Destaining was performed either in destaining solution (10% (v/v) acetic acid, 40% (v/v) methanol, 50% (v/v) H$_2$O) or water for the appropriate amount of time.

3.3.8. *Silver staining*

For a more sensitive detection of proteins, silver staining was performed as summarized in the following protocol provided by Dr. Ying-Ning Chang:

Table 28: Silver staining protocol.

Step	Solution	Time
1. Fixation	60 mL 50% (v/v) acetone, 1.6 mL 50% (w/v) TCA, 25 µL 37% formaldehyde	5 min
2. Rinse	ddH$_2$O	3 x 5 sec
3. Wash	ddH$_2$O	5 min
4. Rinse	ddH$_2$O	3 x 5 sec
5. Pretreat	60 mL 50% (v/v) acetone	5 min
6. Pretreat	100 µL 20% (w/v) Na$_2$S$_2$O$_3$, 60 mL ddH$_2$O	1 min
7. Rinse	ddH$_2$O	3 x 5 sec
8. Impregnate	0.8 mL 20% (w/v) AgNO$_3$, 0.6 mL 37% formaldehyde, 60 mL ddH$_2$O	8 min
9. Rinse	ddH$_2$O	2 x 5 sec
10. Develop	1.2 g Na$_2$CO$_3$, 25 µL 37% formaldehyde, 25 µL 20% (w/v) Na$_2$S$_2$O$_3$, 60 mL ddH$_2$O	variable
11. Stop	1% (v/v) glacial acetic acid	30 sec
12. Store	ddH$_2$O	variable

3.3.9. anti-*His Western blotting*

For *anti*-His Western blotting, proteins were blotted to a PDVF membrane (Roti®-Fluoro PVDF membrane, pore size 0.2 µm, Roth) for 30 min at 25 V/0.1 A using the *Bio-Rad Trans-Blot Turbo* (Bio-Rad). Membranes were blocked with 2.5% (w/v) milk powder (Rapilait, Migros) in PBS buffer supplemented with 0.1% (v/v) Tween-20 (PBST) for one hour at room temperature or at 4°C overnight. After brief washing with PBST, the membranes were incubated with conjugated *anti*-His horseradish peroxidase antibody (Roche/Sigma-Aldrich) for 1 hour at room temperature. Membranes were washed 3 times for 5 min with PBST and His-tagged proteins were visualized by enhanced chemiluminescence at the *ImageQuant LAS 4000*.

3.3.10. Analytical oxidative crosslinking

For analytical oxidative crosslinking, GFP-fusion proteins of cysteine mutants were expressed in 96-well format as described in section 3.3.1 and membrane vesicles were generated in the presence of 10 mM DTT. DTT was subsequently removed using *Bio-Spin® 6* columns (Bio-Rad). Whole cell lysates were treated with 0.5 mM freshly prepared copper phenanthroline (CuPhen) and/or 0.5 mM PEG-5000 maleimide (PEG-5K) plus 1% (w/v) SDS for 30 min at room temperature. Samples were supplemented with appropriate amounts of 5x non-reducing sample buffer and proteins were separated on 8% or 10% SDS-PAGE and visualized by *in gel* GFP fluorescence.

3.3.11. Preparative oxidative crosslinking

For crosslinking and reconstitution of inward-locked SLC26Si (SLC26Si^{IL}, SLC26Si(L49C/S153C/C578S)) for cooperativity studies (section 3.4.8), proteins were purified as described in section 3.3.3, but including 5 mM ß-mercaptoethanol (BME) in all buffers. After SEC, protein-containing fractions were combined and BME was removed using PD-10 desalting columns (Bio-Rad) equilibrated with 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM in 3 mL elution volume. Yield was determined spectrophotometrically. Samples were incubated with a 30x excess of freshly prepared CuPhen for 45 min at room temperature with gentle agitation. The crosslinking reaction was stopped by adding EDTA pH 7.5 from a 0.5 M stock to a final concentration of 20 mM. Proteins were immediately used for reconstitution to soy PC as described in section 3.3.14.

3.3.12. Unspecific crosslinking using glutaraldehyde

The oligomeric state of reconstituted proteins was tested by unspecific crosslinking using glutaraldehyde (GA). For this, proteoliposomes were thawed at room temperature and incubated with different amounts of GA (0.125% - 0.5% (w/v)) for 20 min at room temperature. As a control, additional samples were pre-incubated with 1% (w/v) SDS for 10 min at room temperature before crosslinking was initiated. To stop crosslinking, GA was removed by adding Tris buffer pH 8.0. Samples were supplemented with 5x sample buffer and 2 µg of protein were applied to 8% or 10% SDS-PAGE. Proteins were visualized by silver staining (section 3.3.8).

3.3.13. Site-directed spin-labeling of cysteines for EPR studies

For EPR studies, single-cysteine mutants of SLC26Dg in a cysless background (SLC26Dg(C282S)) were generated and expressed and purified as described above (sections 3.3.2 and 3.3.3). All buffers used for purification contained 5 mM BME. After SEC, protein-containing fractions were combined and BME was removed using PD-10 desalting columns (Bio-Rad) equilibrated with 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM in 1 mL elution fractions. Yield was determined spectrophotometrically. The fractions that contained protein were combined and incubated with a final concentration of 200 µM MTSL spin label (Toronto Research Chemicals, generously provided by Dr. Benesh Joseph) from a 100 mM stock in DMSO (30-60x excess) for 45 min at room temperature with gentle agitation. Proteins were concentrated to a spin label concentration of 100 µM using 0.5 mL 50 kDa MWCO Amicon Ultra Centrifugal Filters (Sigma-Aldrich) and free spin-label was removed using Bio-Spin 6 columns (Bio-Rad). The labeling efficiency was determined by CW-EPR spectroscopy (performed by Dr. Benesh Joseph). Afterwards, single-cysteine mutants were mixed and reconstituted 1:20 (w/v) protein-to-lipid ratio to soy PC as described in section 3.3.14. Proteoliposomes were collected by centrifugation and

resuspended to 100 µM spin label concentration (~120 mg/mL lipid concentration) using 50 mM KPi pH 7.5, 2 mM $MgSO_4$. PELDOR measurements were performed by Dr. Benesh Joseph.

3.3.14. Reconstitution of proteins into liposomes

SEC-pure protein was prepared as described in section 3.3.3. Proteins were reconstituted to L-α-phosphatidylcholine from soy bean (soy PC) 1:50 (w/w) protein-to-lipid ratio for transport studies and 1:20 (w/w) for PELDOR measurements. Reconstitution was performed essentially as described previously (Geertsma *et al.*, 2008b; Chang *et al.*, 2019). In brief, protein was added to destabilized preformed soy PC liposomes in appropriate amounts and detergent was subsequently removed using Biobeads (BioBeads SM-2, Bio-Rad) in multiple incubation steps. Proteoliposomes were collected by centrifugation at ~200,000 g in Ti-60 rotor for 20 min at 15°C and resuspended to 20 mg/mL lipid weight in 50 mM KPi pH 7.5, 2 mM $MgSO_4$. For transport assays, proteoliposomes were subjected to three freeze-thaw cycles in liquid N_2/room temperature and stored in liquid N_2 or at -80°C until use. For PELDOR measurements, proteoliposomes were resuspended in 50 mM KPi pH 7.5, 2 mM $MgSO_4$ to obtain a spin label concentration of ~100 µM.

3.3.15. Reconstitution efficiency test

For all proteoliposome preparations, the efficiency of reconstitution was estimated using a solubilization test. For this, the proteins were re-solubilized from the liposomes using 1% (w/v) dodecyl-ß-D-maltoside (DDM) for 1 hour. Samples were subjected to ultracentrifugation at ~200,000 g for 10 min, 4°C using an Optima™ TLX Ultracentrifuge (Beckman Coulter). The supernatant was transferred to a fresh tube, homogenized by pipetting and protein amounts before and after ultracentrifugation were compared on gel. For this, proteins were visualized either by Coomassie or silver staining and quantified using ImageJ.

3.3.16. Fluorescence-detection size-exclusion chromatography (FSEC)

For fluorescence-detection size-exclusion chromatography (FSEC), GFP fusion proteins of SLC23 homologs were expressed in 96-well plate format as described in 3.3.1. Whole cell lysates were incubated with either 1% (w/v) DDM, 1% (w/v) DM or 2% (w/v) n-octyl-ß-D-glucoside (OG) for 1 hour on ice and subjected to ultracentrifugation at ~200,000 g for 10 min at 4°C (Optima™ TLX Ultracentrifuge, Beckman Coulter). 70 µL of supernatant were homogenized and subjected to a 96-well plate-based, automated FSEC using an Agilent 1200 HPLC system on Superdex 200 5/150 GL Increase column with 50 mM KPi pH 7.5, 150 mM NaCl, 0.02% (w/v) DDM as running buffer.

3.4. Functional assays

3.4.1. Uptake into E. coli BW25113

For *in vivo* radioactive uptake assays of uracil and xanthine, *E. coli* BW25113(*ΔuraA*) (*E. coli* JW2482) and *E. coli* BW25113(*ΔxanP*) (*E. coli* JW3629) from the Keio collection were chosen (Baba *et al.*, 2006). Both strains were purchased from the Coli Genetic Stock Center (CGSC) Strain Collection. *E. coli* BW25113(*ΔuppΔuraA*) and *E. coli* BW25113(*ΔuraAΔrutG*) were generated for this study by genomic knockout (section 3.1.6).

Cells were transformed with pBADcLIC_LacSΔIIAC320A (Geertsma *et al.*, 2005) as negative control or pBXC3GH constructs of wild type or mutant SLC23 transporters. Typically, small amounts of LB/ampicillin (5-20 mL) were inoculated with 1% (v/v) of overnight culture and incubated at 37°C either in 50 mL tubes with gas-permeable lid or 100 mL baffled flasks until OD_{600}~0.5 was reached. Cultures were induced using 10^{-3}% (w/v) L-arabinose and proteins were expressed for 1 hour at 37°C. Cells were harvested by centrifugation at 2500 g for 10 min, cells were washed twice using ice-cold 50 mM KPi buffer pH 7.2, 2 mM $MgSO_4$ and finally resuspended to OD_{600} = 120 using the same buffer. Cells were kept on ice from the harvesting step on until uptake experiments were performed. From the final samples, 1 mg total protein (assuming OD_{600} = 1 is corresponding to 0.3 mg/mL total protein) was stored at -20°C for later determination of expression levels using *in gel* GFP fluorescence of whole cell lysates.

For uptake experiments, cells were diluted 1:100 in outside buffer (50 mM NaPi pH 6.0, 2 mM $MgSO_4$, 0.1% (w/v) glucose) and incubated for 2 min, stirring, before uptake was initiated by adding substrate. If not stated otherwise, cells were incubated at 30°C (for uptake of UraA, RutG or 55PyPi and mutants thereof) or 25°C (for uptake of 51ThXi and 51ThXi mutants), stirring. For each time point, 100 μL aliquots were taken and uptake was terminated by diluting in 2 mL ice-cold stopping buffer (50 mM KPi pH 7.5, 2 mM $MgSO_4$) followed by rapid filtration over 0.45 μm nitrocellulose acetate filters. Filters were washed with additional 2 mL stopping buffer and completely dissolved in 4 mL scintillation liquid (Roth) before radioactivity associated with the filters was measured at the Hidex 300 SL scintillation counter.

3.4.2. Competition assays in whole cells

For competition assays on cells expressing wild type SLC23 transporters, samples were prepared as described above. Uptake initiated by adding 0.1 μM [5,6-^{3}H]-uracil mixed with 0.1 mM of different competitors (purchased from Sigma-Aldrich) was stopped in the linear phase of transport (after 10 sec for 51ThXi and RutG, 15 sec for UraA and 30 sec for 55PyPi, respectively). Internalized radioactivity was

determined as described above for whole cell uptake assays. For evaluation, uptake in presence of competitor was related to uptake in absence of competitor.

3.4.3. Uptake into E. coli BW25113 co-expressing UraA and UPRT

For uptake into cells overexpressing both UraA and UPRT, *E. coli* BW25113(*ΔuppΔuraA*) generated for this study (section 3.1.6) was transformed with pBXC3GH_UPRT and pTX-XH_UraA-RFP. Samples were essentially prepared as described for uptakes in the *ΔuraA*-only knockout strain (section 3.4.1), but inducing with 10^{-3}% (w/v) L-arabinose and 250 ng/mL anhydrotetracycline. Whole cell lysates of uptake samples were analyzed for both *in gel* GFP and RFP fluorescence using the blue and green channel (460 nm/Filter510DF10 and 520 nm/Filter575DF20, respectively) at the *ImageQuant LAS 4000.*

3.4.4. Preparation of proteoliposomes for functional assays

For preparation of all functional assays performed *in vitro*, proteoliposomes generated as described in section 3.3.14 were thawed at room temperature and extruded 11x through 400 nm polycarbonate filters using an Avestin LF-1 extruder. Proteoliposomes were pelleted by centrifugation at ~300,000 g for 20 min at 15°C in a TLA-110 rotor using the Optima^{TM} TLX Ultracentrifuge (Beckman Coulter) and resuspended to a final lipid concentration of 100 mg/mL using appropriate internal buffer. Samples were homogenized using a 26-gauge needle and kept at room temperature until use.

For standard experiments on 51ThXi, as internal buffer 50 mM KPi pH 7.5, 2 mM MgSO$_4$ was used, for SLC26Dg and SLC26Si 50 mM NaPi pH 7.5, 2 mM MgSO$_4$. In case another internal buffer was used, the internal buffer was exchanged by washing the liposomes two times with the respective buffer followed by three freeze-thaw cycles in liquid N$_2$/room temperature.

3.4.5. Uptake into proteoliposomes

For uptake experiments in proteoliposomes, samples prepared as described in section 3.4.4 were diluted 40x into external buffer containing radiolabeled substrate. If not stated otherwise, proteoliposomes were incubated at 10°C for experiments performed on 51ThXi and its mutants and at 30°C for SLC26Dg and SLC26Si and their derivatives. At appropriate time points, 100 µL were taken and essentially treated as described for the whole cell uptakes assays (section 3.4.1) to quantify the amount of internalized substrate. If not indicated otherwise, for uptake experiments on 51ThXi proteoliposomes, 50 mM KPi pH 7.5, 2 mM MgSO$_4$ was used on the outside, for the SLC26 transporters 50 mM KPi pH 6.0, 2 mM MgSO$_4$, 100 nM valinomycin. For uptake assays under reducing conditions,

proteoliposomes were pre-incubated with 10 mM DTT for at least 1 hour at room temperature and 10 mM DTT was included in all buffers.

3.4.6. Competition assays in proteoliposomes

For competition assays on reconstituted 51ThXi, proteoliposomes were prepared as described above. Initial transport rates up to 10 sec were recorded using a final concentration of 1 µM uracil (5% radiolabeled) mixed with 1000x excess of unlabeled competitor.

3.4.7. Counterflow, exchange and efflux assays in proteoliposomes

For counterflow, 51ThXi proteoliposomes were loaded with 10 µM unlabeled uracil for at least 30 min at room temperature. Samples were diluted 40-fold in outside buffer containing 1 µM partially radiolabeled uracil and internalized radioactivity was recorded over up to 4 min. For exchange and efflux assays, proteoliposomes were pre-incubated with 10 µM partially radiolabeled uracil and either diluted 40-fold in buffer containing no substrate (for efflux) or equal amounts of unlabeled substrate (for exchange). To investigate the substrate specificity of reconstituted 51ThXi, exchange of uracil for 5-fluorouracil, cytosine and 6-methyluracil was also tested.

3.4.8. Cooperativity experiments in proteoliposomes

Cooperativity experiments in proteoliposomes were performed on 51ThXi and SLC26Si by mixing wild type protein with an inactive mutant in different ratios and measuring the associated transport activity. Proteins were expressed and purified separately and mixed right before reconstitution. Reconstitution efficiencies for all protein mixes were monitored on gel. Substrate internalization was measured in the linear phase of transport and linear fits were performed using SigmaPlot 10.

3.4.9. Uracil phosphoribosyltransferase activity assay

To determine the activity of purified UPRT, conversion of uracil to UMP was measured adapting an uracil quantification assay described by Shibata *et al.*. When incubating uracil with benzamide oxime (BAO) and $K_3[Fe(CN)_6]$ under alkaline conditions, a blue fluorescent product of unknown nature is formed (Shibata *et al.*, 2010). For the activity assay, 0.01 µM UPRT was incubated in 25 µL total volume with 100 µM uracil, 0.5 mM PRPP, 0.1 mM GTP, 2 mM $MgSO_4$ in either 20 mM Tris or 50 mM KPi buffer at different pH for up to 30 min at room temperature. Tris buffer was used from pH 6.8 to 9.0, KPi buffer from pH 6.0 to 7.5[3]. To stop the reaction, samples were placed on ice and 25 µL of each 4 mM BAO, 8 mM $K_3[Fe(CN)_6]$ and 4 M KOH were added. Samples were incubated at 90°C for 2 min to enable

---

[3] For pH 7.0 and 7.5, UPRT activity was determined using both Tris and KPi buffer with similar outcome.

formation of the blue fluorescent product. The amount of uracil in the samples was determined by measuring fluorescence (excitation 320 nm, emission 410 nm) in a black 96-well plate at the Tecan reader. To determine background fluorescence, the assay was repeated omitting PRPP. Fluorescence was converted to uracil concentration based on an uracil standard and initial reaction rates were determined by linear fitting using SigmaPlot 10.

3.4.10. Differential scanning fluorimetry (DSF)

For differential scanning fluorimetry (DSF), 18 µM SEC-pure protein (0.5 mg/mL) in 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM was incubated with or without 1 mM of different uracil analogs (purchased from Roth) for 5 min one ice. 20 µg CPM dye (from a 1 mg/mL stock in DMSO, Thermo Scientific) was added and samples were centrifuged at 13,000 g for 5 min and 4°C. Supernatant was subjected to thermal melting in a RT-PCR cycler (Rotor-Gene Q, Qiagen). Protein was incubated from 25°C to 95°C in steps of 1°C with 10 or 20 seconds incubation per temperature step and CPM fluorescence was recorded (Excitation 380 nm, Emission 460 nm). All measurements were performed in triplicates.

3.4.11. Determination of melting temperatures using analytical SEC

To determine melting temperatures based on A_{280} in size-exclusion chromatography, 18 µM IMAC-pure protein was incubated with or without 1 mM uracil for 5 min at different temperatures from 4°C to 90°C in a PCR cycler with heated lid. Samples were centrifuged for 5 min at 13,000 g at 4°C and 70 µL of supernatant were used for 96-well plate-based, automated SEC using an Agilent 1200 HPLC system. Protein was separated on Superdex 200 5/150 GL Increase column using 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM as running buffer. For determination of melting temperatures, protein amount in the supernatant was determined by integration of protein peak areas, plotted against the respective temperature and sigmoidal fitting was performed using SigmaPlot 10.

4. Results

4.1. Characterization of two bacterial uracil transporters as highly expressed, stable UraA homologs

To identify new target proteins with superior biochemical characteristics, a homolog screening on a number of bacterial and archaeal UraA homologs was conducted. Details on the screening procedure can be found in my own master thesis that I produced under the supervision of Ass. Prof. Eric Geertsma prior to the work presented within this thesis ("Structural and functional characterization of UraA-like transporters", Goethe University Frankfurt am Main, Dec 2015).[4]

A   UraA (*Escherichia coli*)
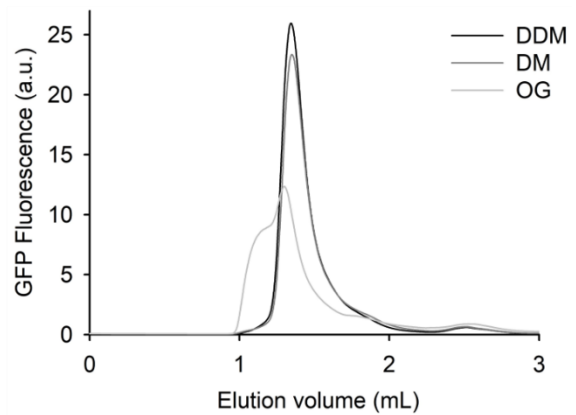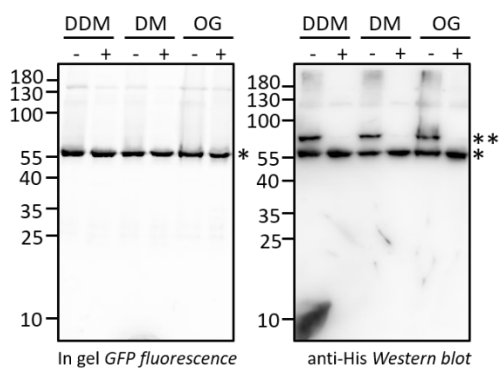
B   51ThXi (*Thalassospira xiamenensis*)

---
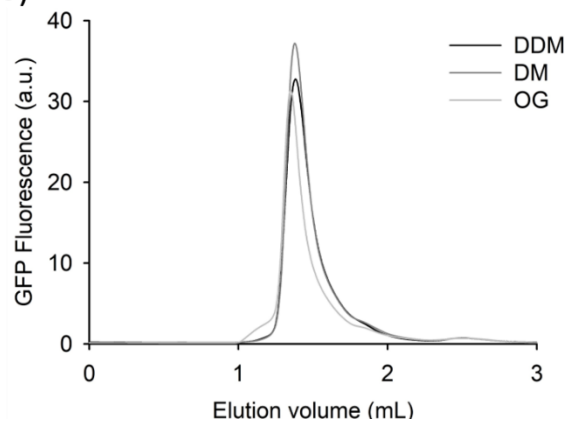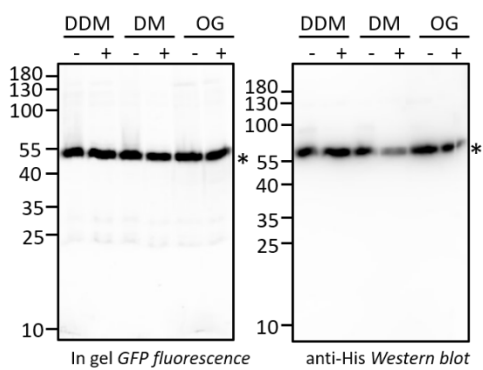
[4] The data shown here was obtained during my PhD work.

C   55PyPi (*Pyramidobacter piscolens*)
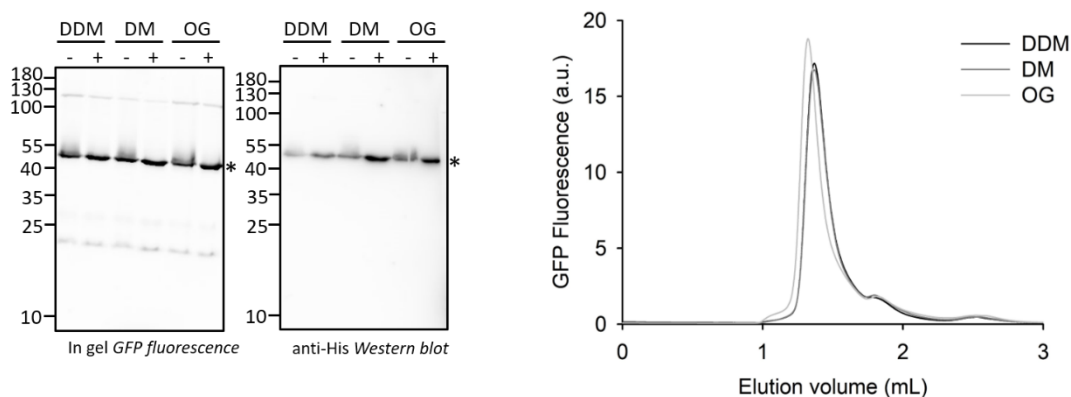
Figure 11: **Folding quality of UraA and two bacterial homologs in different detergents**. (A) Solubility of *E. coli* UraA fused to GFP was assessed by differential electrophoretic mobility (left) and FSEC (right). Well-folded, fluorescent protein is indicated by an asterisk (*), non-fluorescent, aggregated protein by two asterisks (**). Minus (-) and plus (+) indicate samples taken before and after ultracentrifugation. (B) Same as in (A), but for the UraA homolog 51ThXi from *Thalassospira xiamenensis*. (C) Same as in (A), but for the UraA homolog 55PyPi from *Pyramidobacter piscolens.* DDM, dodecyl-ß-D-maltoside. DM, decyl-ß-D-maltoside. OG, octyl-ß-D-glucopyranoside.

In brief, 60 UraA homologs were analyzed using GFP as folding indicator (Waldo *et al.*, 1999; Drew *et al.*, 2005). Protein quality was judged based on differential electrophoretic mobility of folded and unfolded GFP fusions as well as fluorescence-detection size exclusion chromatography (FSEC) (Kawate and Gouaux, 2006; Geertsma *et al.*, 2008a) . The latter was also used to test the solubility in different detergents. In the end, two bacterial UraA homologs, 51ThXi from *Thalassospira xiamenensis* and 55PyPi from *Pyramidobacter piscolens*, were selected. They show high expression levels, little aggregation as assessed by *in gel* GFP fluorescence and *anti*-His Western blotting as well as monodisperse peaks in FSEC using DDM, DM, and OG for solubilization (**Figure 11**).

Initial uptake experiments in whole cells (performed during my master thesis) revealed that both 51ThXi and 55PyPi are uracil transporters. Subsequently, for both homologs as well as for UraA assay conditions were identified that show robust time-dependent internalization of uracil into *E. coli* BW25113(*ΔuraA*) over cells expressing LacS (the lactose transporter of *Streptococcus thermophilus*, a non-related membrane protein) (**Figure 12**). These conditions were used for all *in vivo* uptake experiments described in this work, if not stated otherwise.
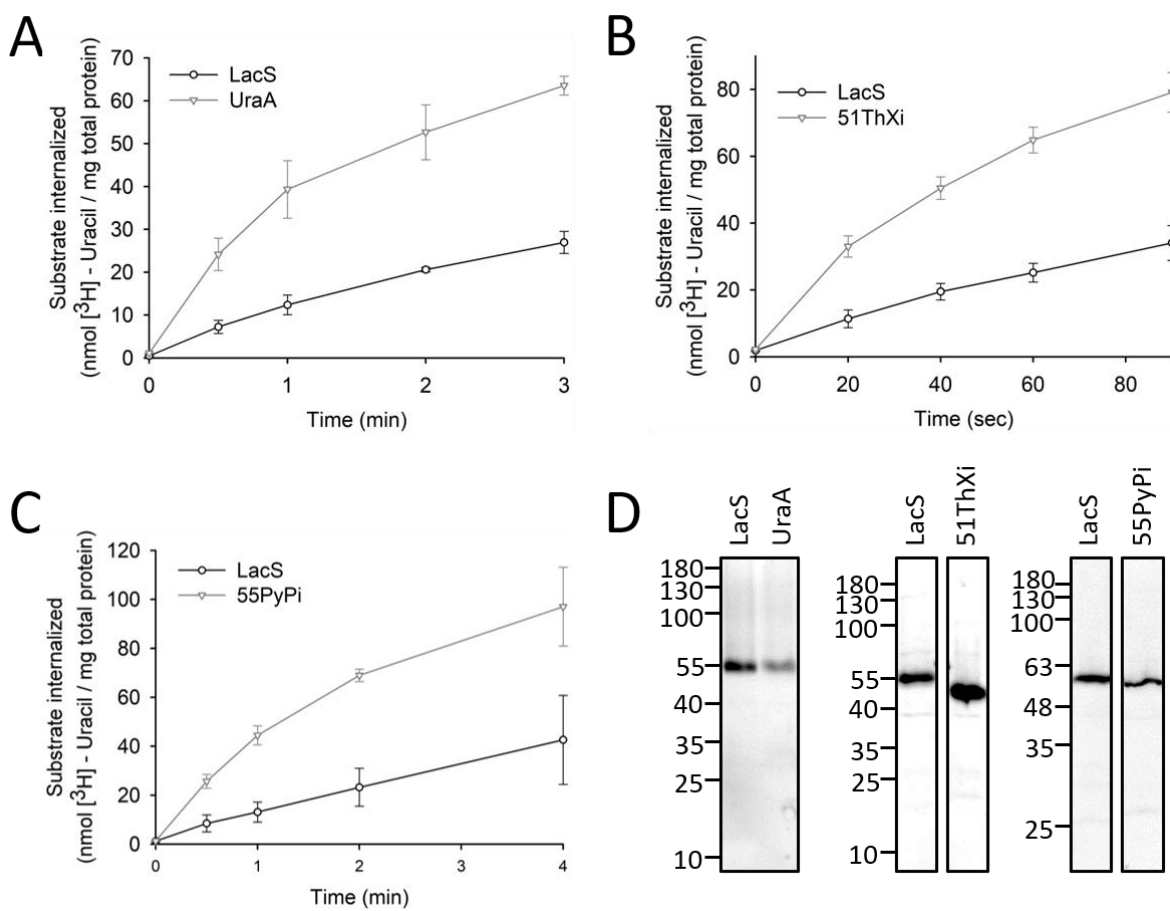
Figure 12: Optimized whole cell uptake assays for UraA, 51ThXi and 55PyPi. (A) Uracil uptake of UraA in whole cells at 30°C using 2 µM substrate. (B) Uracil uptake of 51ThXi in whole cells at 25°C using 1 µM substrate. (C) Uracil uptake of 55PyPi in whole cells at 30°C using 4 µM substrate. For all whole cell uptake experiments, *E. coli* BW25113*(ΔuraA)* was used and proteins were expressed as GFP-fusions. Uptakes were determined in triplicates, error bars shown represent standard deviations. LacS was used as a negative control. (D) Whole cell lysates of uptake samples were separated on 12% SDS-PAGE and protein expression levels were assessed using *in gel* GFP fluorescence.

4.2. Proton-coupling in SLC23 nucleobase transporters

4.2.1. Abolishment of uracil uptake in whole cells in presence of protonophore

SLC23 proteins have been suggested to being symporters, coupling the substrate transport to the proton motive force, or, in the case of mammalian SLC23 ascorbate transporters, to the sodium gradient (section 1.6). Proton-coupling in prokaryotic SLC23 transporters has been investigated exclusively *in vivo* based on radioactive uptake assays. In all cases, uptake was drastically reduced for cells treated with the protonophore CCCP. Mostly, this has been reported for the *E. coli* nucleobase transporters, in particular the uracil transporter UraA, the uracil/xanthine transporter RutG, the xanthine transporters XanQ (YgfO) and XanP (YicE) as well as UacT (YgfU), which transports both xanthine and uric acid (Karatza and Frillingos, 2005; Papakostas and Frillingos, 2012; Botou *et al.*, 2018). Also homologs from higher organisms, such as *Aspergillus nidulans*, *Candida albicans* or *Arabidopsis thaliana* showed similar behavior (Goudela *et al.*, 2005; Kosti *et al.*, 2012; Niopek-Witz *et al.*, 2014). We observe abolishment of uracil uptake in the presence of 10 µM CCCP also for the UraA homolog 51ThXi from *Thalassospira xiamenensis* (**Figure 13A**).
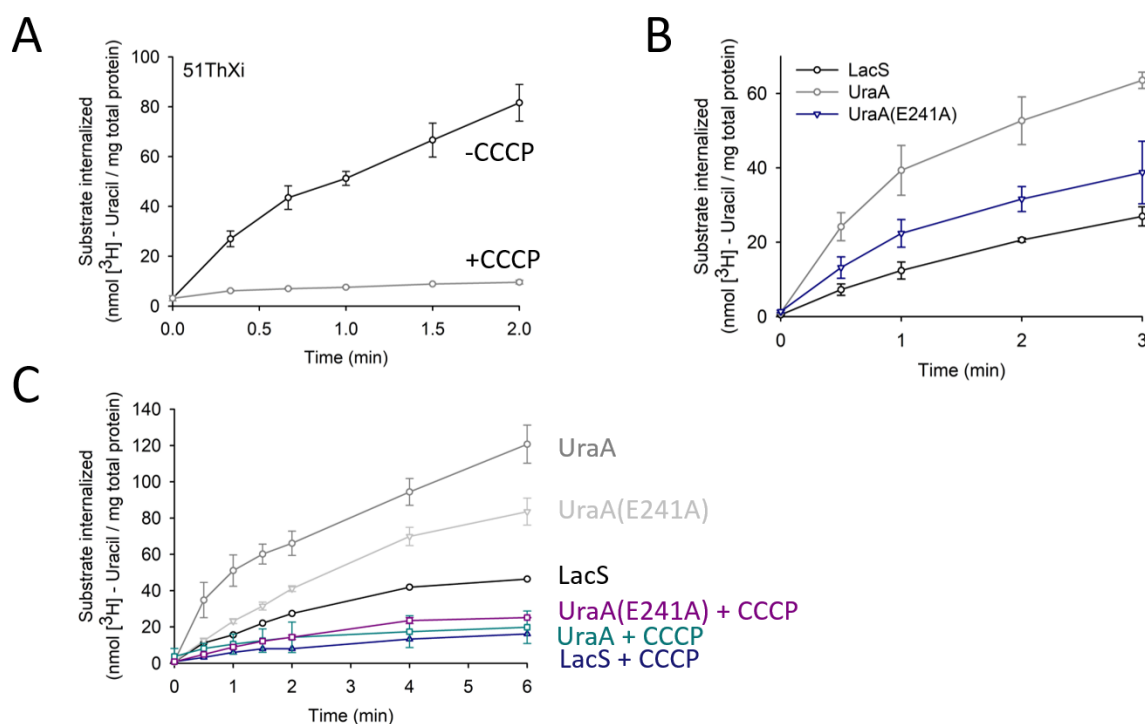
Figure 13: **Effect of CCCP on uracil transport in whole cells**. (A) Uptake of 51ThXi in whole cells is abolished when cells are pre-incubated with the protonophore CCCP. (B) Residual uracil uptake by UraA(E241A) (dark blue) compared to the LacS negative control (black). (C) Uptake of 4 µM uracil in cells expressing LacS (black), UraA (dark grey) and UraA(E241A) (light grey) over a prolonged time period in absence of CCCP, and in cells preincubated with 10 µM CCCP (LacS, dark blue; UraA, dark teal; UraA(E241A), dark pink). For all uptake experiments, *E. coli* BW25113*(ΔuraA)* was used and proteins were expressed as GFP fusions. All uptakes were performed in triplicates, except for LacS in (C) which was measured only once. Error bars represent standard deviations.

The substrate binding site of all SLC23 transports contains an extremely conserved glutamate residue located on TM8. Furthermore, also in the structurally related, but functionally unrelated SLC4 and SLC26 families, almost without exception either glutamate or aspartate is found at the corresponding position[5]. These families have been shown to accommodate both secondary active transport, using protons or sodium ions as coupling ions, as well as uniport (Chang and Geertsma, 2017). The extreme conservation of the negative charge in TM8 in all these different transporter families suggests a relevance for ion coupling. For UraA, Lu *et al.* report a complete abolishment of uracil transport for a mutant in which the TM8 glutamate is replaced by alanine (UraA(E241A)) (Lu *et al.*, 2011). In contrast, in a similar uptake experiment, we observe residual activity of UraA(E241A) compared with the negative control that might be attributed to substrate equilibration rather than accumulation (**Figure 13B**)[6]. In order to investigate the role of E241 for proton-coupling, we repeated the experiment, using a higher substrate concentration (4 µM instead of 1 µM) to increase the time period of the initial transport phase, and comparing uptake in absence and presence of CCCP (**Figure 13C**). In case E241 is responsible for proton-coupling, the activity of UraA(E241A) is expected to be unaffected by CCCP, while UraA activity would be decreased to the one of the mutant. Incubating the cells with 10 µM CCCP, however, led to a drastic reduction of uracil transportation for both UraA and UraA(E241A) to similar levels as well as, notably, for the LacS negative control. The high background uracil uptake and the effect of the protonophore on the negative control demonstrate that this kind of cell-based uptake assay in *E. coli* is unsuitable to investigate proton-coupling for SLC23 uracil transporters.

4.2.2. Effect of outside pH on whole cell uracil transport

Although a reduction of uracil transport in presence of CCCP can be observed for both UraA and 51ThXi (**Figure 13**), the effect that is observed for the LacS negative control prevents an unambiguous conclusion on proton-coupling in the SLC23 family. Therefore, the effect of different outside pH on whole cell uracil transport was probed. Assuming that SLC23 transporters are proton-coupled, on a general basis an outside buffer with pH 6.0 was used for standard whole cell uptake assays. This is expected to create a steep proton gradient over the membrane, supporting substrate accumulation. A different outside pH might therefore be able to modulate the uptake rate of uracil. Unexpectedly, different outside pH did not markedly reduce uracil uptake in cells expressing 51ThXi (**Figure 14A**), despite the drastic reduction of uptake that had been observed in the presence of CCCP that suggests proton-coupling for this protein (**Figure 13A**). Moreover, the same was observed for the LacS negative control (**Figure 14B**). This indicates that neither uracil uptake by 51ThXi nor the background uracil

---

[5] Human SLC26 members harbor a neutral amino acid at that position.
[6] UraA(E241A) was generated and kindly provided by Benedikt Kuhn.

uptake observed in cells expressing the LacS negative control is influenced by the steepness of the pH gradient over the membrane, which is more in line with transport activity that is not proton-coupled.
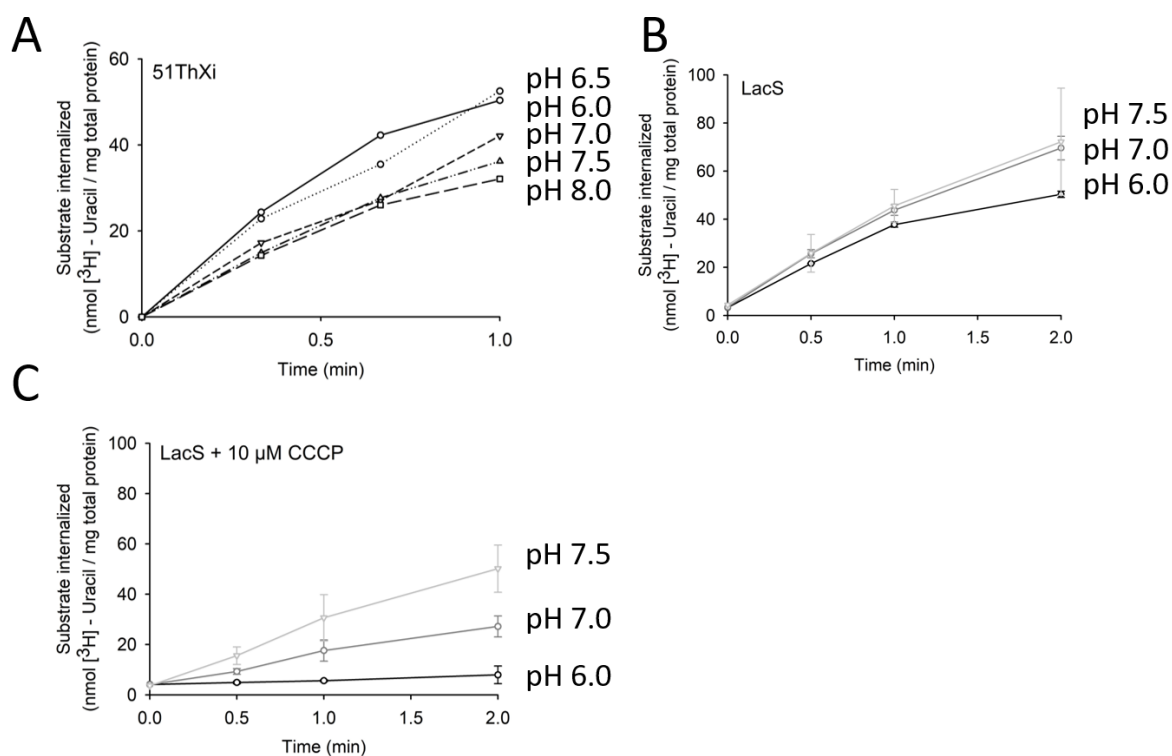
Figure 14: **Effect of different outside pH on uracil uptake in *E. coli* BW25113(*ΔuraA*)**. (A) Uracil uptake of 51ThXi in cells was tested using outside buffer with different pH (6.0-8.0). (B) Uracil uptake of cells expressing LacS-GFP using outside buffer with different pH (6.0-7.5). (C) Same as in (B), but measured in the presence of 10 µM CCCP. Uracil uptake was determined in duplicates at pH 6.0 and 7.5 (errors are mean deviations) and in triplicates for pH 7.0 (errors are standard deviations).

Unexpectedly, an increase in uracil uptake upon an increase of outside pH was observed for cells expressing LacS in presence of 10 µM CCCP (**Figure 14C**). Uptake recorded using pH 7.5 on the outside almost reached the level of uptake measured in the absence of the protonophore (**Figure 14B**). Apparently, the inactivation of uracil transport in presence of proton-protonophore is restored when the pH is increased from 6.0 over 7.0 to 7.5. This again indicates that (background) uracil uptake is not coupled to the proton-gradient and also suggests a more indirect effect of the protonophore on uracil transport in *E. coli*.

4.2.3. Overexpression increases background uracil uptake

In order to reduce the high background for uracil uptake observed in *E. coli* BW25113(*ΔuraA*) which complicates the interpretation of results from *in vivo* uracil transport experiments, we constructed the *E. coli* BW25113(*ΔuraAΔrutG*) double knockout. *E. coli* RutG was demonstrated before to transport uracil, thymine and xanthine (Botou *et al.*, 2018). Interestingly, cells overexpressing the LacS negative

control still showed elevated levels of uracil uptake in *E. coli* BW25113(*ΔuraAΔrutG*), in contrast to uninduced cells (**Figure 15**). Strikingly, lower uracil uptake levels were observed for uninduced cells regardless of whether they carried the plasmid for LacS or UraA expression. As the LacS substrate specificity includes a number of galactosides, it appears unlikely that uracil is a substrate for this transporter (Veenhoff and Poolman, 1999).
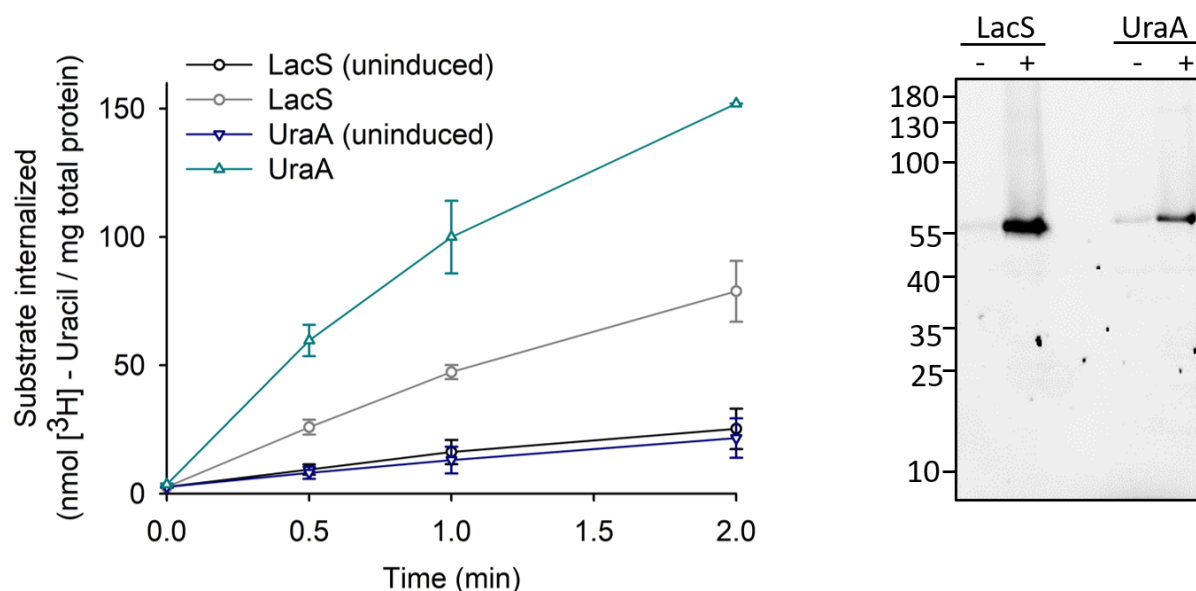
Figure 15: Background uptake in *E. coli* BW25113*(ΔuraAΔrutG)*. *E. coli* BW25113(*ΔuraAΔrutG*) cells transformed with plasmids for LacS or UraA were either not induced (black and blue) or induced with 10^{-3}% (w/v) L-arabinose and assayed for transport of uracil (light grey and teal). Uptakes were performed in triplicates, error bars shown are from standard deviations. Expression levels of proteins were determined according to *in gel* GFP fluorescence of whole cell lysates, with minus (-) and plus (+) indicating no induction and induction, respectively.

The remaining background uracil uptake in *E. coli* BW25113(*ΔuraAΔrutG*), in which both known uracil transporters have been knocked out, suggests the involvement of one or more additional endogenous *E. coli* uracil transporter. A number of SLC23 homologs have been identified in *E. coli* next to UraA and RutG based on sequence alignments and subsequent functional studies. For most of these, a narrow substrate specificity including mostly one or two main substrates have been assigned, based on whole cell uptake assays using radiolabeled substrates and competition assays. These include UacT, a uric acid transporter that also binds xanthine (Papakostas and Frillingos, 2012), two xanthine permeases called XanQ and XanP (Karatza and Frillingos, 2005), and four more distantly related ones specific for guanine/hypoxanthine (GhxQ and GhxP) and adenine (AdeP and AdeQ) (Papakostas *et al.*, 2013). The latter four share only low sequence identity with UraA (14% and 15% for GhxQ and GhxP, respectively, and 11% each for AdeP and AdeQ) according to AlignMe (Stamm *et al.*, 2014) and are not identified as SLC23 homologs using the BlastP algorithm (Altschul *et al.*, 1997). In general, SLC23 proteins are reported to have a narrow substrate specificity and to transport either pyrimidine or purine bases

(Frillingos, 2012; Chaliotis *et al.*, 2018). However, for many of them only a limited number of potential substrates were tested, mostly in competition assays that cannot report on active transport. We tested uracil uptake in *E. coli* BW25113(*ΔuraA*) for a set of proposed SLC23 transporters (**Figure 16**). Despite the reported substrate specificities, for a number of them significant uracil uptake rates were determined. It is currently unclear whether this is caused by actual uracil transportation by these homologs or by different upregulation of one (or more) other uracil transporters in response to different expression stress, or large variations in the background for different proteins based on not yet identified mechanisms. So far, all reported functional data for SLC23 uracil transporters has been derived from whole cell assays using empty plasmid as a negative control. It has been shown before that membrane protein overexpression has an impact on the membrane proteome (Wagner *et al.*, 2007). Our results suggest that background uracil uptake is also influenced by membrane protein overexpression and demonstrate the difficulty to identify a suitable negative control. This underlines the necessity for a more defined *in vitro* transport assay to study the transport mechanism of SLC23 proteins in detail.

1: LacS
2: UraA (uracil)
3: RutG (uracil, xanthine)
4: XanQ (xanthine)
5: XanP (xanthine)
6: UacT (uric acid)
7: YbbY (unknown)
8: GhxQ (guanine)
9: GhxP (guanine)
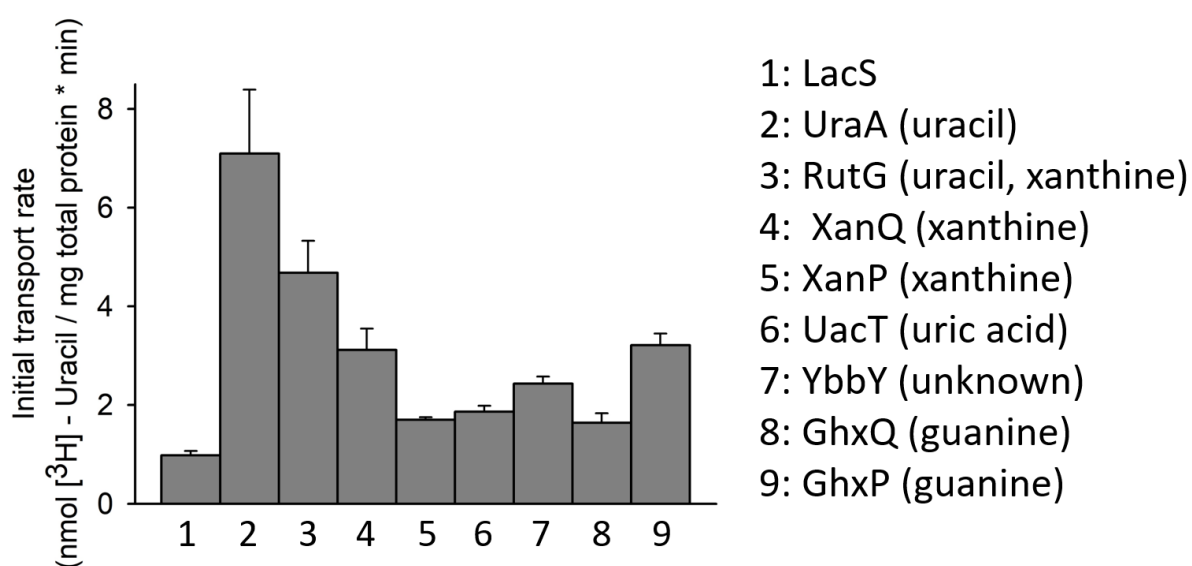
Figure 16: Uracil uptake in *E. coli* BW25113(*ΔuraA)* expressing different putative SLC23 homologs. Initial transport rates were determined by linear fitting to the initial transport phase using SigmaPlot10. Substrates as identified so far are indicated in brackets.

4.2.4. *Purification and reconstitution of UraA, 51ThXi and 55PyPi*

As outlined in section 4.2.3, uracil uptake in whole cells cannot be determined with sufficient accuracy. In order to be able to measure activity of uracil transporters in a more defined environment, *E. coli* UraA and the two UraA homologs 51ThXi and 55PyPi were purified and reconstituted 1:50 (w:w) protein-to-lipid to L-α-phosphatidylcholine from soy bean (soy PC) (**Figure 17**).
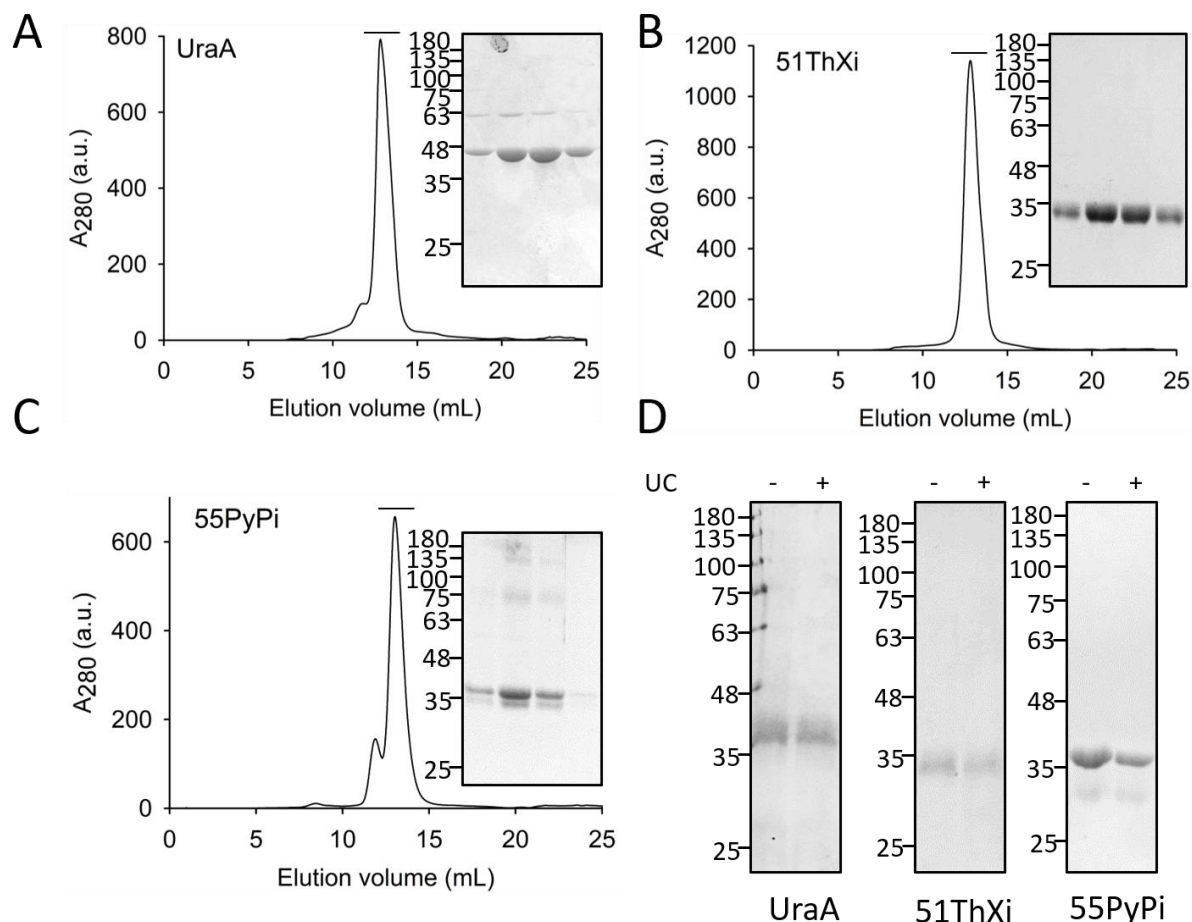
Figure 17: Purification and reconstitution of UraA, 51ThXi and 55PyPi. (A) Size-exclusion chromatogram of UraA on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. UraA-GFP was overexpressed in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) Size-exclusion chromatogram of 51ThXi on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. 51ThXi-GFP was overexpressed in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (C) Size-exclusion chromatogram of 55PyPi on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.02% (w/v) DDM. 55PyPi-GFP was overexpressed in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (D) Reconstitution efficiency tests for UraA, 51ThXi and 55PyPi reconstituted at a 1:50 (w:w) protein-to-lipid ratio in soy PC. To estimate the reconstitution efficiency, protein was resolubilized from the liposomes using 1% (w/v) DDM and protein amount before and after ultracentrifugation (UC) was compared. Samples taken before UC are indicated with a minus (-), samples taken from supernatant after UC with a plus (+). Samples were subjected to 12% SDS-PAGE and proteins were visualized by Coomassie staining.

To estimate the reconstitution efficiency, the proteoliposomes were resolubilized using 1 %(w/v) DDM and subjected to an ultracentrifugation step. Assuming that well-incorporated protein can be detergent-solubilized in contrast to not well-incorporated, aggregated protein, we conclude that all three proteins reconstituted to soy PC with high efficiency (**Figure 17D**).

4.2.5. In vitro *uracil uptake of 51ThXi suggests a uniporter function*

UraA, 51ThXi and 55PyPi were tested for *in vitro* uracil uptake. In an initial experiment applying a pH gradient, internalized radioactivity obtained for UraA and 55PyPi proteoliposomes was very low, in contrast to the one measured for proteoliposomes of 51ThXi which gave rise to extremely high signal compared to liposomes containing no protein (supplement **Figure S1**). Despite numerous efforts, no *in vitro* uracil uptake assay for UraA could be established. Neither using higher or lower uracil concentrations, inward- or outward-directed pH- or sodium gradients, membrane potential, higher or lower temperatures during uptake, nor reconstitution in different protein-to-lipid ratios (1:20, 1:50, 1:500), purification using DDM instead of DM or reconstituting to *E. coli* polar lipids instead of soy PC resulted in any measurable uracil uptake activity of reconstituted UraA, although the protein appeared to be well-incorporated into the liposomes. Similar observations were made for reconstituted 55PyPi, although fewer conditions were tested for this protein (effect of pH gradient, membrane potential, different substrate concentrations and different assay temperatures). The reason for this inactivity in proteoliposomes is still unknown. Since high internalization of radiolabeled uracil into proteoliposomes of 51ThXi was observed, this protein was subsequently used for establishing an *in vitro* uracil uptake assay.
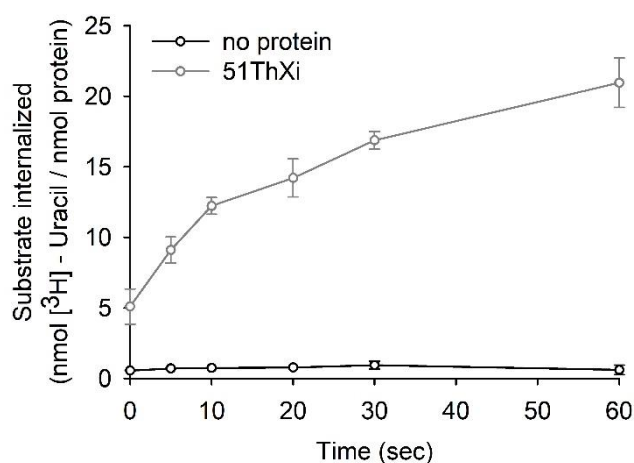
Figure 18: *In vitro* **uptake of uracil in 51ThXi proteoliposomes**. (A) Uptake of 1 µM uracil at 10°C into proteoliposomes of reconstituted 51ThXi (grey) or liposomes without protein (black). For inside and outside buffer, 50 mM KPi pH 7.5, 2 mM $MgSO_4$ was used.

Reconstituted 51ThXi showed robust internalization of uracil compared to empty liposomes (**Figure 18**). Exceptionally, assays were conducted at 10°C in order to obtain a linear phase of transportation. Higher temperatures led to immediate saturation of transport (**Figure S2**). Strikingly, initial transport rates of uracil internalization in 51ThXi proteoliposomes were substantially unaffected by a number of different driving forces tested (**Figure 19A**). Conditions tested included pH- and

sodium-gradients as well as positive and negative membrane potentials, generated by using the potassium ionophore valinomycin. As uracil is not charged under physiological conditions, proton-symport would imply an accumulation of positive charges on the inside of the proteoliposomes that would be counteracted by a negative-inside membrane potential. While transport experiments in whole cells indicated proton-coupling for 51ThXi (as uracil uptake was largely reduced in the presence of the protonophore CCCP, section 4.2.1), the *in vitro* setting suggests a uniporter function instead.
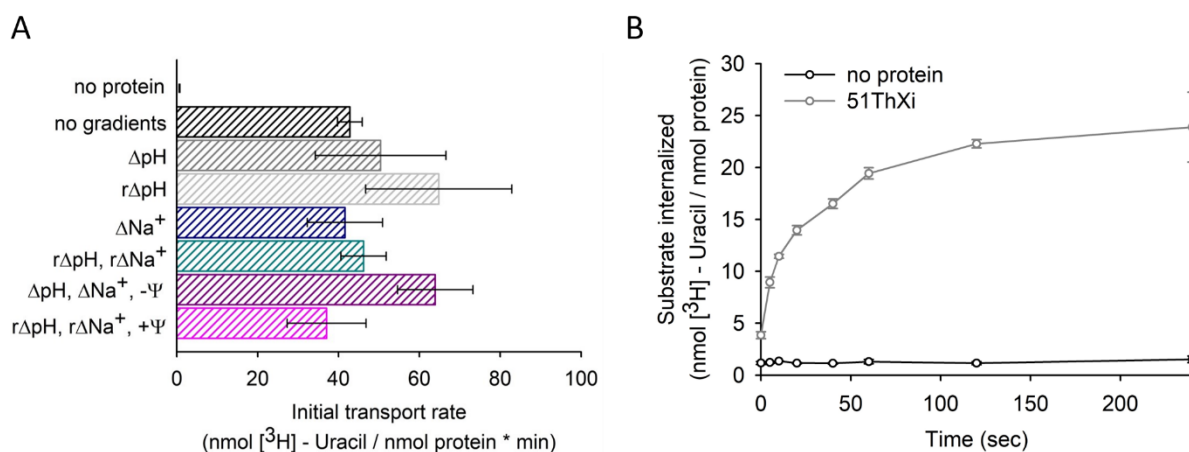
Figure 19: Uniporter function of reconstituted 51ThXi. (A) Uracil uptake of 51ThXi proteoliposomes was measured using different driving forces. An inward-directed proton gradient (ΔpH) or outward-directed pH-gradient (rΔpH) were applied using KPi buffer pH 7.5 and pH 6.0. Inward- or outward-directed sodium gradients (ΔNa$^+$ or rΔNa$^+$) were generated using combinations of equimolar sodium and potassium buffer with equal pH. -Ψ refers to an inside negative, +Ψ to an inside positive membrane potential, which were generated using K$^+$-diffusion in the presence of valinomycin. Uptakes were measured at 10°C within the linear rate of transport (up to 10 seconds) and linear fits were performed using SigmaPlot 10. (B) Counterflow transport of reconstituted 51ThXi. Internalization of 1 μM radiolabeled uracil was determined into liposomes precharged with 10 μM unlabeled uracil. Transport assays were recorded in triplicates, errors represent standard deviations.

The immediate saturation of radioactivity associated with the proteoliposomes at temperatures above 10°C and the low level of substrate accumulation at lower temperatures might also be interpreted as substrate binding instead of transport. However, the internalized radioactivity at saturation was at least 10-fold higher than the number of expected binding sites, which is one per protomer, following the structures of UraA and UapA. Another indication for transport instead of binding can be derived from a counterflow transport experiment: Substrate transport that is not coupled to a secondary gradient can only lead to equilibration instead of accumulation. Counterflow, in which the liposomes are pre-charged with unlabeled substrate, will eventually also lead to substrate equilibration but might exhibit slower initial rates (Xie, 2008). Indeed, counterflow transport of reconstituted 51ThXi showed a prolonged initial phase with saturation at similar levels as observed for uracil uptake (**Figure 19B**).

4.2.6. Translocation of the empty carrier as the rate-limiting step

In alternating access mechanisms, transport proteins facilitate substrate translocation by exposing the substrate binding site alternatingly to the two sides of the membrane. As shown in section 4.2.5, 51ThXi likely operates by mediating facilitative diffusion of uracil without coupling to a secondary substrate. In general, a tightly coupled transporter can undergo the transition between inward- and outward-facing conformation either fully loaded, or empty. Depending on the number of substrates, multiple other conformations can be sampled. In the case of 51ThXi, only one substrate is transported, resulting in the simple reaction scheme shown in **Figure 20A**. By comparing the efflux to the exchange rate of a transporter, the rate-limiting step of transportation can be determined. If the re-localization of the empty transporter is limiting, efflux will be slow compared to exchange, in which substrate is present at both sides of the membrane. For measuring efflux and exchange of reconstituted 51ThXi, proteoliposomes were pre-loaded with 10 µM partially radiolabeled uracil. For efflux, proteoliposomes were diluted in outside buffer containing no additional substrate, whereas for exchange, the same concentration of unlabeled uracil was present on the outside. Under these conditions, substrate exchange was considerable faster than efflux, which is consistent with a mechanism in which re-localization of the empty carrier is rate-limiting (**Figure 20B)**.
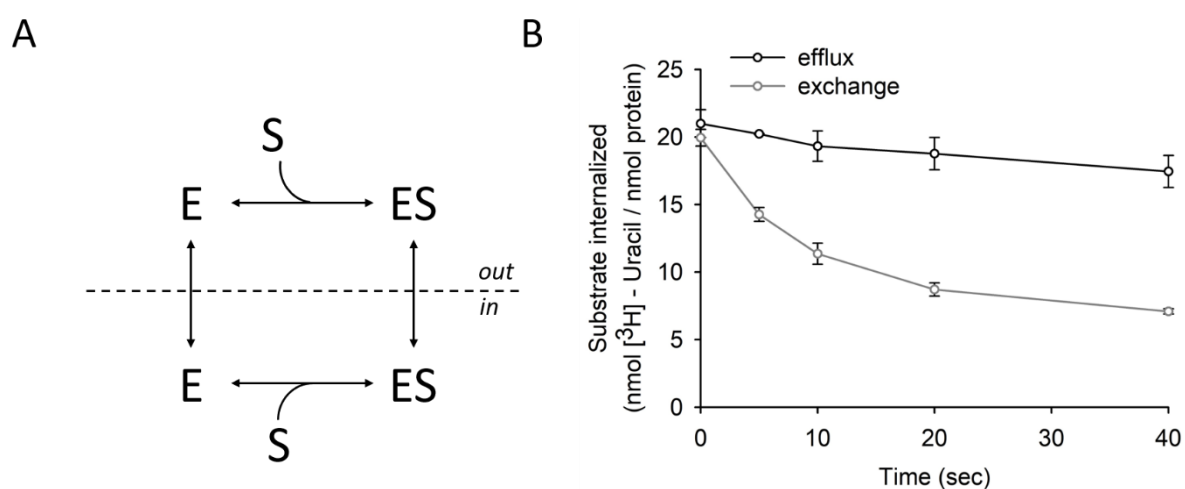
Figure 20: **Substrate efflux and exchange in 51ThXi proteoliposomes.** (A) Schematic transport cycle for an alternating access uniporter. The transporter (E) can switch between inside and outside conformation either empty, or loaded with substrate (S). (B) Comparison of uracil efflux (black) and exchange (grey) of reconstituted 51ThXi. In both cases, the proteoliposomes were preloaded with 10 µM partially radiolabeled uracil. For efflux, proteoliposomes were diluted 40x in buffer containing no substrate, for exchange in buffer containing an equal concentration of unlabeld uracil. Experiments were performed at 10°C in triplicates, errors shown are from standard deviations.

4.2.7. *Activity of uracil phosphoribosyltransferase is crucial for uracil uptake in E. coli*

The uniporter function of 51ThXi demonstrated in proteoliposomal transport assays is in stark contrast to the observed apparent proton-dependent uracil uptake in whole cells. To dissolve this contradiction, we analyzed the *E. coli* uracil metabolism in more detail. Uracil uptake in *E. coli* was shown before to depend on the activity of uracil phosphoribosyltransferase (UPRT) (Andersen *et al.*, 1995). The activity of several UPRT paralogs has already been shown to be strictly pH-dependent. Additionally, it has been demonstrated before that the cytoplasmic pH of *E. coli*, when being exposed to acidic pH, initially drops but is rapidly recovered to neutral; when the cells are pre-incubated with a protonophore, however, it will stay close to the pH of the outside buffer (Wilks and Slonczewski, 2007; Martinez *et al.*, 2012). In section 4.2.2 we showed that uracil transport is reduced in presence of CCCP only when an acidic outside pH is used, but not at pH 7.5. We therefore explored whether an inactivation of UPRT due to low outside pH in combination with a protonophore could cause a phenotype in whole cell uracil uptake assays that mimics the phenotype of a proton-symporter.

The *uraA* gene of *E. coli* was identified based on selection of mutants that showed resistance towards 5-fluorouracil (5FU) (Andersen *et al.*, 1995). 5FU becomes toxic to the cell when it is metabolized to 5-fluorouridine monophosphate, the 5-fluorinated form of dUMP. As such, it is a potent inhibitor of thymidylate synthase by forming a covalent complex with the enzyme, and thus is commonly used in anticancer treatments (Sadee and Wong, 1977). Inhibition of thymidylate synthase is leading to accumulation of dUMP and subsequent misincorporation of dUTP instead of dTTP into DNA. Finally, frequent DNA double strand breaks caused by misincorporation of nucleotides and aberrant repair mechanisms will induce apoptotic cell death. This mechanism is also referred to as "thymineless cell death" (El-Awady *et al.*, 2010; Gmeiner, 2012; Mani *et al.*, 2018).

The first step of uracil metabolism in most organisms is the formation of uridine monophosphate (UMP) by glycosylation at N1 of uracil using the high-energy compound phosphoribosyl pyrophosphate (PRPP) (Rasmussen *et al.*, 1986) (**Figure 21**). This reaction is catalyzed in *E. coli* by uracil phosphoribosyltransferase (UPRT) that is encoded by the *upp* gene. UPRT homologs are found in many lower eukaryotes and microorganisms, whereas in mammals, a similar function is carried out by the bifunctional UMP synthase (Lundegaard and Jensen, 1999; Li *et al.*, 2007). Andersen *et al.* found *E. coli* mutants with either defective *uraA* or *upp* genes both resistant towards 5FU, suggesting that this compound is a substrate for both proteins (Andersen *et al.*, 1992, 1995). In the same studies, the two genes were shown to map together in one bicistronic operon, with *upp* upstream of *uraA*. UraA was identified as an inner membrane uracil transporter. Notably, strains lacking the *upp* but with intact *uraA* showed very reduced levels of uracil uptake in whole cells. The importance for uracil

phosphoribosyltransferase activity for uracil uptake had in fact already been suggested long before UraA, the actual uracil transporter, was identified (Beck *et al.*, 1972; Burton, 1977).
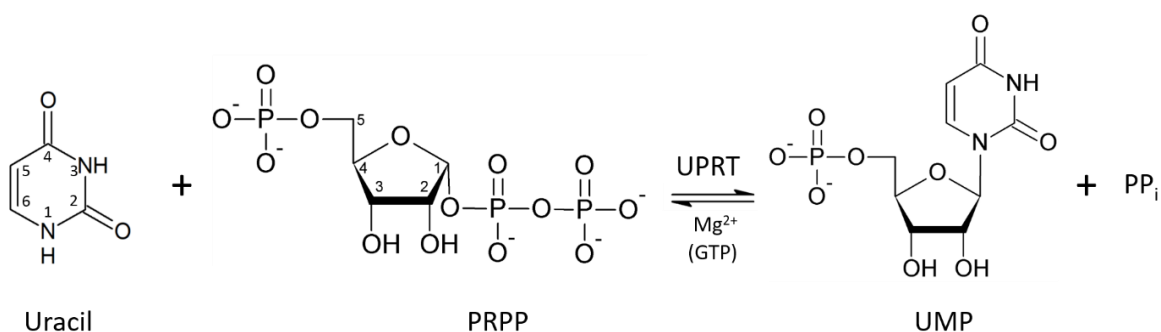
Figure 21: Formation of uridine monophosphate (UMP) from uracil and phosphoribosyl pyrophosphate (PRPP). Uracil phosphoribosyltransferase (UPRT) catalyzes the glycosylation of uracil at N1 in a reaction that is depending on divalent cations (typically Mg^{2+}) and allosterically enhanced by GTP (Jensen and Mygind, 1996).

In order to elucidate the role of the cytosolic uracil phosphoribosyltransferase for uptake by UraA, we generated a double knockout of both genes in *E. coli* BW25113 (*E. coli* BW25113(*ΔuppΔuraA*). *E. coli* BW25113 (F⁻ *Δ(araD-araB)567 ΔlacZ4787 (::rrnB-3) λ⁻ rph-1 Δ(rhaD-rhaB)568 hsdR514*) was chosen as it is the parental strain of the Keio collection and the respective *uraA* knockout (*E. coli* JW2482) (Baba *et al.*, 2006). Since *upp* and *uraA* are located in one operon separated by less than 100 nucleotides, both genes were replaced as a whole with a single recombination step.
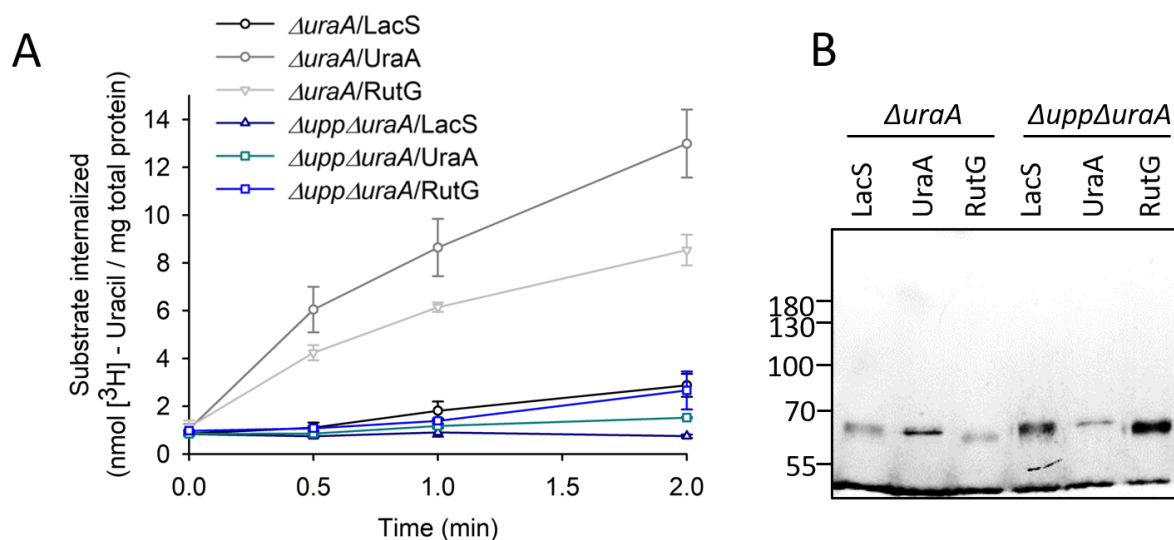
Figure 22: Uracil uptake into *E. coli* BW25113(*ΔuppΔuraA*) expressing GFP fusions of different uracil transporters. (A) Comparison of uptake of 0.5 µM uracil in *E. coli* BW25113(*ΔuraA*) or *E. coli* BW25113(*ΔuppΔuraA*) expressing LacS (black and dark blue), UraA (dark grey and teal) or RutG (light grey and blue). (B) Expression levels of samples used for whole cell uptake assays shown in (A). Whole cell lysates were separated on 12% SDS-PAGE and proteins were visualized by *in gel* GFP fluorescence.

The *E. coli* BW25113(*ΔuppΔuraA*) double knockout was transformed with pBXC3GH constructs of different uracil transporters and tested for uptake activity (**Figure 22A, Figure S3**). For all whole cell assays presented in this thesis, cells expressing LacS (the lactose transporter of *Streptococcus thermophilus*) from araBAD promoter were used as a negative control. *E. coli* BW25113(*ΔuraA*) overexpressing LacS was found to show increased uracil background uptake compared to cells not overexpressing any (membrane) protein, presumably due to upregulation of other endogenous uracil transporters (section 4.2.3). One known endogenous *E. coli* uracil transporter is RutG that also transports xanthine (Botou *et al.*, 2018). Next to uptake of cells overexpressing LacS, activity of UraA and RutG was compared in *E. coli* BW25113(*ΔuraA*) and *E. coli* BW25113(*ΔuppΔuraA*). Notably, although all proteins were expressed (**Figure 22B**), almost no uracil uptake was observed for neither of the uracil transporters in cells lacking *upp*, compared to robust substrate internalization for both UraA and RutG in the *ΔuraA* only strain. This not only confirms the importance of UPRT for uptake by UraA, but also emphasizes a general relevance of UPRT for uracil uptake in *E. coli*.

In case UPRT activity is required for uptake of uracil in *E. coli*, overexpression of the enzyme from plasmid should be able to restore uracil internalization in a *Δupp* strain. To test this, both UraA and UPRT were co-expressed in *E. coli* BW25113(*ΔuppΔuraA*). To tightly control expression levels of both proteins as well as to enable unambiguous identification on gel, UPRT was expressed as GFP fusion from araBAD promoter and UraA was fused to RFP and expressed from a tetracycline-inducible promoter. Both plasmids contained compatible origins of replication (ColE1 and p15A, respectively) and antibiotic markers. Anhydrotetracycline was used for induction instead of tetracycline, as this compound is not an antibiotic for *E. coli* (Skerra, 1994). High levels of UPRT-GFP overexpression were seen when cells were induced with L-arabinose. Expression of UraA-RFP was independent of different anhydrotetracycline concentrations tested and only mildly affected by simultaneous overexpression of the soluble UPRT-GFP (**Figure 23**).
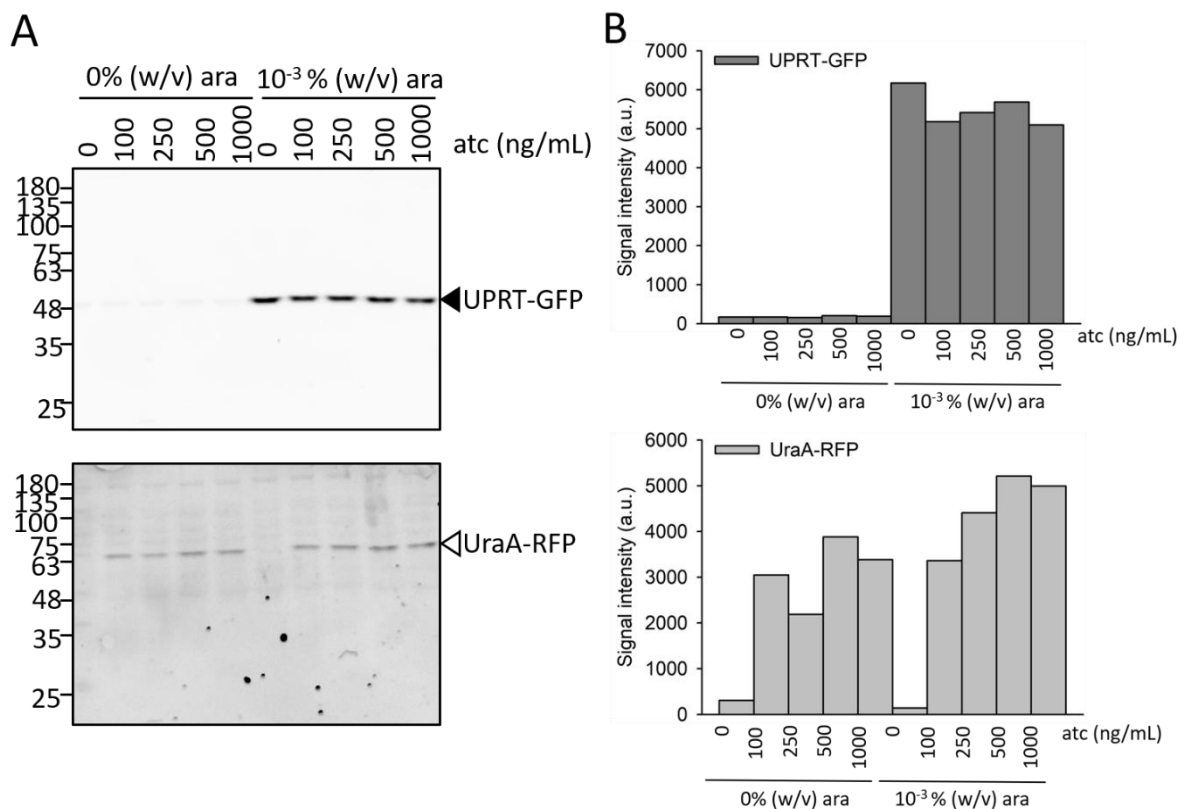
Figure 23: Co-expression of UPRT-GFP and UraA-RFP in *E. coli* BW25113(*Δupp ΔuraA*). (A) Whole cell lysates of *E. coli* BW25113(*ΔuppΔuraA*) pBXC3GH_UPRT/pTX-XH_UraA-RFP induced with different concentrations of L-arabinose (ara) and anhydrotetracycline (atc) for 2 hours were subjected to 12% SDS-PAGE and proteins were visualized by *in gel* fluorescence using the blue and green channel (460 nm/Filter510DF10 and 520 nm/Filter575DF20), respectively. Illumination for 2 min for visualization of UPRT-GFP, for 8 min for UraA-RFP. (B) Signal intensities from gels shown in (A) were quantified using ImageJ. Experiment was performed on two different colonies with similar outcome.

Background expression was also tested in a radioactive uracil uptake experiment, comparing substrate internalization in *E. coli* BW25113(*ΔuppΔuraA*) transformed with only one, or with both expression plasmids (**Figure 24A**). Uptake levels of cells transformed with only pBXC3GH_UPRT or both expression plasmids (but overexpressing only one of the proteins) were considerably higher than in cells carrying only the plasmid for UraA overexpression (**Figure 24A**). This enhanced uracil uptake is most probably caused by background activity of UPRT and reflects its relevance for *in vivo* uracil uptake. Consequently, internalization of uracil was drastically increased in *E. coli* BW25113(*ΔuppΔuraA)* overexpressing both UPRT and UraA compared to the same strain overexpressing only one of these proteins (**Figure 24B**). Uptake rates were also significantly higher than what was observed in any uracil uptake experiments performed in *E. coli* BW25113(*ΔuraA*) harboring the endogenous gene for UPRT. This suggests that the rate of phosphoribosylation by UPRT, rather than the transport rate of UraA, might be the limiting step of uracil uptake *in vivo*.
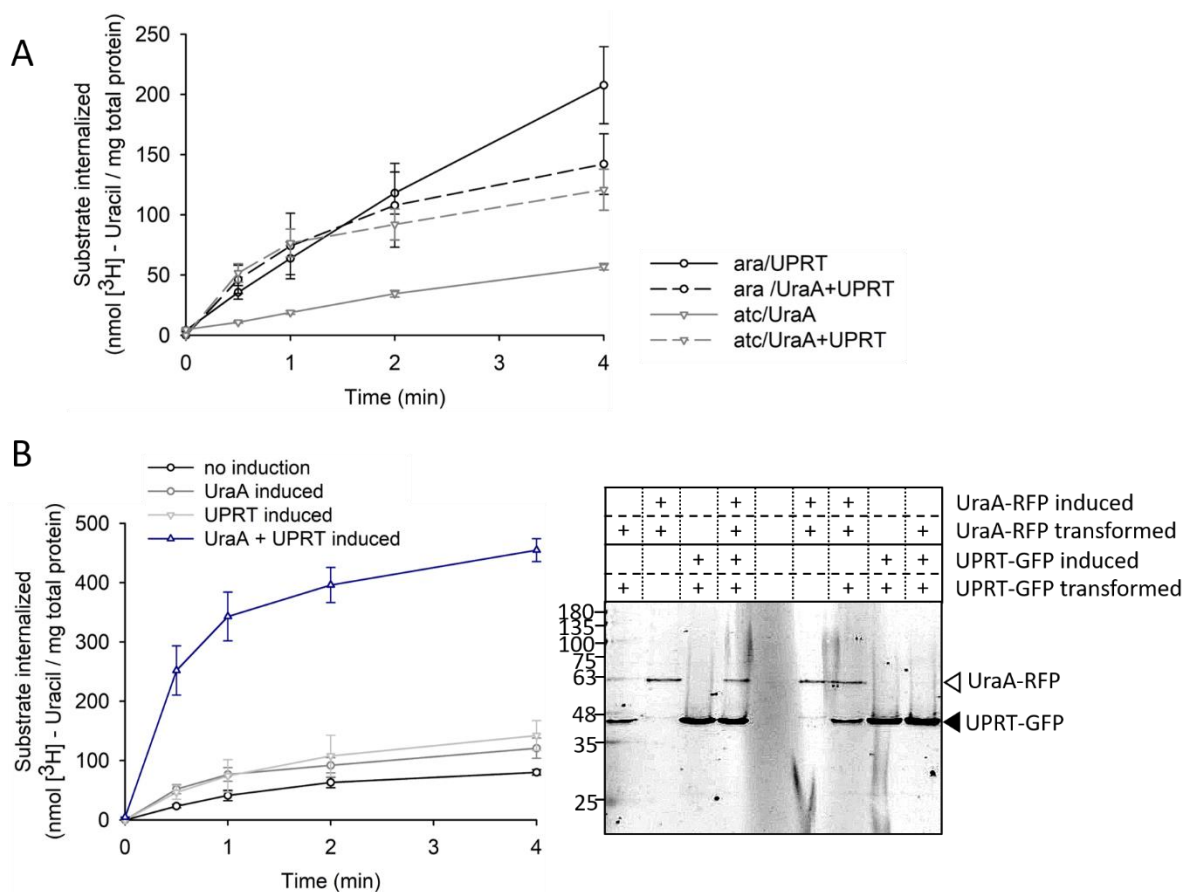
Figure 24: Uracil uptake in *E. coli* BW25113 *Δupp ΔuraA* overexpressing both UPRT and UraA. (A) Uracil uptake in *E. coli* BW25113(*ΔuppΔuraA*) transformed with either only pBXC3GH-UPRT (solid black line), only pTX-XH_UraA-RFP (solid grey line) or both expression plasmids (dashed lines). Cells were induced either only with L-arabinose (ara) (circles) or only with anhydrotetracycline (triangles). (B) Uracil uptake in *E. coli* BW25113(*ΔuppΔuraA*) transformed with pBXC3GH_UPRT and pTX-XH_UraA-RFP. Cells were either not induced (black), induced with atc (UraA induced, dark grey), induced with ara (UPRT induced, light grey), or induced with both atc and ara (UraA + UPRT induced, blue). Expression levels were determined in whole cell lysates of uptake samples from cells transformed with only pBXC3GH_UraA, pTX-XH_UraA-RFP, or both, with and without adding inducers, as indicated by a plus (+). Samples were subjected to 12% SDS-PAGE and proteins were visualized by *in gel* fluorescence using the blue (460 nm/Filter510DF10) and green (520 nm/Filter575DF20) channels. Gels were merged using ImageJ to illustrate co-expression of both proteins.

4.2.8. Trapping of uracil as UMP inside the cell

Human glucose transporters of the SLC2 protein family (GLUT1 - GLUT14) mediate facilitative diffusion of metabolites, mostly hexoses, over cellular membranes. They transport substrates down the concentration gradient, but have also been shown to mediate exchange *in vitro* once the equilibrium is reached. *In vivo*, the sugars are rapidly phosphorylated, e.g. by glucokinase or hexokinase, and as a result are no longer recognized by the GLUT transporters. This way, a steep concentration gradient is maintained. This mechanism is essential for example for ensuring rapid glucose uptake from the bloodstream into various tissues as a response to cellular stimuli. In most cases, transport of glucose is the rate-limiting step and glucose uptake Is mostly regulated over expression levels of the GLUTs. For GLUT2, however, it has been shown that downstream phosphorylation by glucokinase governs the

overall glucose influx as the first step of glycolysis in liver and pancreas (Printz *et al.*, 1993; Uldry and Thorens, 2004; Mueckler and Thorens, 2013; Sternisha and Miller, 2019; Holman, 2020).

To explore the hypothesis that phosphoribosylation of uracil by UPRT might serve a similar role by trapping uracil as UMP inside the cell, we tested whether UMP is a substrate for UraA. A "nucleic acid base transport metabolon" including these two enzymes was already suggested in 2012 by Moraes and Reithmeier (Moraes and Reithmeier, 2012). Cells to which an excess of unlabeled uracil was added after incubation with labeled substrate for a certain amount of time ("cold chase" experiment) fail to expel the radiolabel, meaning the labeled substrate is not exchanged for the unlabeled one (**Figure 25A**). This could be taken as a hint that indeed uracil is modified into a metabolite that is no longer a substrate for UraA, for example UMP. UMP was found to neither compete for uracil in whole cells, nor was it stabilizing purified UraA in a thermal shift experiment (**Figure 25B**). This is in line with a potential mechanism in which internalized uracil is rapidly converted to UMP by UPRT, which then in turn would remain inside the cell as it is no longer a substrate for UraA.
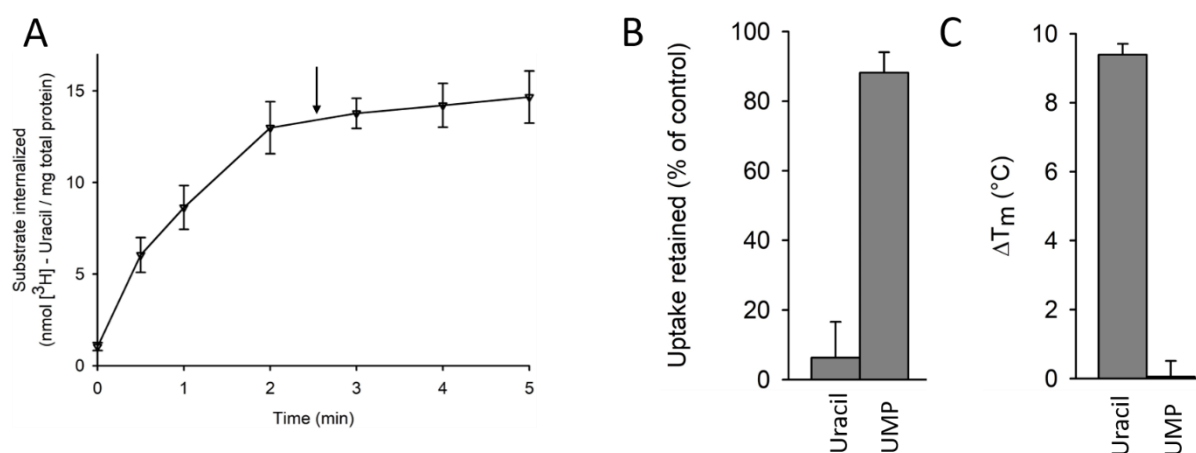
Figure 25: UMP is not a substrate for UraA. (A) "Cold chase" experiment on *E. coli* BW25113(*ΔuraA*) overexpressing UraA-GFP. Cells were incubated at 30°C with 0.5 μM uracil. After 2.5 min, indicated by an arrow, 1 mM unlabeled uracil was added from the outside. (B) UMP does not compete for uracil in whole cell competition assays. Internalization of 0.1 μM uracil in cells expressing UraA-GFP within 15 seconds was measured in presence of 1000x excess of unlabeled uracil or UMP. Amount of internalized substrate in presence of competitor was normalized to the one in absence of competitor. (C) No stabilizing effect of UMP on purified UraA is observed in a thermal shift experiment. 0.5 mg/mL UraA was incubated with 1 mM uracil or UMP and the protein melting temperature was determined by differential scanning fluorimetry monitoring CPM dye fluorescence change over temperature increase. Melting temperatures in presence of uracil or UMP were compared to melting temperature obtained of protein that was not pre-incubated with any substrate to determine ΔT_m in °C.

The "nucleic acid base transport metabolon" as it was suggested by Moraes and Reithmeier implies a physical interaction between the two enzymes that would allow "channeling" of the substrate from the transporter to the modifying enzyme. Genes for uracil phosphoribosyltransferases close to the

ones for putative uracil permeases can be found for many organisms. For a third of the 36 bacterial and archaeal genomes that were used for the homolog screening preceding the work for this thesis (section 4.1), a UraA homolog was found close to a potential homolog of UPRT (supplement **Table 30**). Although gene association does not imply protein association, this might indicate that the mechanism of uracil conversion to UMP is not unique to *E. coli*. The importance of UPRT activity for uracil uptake, however, was not restricted to UraA, but also found for RutG, another *E. coli* uracil transporter, and 51ThXi and 55PyPi, UraA homologs from other organisms (**Figure S3**). This opposes a direct physical association between the transporter and UPRT and points to a more general role of uracil modification by UPRT, or possibly also other uracil modifying enzymes, for uracil uptake *in vivo*. In line with this notion, no conserved patches could be identified on UPRT or the uracil transporters that could function as a potential interaction platform.

4.2.9. *Activity of* Escherichia coli *uracil phosphoribosyltransferase is pH-dependent*

In section 4.2.7, evidence for the importance of the activity of uracil phosphoribosyltransferase (UPRT) for uracil uptake in *E. coli* has been presented. Activities of UPRTs from other organisms have been reported to be strongly pH-dependent (Linde and Jensen, 1996; Turner *et al.*, 1998; Jensen *et al.*, 2005; Soysa *et al.*, 2013). For *E. coli* UPRT, the pH optimum has been suggested to be 7.5 - 8.5, although direct experimental proof for this has not been provided (Rasmussen *et al.*, 1986). It also has been reported that *E. coli* cells exposed to acidic pH will recover to neutral luminal pH within a short time frame, but not when they are pre-incubated with protonophore. In that case, the inside pH will stay close to that on the outside (Wilks and Slonczewski, 2007; Martinez *et al.*, 2012). Assuming that indeed UPRT activity is governing uracil uptake in *E. coli* and UPRT is inactivated at low pH, this might pose an explanation to the contradictory results presented in sections 4.2.1 and 4.2.2 on the reduced uracil uptake in whole cells in presence of CCCP.

In order to investigate the pH-dependency of *E. coli* UPRT, the enzyme was overexpressed as C-terminal GFP-fusion protein with 10x His-tag and purified by IMAC and SEC. High yields of monodisperse, pure protein were obtained (**Figure 26A**). Notably, using buffer with pH 8.0 for the purification improved both protein yield and quality compared to using buffer with pH 7.5. As an initial quality control, the melting temperature of the purified protein was determined in a differential scanning fluorimetry (DSF) experiment (**Figure 26B**).
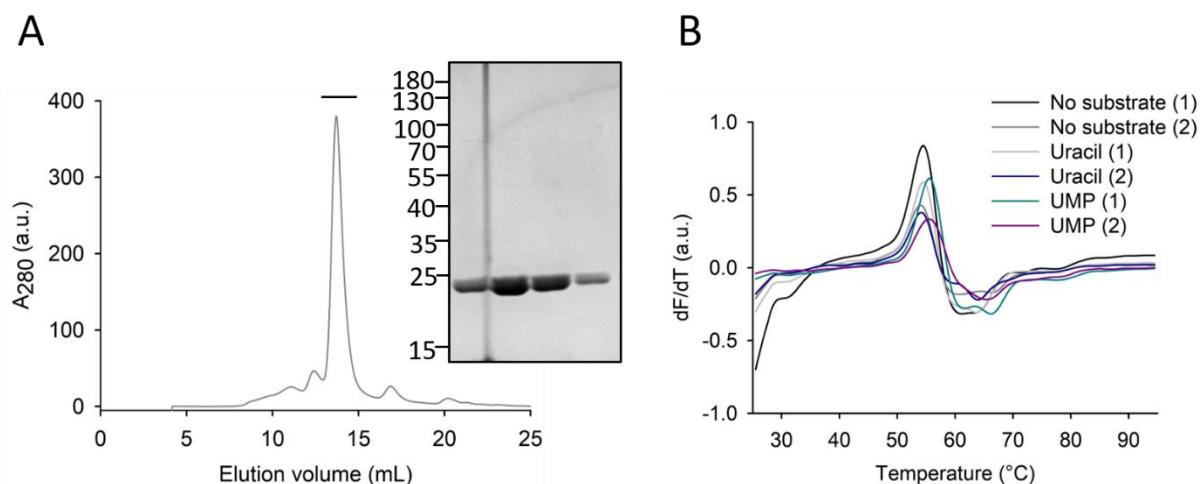
Figure 26: Purification and initial characterization of *E. coli* uracil phosphoribosyltransferase. (A) Size-exclusion chromatogram of *E. coli* UPRT. The protein was overexpressed with C-terminal GFP and 10x His-tag and purified by IMAC and SEC. Protein was subjected to 12% SDS-PAGE and size and purity were visualized by Coomassie staining (inset; M_W = 23.4 kDa). (B) Melting temperature determination of purified UPRT. 0.8 mg/mL SEC-pure protein were incubated without substrate, with 1 mM uracil or 1 mM UMP for 20 min on ice and change of fluorescence of SYPRO Orange was recorded upon temperature increase. Samples were heated from 25°C to 95°C in steps of 1°C with 2 min incubation per degree. The experiment was done in duplicates.

The recorded melting temperature (~54.5°C) is close to the one reported by Lundegard and Jensen that was determined by differential scanning calorimetry (DSC) (57.5°C) (Lundegaard and Jensen, 1999). A small shift (+1.5°C) was observed when preincubating with UMP, but not with uracil. This is in line with the suggested higher affinity of UPRT for UMP over uracil (Linde and Jensen, 1996).

To overcome the use of radioactive uracil (Rasmussen *et al.*, 1986; Ghim and Neuhard, 1994; Jensen and Mygind, 1996; Linde and Jensen, 1996; Carter *et al.*, 1997; Turner *et al.*, 1998; Barchue *et al.*, 1999; Lundegaard and Jensen, 1999; Jensen *et al.*, 2005) and non-quantitative procedures (Carter *et al.*, 1997; Li *et al.*, 2007; Yata *et al.*, 2012; Soysa *et al.*, 2013) for determining UPRT activity, a novel assay was developed. This assay is based on a uracil quantification reaction described by Shibata *et al.* (Shibata *et al.*, 2010). Uracil will form a blue fluorescent product when being incubated at 90°C with benzamide oxime (BAO) and potassium hexacyanoferrate ($K_3[Fe(CN)_6]$) under alkaline conditions. The nature of the blue fluorescent product is unknown, but it can be quantified by measuring emission at 410 nm (excitation 320 nm). Importantly, the assay is insensitive for UMP since no blue fluorescence can be measured when UMP is used instead of uracil (**Figure 27A**). It was also tested whether the other components required for the enzymatic reaction, in particular PRPP or GTP, would interfere with the uracil quantification reaction (**Figure 27B**). No additional fluorescence was recorded. The concentration of PRPP used in the assay (0.5 mM) is well above the reported K_M of 73 ± 19 µM (Lundegaard and Jensen, 1999). GTP activates UPRT by lowering the binding affinity for PRPP; the K_M

for GTP has not been determined, but concentrations as used herein (0.1 mM) are in line with other reported UPRT activity assays (Jensen and Mygind, 1996).
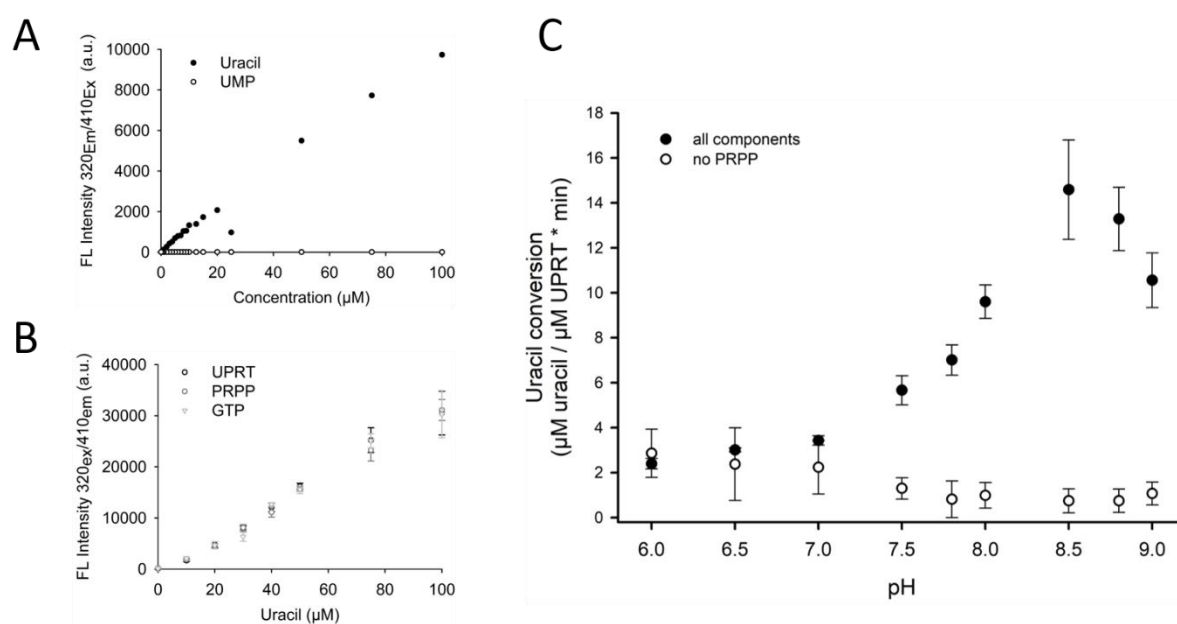
Figure 27: Activity assay for *E. coli* uracil phosphoribosyltransferase. (A) The uracil quantification reaction described by Shibata *et al.* (Shibata *et al.*, 2010) is specific for uracil over UMP. No blue fluorescence can be detected when UMP is used instead of uracil. (B) The assay components used in the UPRT reaction do not interfere with the uracil quantification reaction. Quantification of uracil was tested in the additional presence of 0.1 µM UPRT (black), 0.5 mM PRPP (dark grey), or 0.1 mM GTP (light grey). (C) Activity assay on *E. coli* UPRT at different pH. UPRT was incubated with 100 µM uracil, 0.5 mM PRPP, 2 mM Mg^{2+} and 0.1 mM GTP in buffer with different pH for up to 30 min and remaining uracil amount was quantified (filled circles). To determine background fluorescence, the assay was repeated excluding PRPP that is essential for the conversion of uracil to UMP (open circles). Linear fits to determine initial activities were performed using SigmaPlot 10. Experiment was done in triplicates, error bars shown are derived from standard deviations and error propagation.

The uracil quantification assay was used to determine initial reaction rates of uracil to UMP conversion by UPRT at different pH ranging from 6.0 to 9.0. For pH 6.0 to 7.5, potassium phosphate buffer was used, from pH 7.0 to 9.0 Tris buffer. For pH 7.0 and 7.5 both buffers were used with similar outcome. Overall, the activity of the purified protein was low, but in line with reported values (Jensen and Mygind, 1996; Lundegaard and Jensen, 1999). The activity of UPRT is strongly dependent on the concentration of the reactants as well as the allosteric activator GTP. Additionally, product inhibition and different oligomeric states with differing activities, depending on enzyme concentration and buffer conditions, have been observed (Jensen and Mygind, 1996). The activity of the enzyme that was obtained here appears not high enough for the enzyme to fulfil its role in uracil metabolism that is suggested above; however, the assay conditions here most likely do not well resemble the *in vivo* situation. The purified enzyme exhibited a pH optimum around pH 8.5, and, furthermore, activity dropped rapidly with decreasing pH, with almost no activity observed below pH 7.0. As a conclusion, it appears plausible that the internal acidification of the cytoplasm resulting from the combination of

an acidic assay buffer (pH 6.0) and a protonophore leads to immediate inactivation of uracil phosphoribosyltransferase. As a consequence, uracil uptake by UraA (or other uracil transporters) would decrease, mimicking the phenotype of proton-coupled uracil uptake.

4.3. Substrate recognition in SLC23 uracil transporters

An *in vitro* based transport assay could only be developed for 51ThXi, revealing that this protein works as a uniporter rather than as a proton-symporter. To determine whether 51ThXi is a representative member of the SLC23 family, its transport and binding behavior was investigated and compared to the one of UraA (section 4.3.1). The substrate binding site of SLC23 transporters contains a number of conserved charged or polar residues that were suggested to being involved in proton-coupling. In light of a uniporter mechanism, a new role for these residues needs to be defined. In order to investigate their relevance for substrate transport, a number of substrate binding site residues of UraA and 51ThXi were tested for binding as well as transport of uracil and of a number of uracil analogs (section 4.3.2). Most of the substrate binding site residues are conserved among SLC23 transporters which have been characterized regarding their function. As only few residues within the substrate binding site differ in SLC23 homologs that show distinct specificity for certain nucleobases over others, these residues were suggested to being the molecular determinants for SLC23 transporter specificity (Frillingos, 2012; Papakostas and Frillingos, 2012; Karena *et al.*, 2015; Chaliotis *et al.*, 2018) In this context, we also attempted to change the uracil transporters into transporters specific for xanthine, making use of the available structure of the SLC23 homolog UapA (section 4.3.3).

4.3.1. UraA, 51ThXi and 55PyP exhibit narrow substrate specificities

To explore the substrate specificities of UraA and the two homologs 51ThXi and 55PyPi, differential scanning fluorimetry (DSF) and competition assays in whole cells were performed, testing binding of uracil and a number of closely related uracil derivatives. For binding studies using DSF, proteins were purified similar as for reconstitution (section 4.2.4). Protein melting was probed using the cysteine-reactive dye CPM. CPM dye fluorescence increases upon binding to cysteines, which often get exposed when the protein unfolds. It was shown that binding of ligands can stabilize proteins in such an assay, which manifests in an increase of the melting temperature (Abbott *et al.*, 2017; Bai *et al.*, 2019). Exemplary, thermal melting of protein without addition of substrate and pre-incubated with excess uracil is shown in **Figure 28.** For all three proteins, a T_m shift, indicative of binding, was obtained in the presence of uracil. The melting temperature of UraA ($49.4 \pm 0.2°C$) increased by $9.4 \pm 0.3°C$ to $58.9 \pm 0.2°C$ when the protein was pre-incubated with 1 mM uracil. Purified 51ThXi appeared less stable than UraA, as suggested by a lower melting temperature in absence of stabilizing ligands ($43.4 \pm 0.2°C$). Uracil binding yet caused a strong shift of $11.6 \pm 0.3°C$. 55PyPi was stabilized by uracil by $8.3 \pm 0.7°C$ from $53.9 \pm 0.6°C$ to $62.3 \pm 0.3°C$. The non-symmetrical peak for the fluorescence derivative especially for the apo state of 55PyPi is probably caused by the presence of nine native cysteines in the protein. Using a mutant of 55PyPi, holding only one cysteine buried inside the protein,

was unsuccessful due to protein instability (not shown). Since uracil binding still led to a pronounced

shift of the T_m of 55PyPi, wild type protein was used for analyzing the substrate specificity.
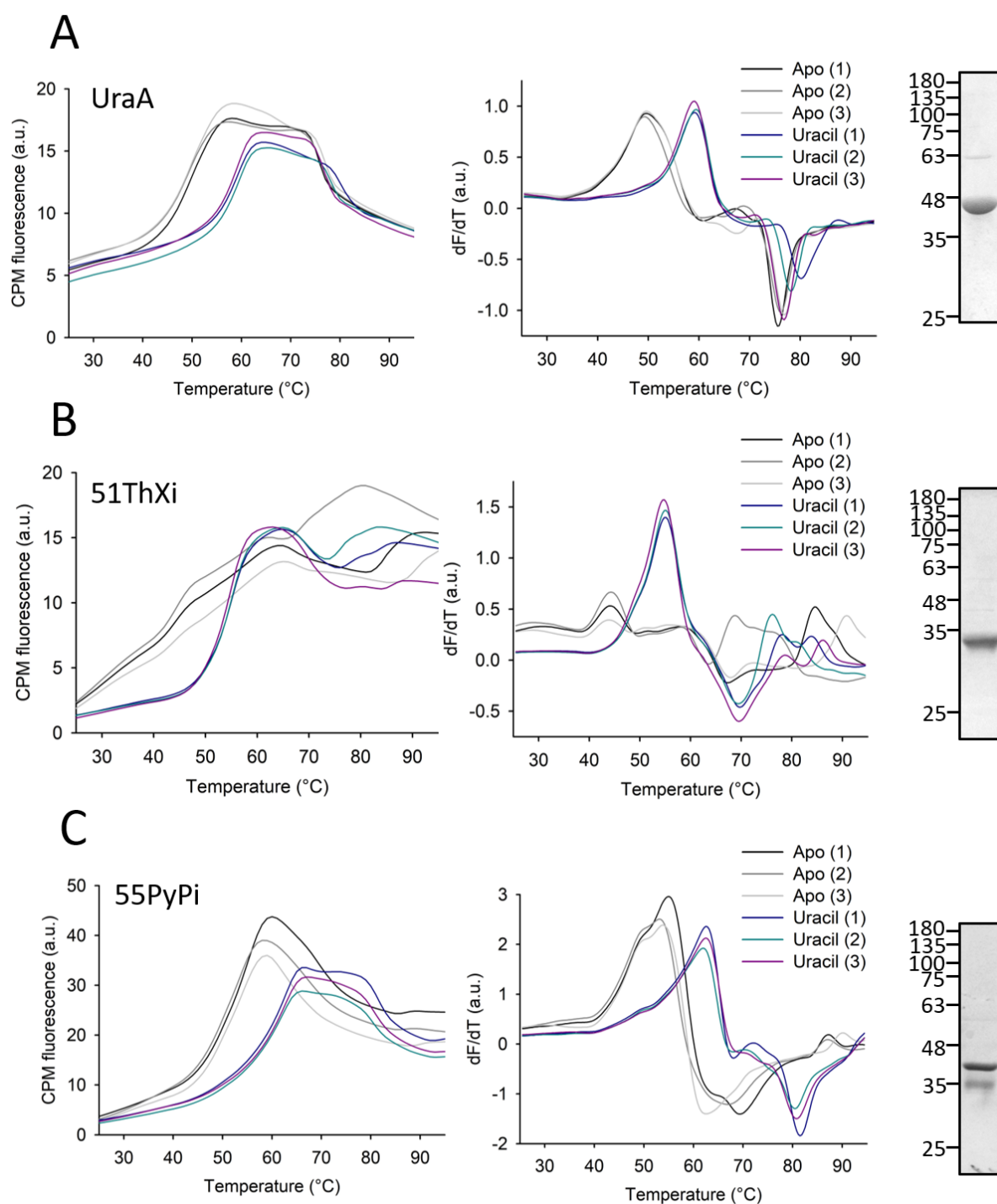
Figure 28: Differential scanning fluorimetry (DSF) to test binding of uracil analogs to purified UraA, 51ThXi and 55PyPi. (A) – (C) 18 µM (0.5 mg/mL) SEC-pure UraA (A), 51ThXi (B) or 55PyPi (C) were subjected to a thermal melting analysis in absence (black and grey lines) and presence of 1 mM uracil (colored lines). The temperature was increased from 25°C to 95°C in steps of 1°C with 20 sec incubation per step, and fluorescence of CPM dye was monitored (left). The melting temperature is found at the maximum of the fluorescence derivative (dF/dT, middle). Purity of protein fractions from SEC used for DSF was monitored by 12% SDS-PAGE and Coomassie staining (right). Melting experiments were performed in triplicates.
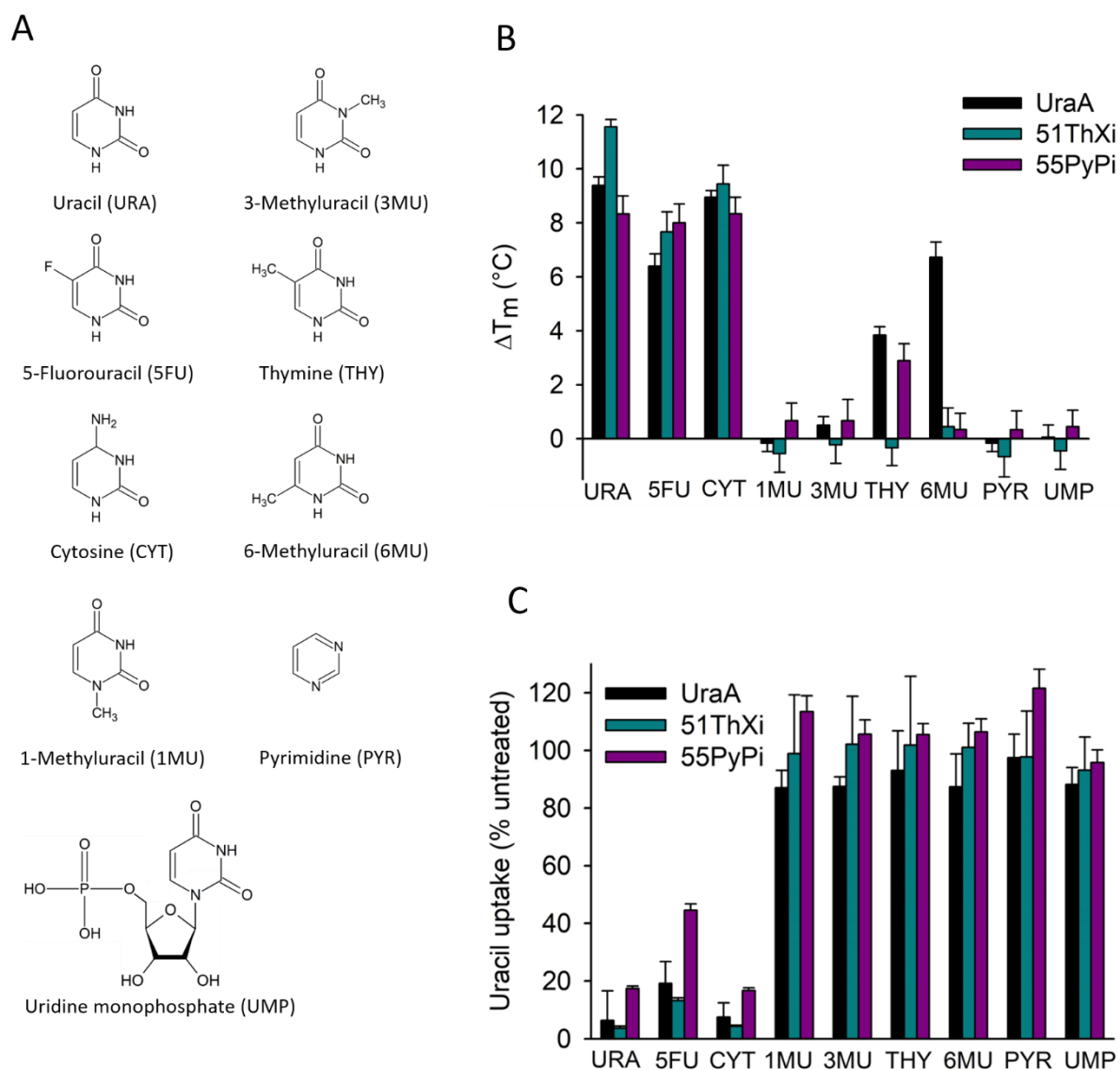
Figure 29: **Substrate specificities of UraA, 51ThXi and 55PyPi**. (A) Substrate analogs used to investigate substrate specificities of UraA, 51ThXi and 55PyPi. (B) Melting temperatures of UraA (black), 51ThXi (teal) and 55PyPi (dark pink) (18 µM/0.5 mg/mL each) were determined in presence of 1 mM of different uracil analogs. Depicted are differences in melting temperatures (ΔT_m) obtained by subtracting the T_m determined without substrate addition from the one derived in presence of the analogs. Melting experiments were performed in triplicates, error bars shown are from error propagation of standard deviations. (C) Competition assays in *Escherichia coli* BW25113(*ΔuraA*) expressing UraA (black), 51ThXi (teal) or 55PyPi (dark pink). Uptake of 0.1 µM uracil was monitored in presence of 0.1 mM of different uracil analogs. Shown is the remaining uracil uptake related to uptake by cells without any addition of competitor. Uptakes were performed in triplicates, errors shown are from error propagation on standard deviations.

Melting temperatures of purified UraA, 51ThXi and 55PyPi were determined in a similar fashion using different uracil analogs (**Figure 29A**). 5-fluorouracil (5FU) is a known anti-cancer agent and was already in 1995 suggested to be a UraA substrate (Andersen *et al.*, 1995) (section 4.2.7). Cytosine (CYT) closely resembles uracil, but is reported to be not a substrate for SLC23 transporters but for the structurally unrelated NCS1 (nucleobase-cation symporter 1) protein family. 1-methyluracil (1MU), 3-methyluracil (3MU), thymine (5-Methyluracil, THY) and 6-methyluracil (6MU) contain an uracil core with additional

methyl groups. Pyrimidine (PYR) harbors the basic ring structure, but does not hold the ketone groups at position 2 and 4 (O2 and O4) compared to uracil. Uridine monophosphate (UMP) is the product of the main downstream reaction of uracil transport in *E. coli*, catalyzed by uracil phosphoribosyltransferase.

For all three proteins, stabilization by uracil, 5-fluorouracil and cytosine was observed (**Figure 29B**). Interestingly, UraA was additionally stabilized by thymine and 6-methyluracil, 55PyPi by thymine, but not by 6-methyluracil, and 51ThXi by neither of those. Contrary, in cell-based assays, competition was seen only for uracil, 5-fluorouracil, and cytosine, but not for thymine or 6-methyluracil for any of the proteins (**Figure 29C**). No competition for pyrimidine was observed, suggesting a relevance of O2 and O4 of uracil for interaction with the substrate binding site. UMP is not a substrate for 51ThXi or 55PyPi, as was already shown for UraA (section 4.2.8). The newly set up *in vitro* uptake assay for 51ThXi also allowed to investigate its substrate specificity in proteoliposomes (**Figure 30**). The results from proteoliposome-based competition assays are in good agreement with the ones performed in whole cells, showing a reduction of uracil uptake in presence of excess 5FU or cytosine (**Figure 29C**, **Figure 30A**).
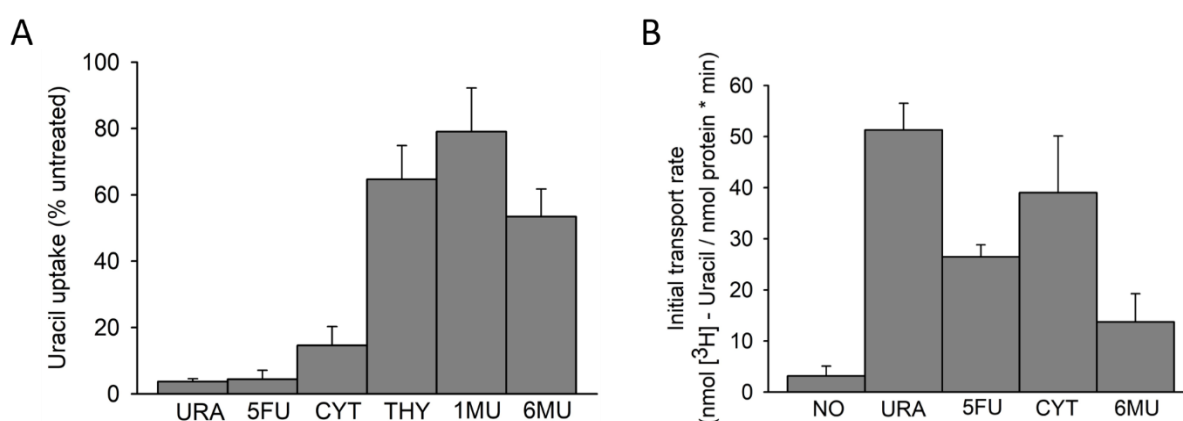
Figure 30 Substrate specificity of reconstituted 51ThXi. (A) Proteoliposome-based competition assay. Uptake of 1 µM uracil was monitored in presence of 1 mM competitor and compared to uracil uptake obtained in the absence of competitor. (B) Transport of 5-fluororuacil and cytosine by 51ThXi. Proteoliposomes were pre-loaded with 10 µM partially labeled uracil and diluted in buffer containing no countersubstrate (NO) or 10 µM of uracil analogs. Initial transpor rates were determined by linear fitting using SigmaPlot 10. All experiments were performed in triplicates and errors are derived from standard deviations and error propagation. URA, uracil; CYT, 5FU, 5-fluorouracil; CYT, cytosine; THY, thymine; 1MU, 1-methyluracil; 6MU, 6-methyluracil.

To test which of the uracil analogs are not only binding, but actually transported by 51ThXi, exchange assays were performed using some of the competitors (**Figure 30B**). In all cases, proteoliposomes were loaded with 10 µM radiolabeled uracil and diluted in outside buffer containing either no competitor or 10 µM each of uracil, 5-fluorouracil, cytosine, or 6-methyluracil. Exchange was observed not only for

uracil, but also for 5-fluorouracil and cytosine, meaning all these compounds that were shown to bind to 51ThXi are also transported.

A

```
55PyPi    ISLDNVNAAKWFQFTGFMPWGMPKFSTVAFGAILAGFFSVILESIGDYFNVCNAAGLPDP    281
UraA      VDTTPIINAHWFALPTLY---TPRFEWFAILTILPAALVVIAEHVGHLVVTANIVKKDLL    258
51ThXi    IDFTELSQAAWFGMPSFV---TPRFESSAMFMIAPVAFILVAENLGHIKAIGGMTKQNMD    273
          :.   :  * ** :  :     *:*.  *:  *     : :: * :*.    . .

55PyPi    TEQQISRGIRAEGLGCIFGGLTGAVACTSYTENIGLIGLTGVASRWVVRVGAILLIGMSM    341
UraA      RDPGLHRSMFANGLSTVISGFFGSTPNTTYGENIGVMAITRVYSTWVIGGAAIFAILLSC    318
51ThXi    R--YIGRGFVGDGLATMLAGSGGGTGVTTYVENMGVMAVTRVFSTLIFVIAAVIAILLGF    331
          :  *.: .:**. ::.* *.. *:* **:*::.:* * *  :. .*:: * :.
```

B
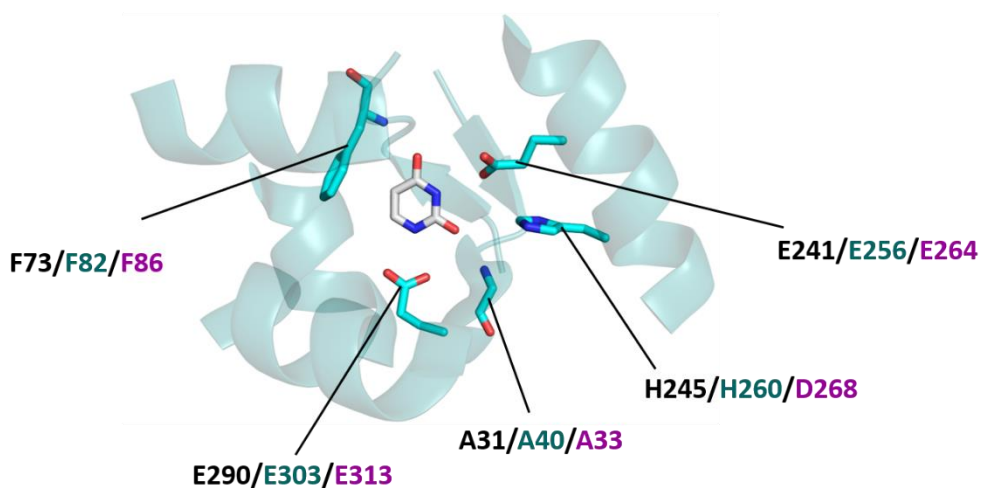
Figure 31: Substrate binding sites of 51ThXi and 55PyPi. (A) Sequence alignment of UraA, 51ThXi and 55PyPi. Full-length sequences were aligned using Clustal Ω (Sievers and Higgins, 2018). E241, H245 an E290 of UraA that are in hydrogen-bonding distance with uracil and corresponding residues in 51ThXi (E256, H260 and E303) and 55PyPi (E264, D268 and E313) are indicated. The overall sequence identity to UraA is 33.41% for 51ThXi and 22.90% for 55PyPi according to AlignMe (Stamm *et al.*, 2014). (B) Substrate binding site of UraA with residues suggested to form electrostatic interactions with uracil, either over the side chains (E241, H245, E290) or the backbone (A31, F73). Corresponding residues in 51ThXi are annotated in dark teal, for 55PyPi in dark pink.

In general, with the exception of H245/H260 in UraA/51ThXi that is D268 in 55PyPi, the residues found in the uracil binding site are identical for the three proteins (**Figure 31**). Based on the UraA structure, it can be hypothesized that binding of 1MU and 3MU is precluded due to interactions of the substrate with E290 and E241, respectively. An additional methyl group at these positions is not allowed either because it disturbs these interactions, or because of sterical clashes with the protein. The (almost) identical architecture of the binding side and the similar binding behavior of 51ThXi and 55PyPi suggest that they bind substrate in the same orientation as UraA. It also indicates that the interaction of His245/H260 with O2 of uracil is not crucial for substrate binding, as aspartate in 55PyPi is also tolerated at this position. In line with this, we find that the UraA(H245D) and 51ThXi(H260D) mutants

are fully capable of uracil transport (section 4.3.2.4). The difference in specificity in binding observed for UraA and 55PyPi in DSF compared to competition in whole cell experiments (stabilization in DSF by thymine and 6-methyluracil for UraA, by thymine for 55PyPi, but no competition for neither compound (**Figure 29B,C**) is less straightforward to explain. The competition assay in whole cells probes the outward-facing binding site, while in the thermal shift experiment on detergent-solubilized protein both binding sites are, in principle, accessible. The differences in binding pattern observed using the two different assays indicate that the substrate binding sites might differ in the distinct conformations of the protein. Whether the purified proteins favor a certain conformation in the detergent-solubilized state, however, is currently not known. Another, more straightforward explanation would be the existence of a second binding site, that only becomes apparent in certain conformations. No in-depth experiments were conducted to explore this possibility; in DSF experiments, no distinct additional stabilization was observed when UraA was incubated with two substrates that both stabilized the protein on their own (**Figure S5**).

4.3.2. *Conservation of the UraA substrate binding site*

Uracil is coordinated by a number of polar and non-polar interactions within the substrate binding site of UraA. In particular, E241 and H245 on TM8 and E290 on TM10 are in suitable distances to form hydrogen bonds with the substrate. Additionally, Y288 probably forms π-π-stacking interactions with the pyrimidine ring. In consequence, mutating either E241, H245 or E290 to alanine led to severe impairment in substrate transport. SPA assays indicated that all mutants lost the ability to bind the substrate (Lu *et al.*, 2011; Yu *et al.*, 2017). As SLC23 transporter were so far thought of as proton-dependent symporters, the charged substrate binding site residues were consequently suggested to being involved in proton translocation. Different studies, however, come to different conclusions on which of the residues is mainly responsible for coupling (Yu *et al.*, 2017; Yang *et al.*, 2018).

To analyze the conservation of substrate binding site residues of UraA, the ConSurf server was used, including 500 automatically picked UraA homologs (**Figure 32**). Almost all of the residues in the vicinity of uracil were annotated with the highest conservation score possible. This includes the three UraA residues E241, H245 and E290 located on TM8 and 10, respectively, that are involved in polar or electrostatic interactions with the substrate. The two glutamates appear to be, as is illustrated in the WebLogo depiction on the right of **Figure 32**, fully conserved, while for the residue at position 245 of UraA histidine or aspartate is found in roughly equal proportions. This is reflected, according to sequence alignments, in the two newly identified uracil transporters 51ThXi and 55PyPi: both contain glutamates on TM8/10 corresponding to E241/E290 in UraA (E256/E303 in 51ThXi and E264/E303 in 55PyPi), but 55PyPi harbors D268 at the position of H245/H260 in UraA and 51ThXi, respectively

(**Figure 31B**). To further substantiate the importance of the conserved substrate binding site residues for the function of SLC23 uracil transporters, different mutants of UraA, 51ThXi and 55PyPi were generated and analyzed for binding as well as for transport, as will be described in the following.
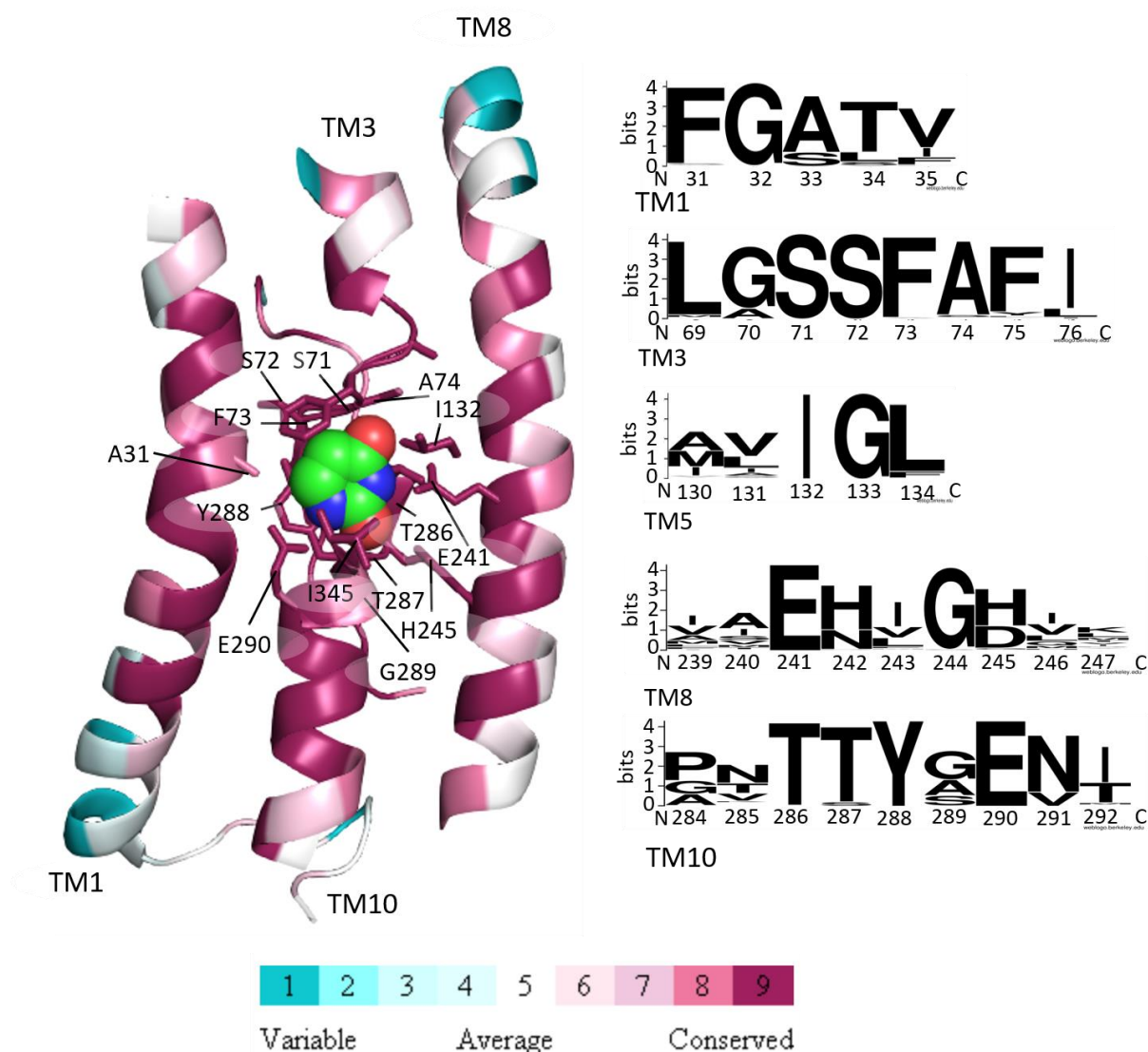
Figure 32: Conservation of the UraA substrate binding site. The conservation of the UraA substrate binding site was analyzed using the ConSurf server (https://consurf.tau.ac.il/), including 500 automatically picked UraA homologs with sequence identities from 36.34% - 90.95% from the UNIREF90 database. Colors are based on the default ConSurf scheme as indicated on the bottom, with dark pink for extremely conserved residues and dark cyan for least conserved residues. Only helixes 1, 3, 8 and 10 are shown as they (almost) exclusively contain the uracil binding residues. Uracil is shown in spheres, residues within 5 Å of uracil are shown in sticks and annotated. Sequence conservation is also depicted on the right using WebLogo (https://weblogo.berkeley.edu/logo.cgi).

4.3.2.1. Removing charged substrate binding site residues affects transport

As outlined in section 4.3.2, the substrate binding site of SLC23 uracil transporters is highly conserved, but the role of the conserved residues for substrate translocation is not well understood. To address this issue, several binding site residues in UraA, 51ThXi and 55PyPi were mutated and tested for binding in DSF as well as transport in a cellular context. These mutants in particular include the residues suggested to being involved in electrostatic interactions with the substrate (E241, H245 and E290 in UraA) that were not only replaced by alanine, but also other, more conservative amino acids (section 4.3.2.3). Lu *et al.* show that mutating either of these to alanine abrogates uracil transport. In a scintillation proximity assay (SPA), none of the mutants was able to bind the substrate (Lu *et al.*, 2011). As already described in section 4.2.1, in the uptake experiments performed for this study, a small but noticeable residual activity of UraA(E241A) was observed, in contrast to the complete abolishment of uracil transport reported by Lu *et al.*. Assuming that UraA is a proton-coupled symporter, they suggest that one or more of the charged substrate binding site residues are involved in coupling. The only SLC23 transporter for which, to date, an *in vitro* assay could be established, 51ThXi, was demonstrated in section 4.2.5 to function as a uniporter rather than as a symporter. In order to get more insight into the relevance of the conserved substrate binding site residues for substrate transport in the SLC23 family, the respective mutants were generated for UraA (E241A, H245A and E290A)[7], 51ThXi (E256A, H260A and E303A) and 55PyPi (E264A, D268A and E313A) and tested for uracil uptake in whole cells (**Figure 33**, **Figure S6**). Almost all of the mutants showed elevated uracil uptake compared to the cells expressing LacS as negative control. Interestingly, uracil uptake for cells expressing the 55PyPi(D268A) mutant was close to that of the wild type protein while both UraA(H245A) and 51ThXi(H260A) were affected in transport. This suggests a different function of the residue at that position in 55PyPi than in the other uracil transporters, but in the light of the high degree of conservation of the substrate binding site this differing role remains elusive at the moment.

---

[7] UraA(E241A) was generated and kindly provided by Benedikt Kuhn.
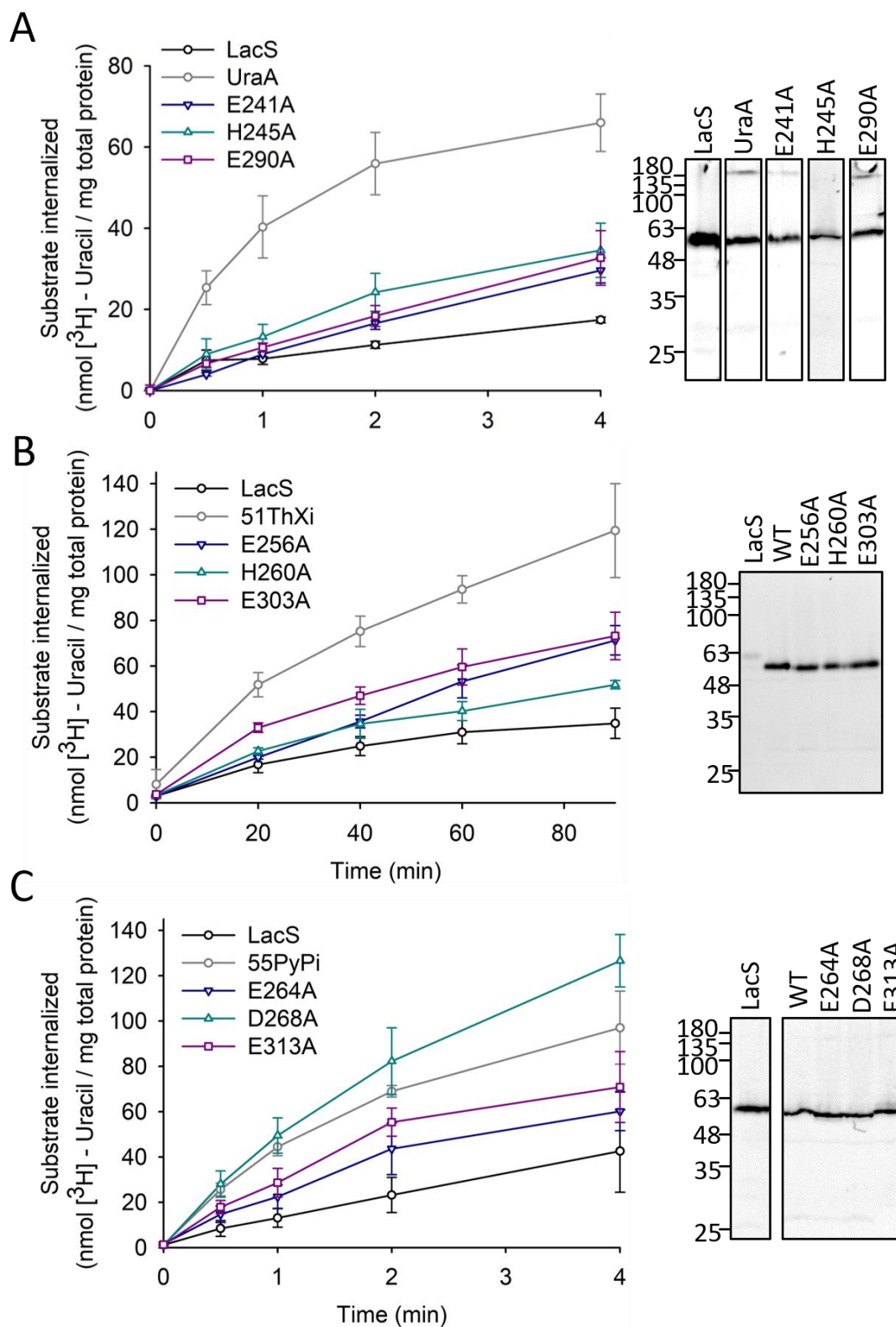
Figure 33: Uracil uptake of substrate binding site mutants of UraA, 51ThXi and 55PyPi. (A) Uracil uptake of *E. coli* BW25113(*ΔuraA*) overexpressing GFP-fusions of LacS as negative control (black), wild type UraA (grey) or alanine mutants of substrate binding site residues of UraA (E241A, dark blue, H245A, teal, or E290A, dark pink). (B) Same as in (A), but testing respective substrate binding site mutants of 51ThXi. (C) Same as in (A) and (B), but testing respective substrate binding site mutants of 55PyPi. All uptakes were determined in triplicates, errors shown represent standard deviations. Expression levels of proteins were visualized by *in gel* GFP fluorescence on whole cell lysates.

4.3.2.2. The substrate binding site residues are relevant for protein stability

The reduction of uracil uptake observed for the substrate binding site mutants of UraA, 51ThXi and 55PyPi might be explained by an impaired binding of uracil. In order to test this hypothesis, the respective mutants of UraA and 51ThXi were purified with the intention to test substrate binding in DSF experiments. Interestingly, a T_m determination for UraA(E241A) proved difficult although after SEC a monodisperse peak was obtained, indicating good protein quality. The characterization of the UraA(E241A) mutant was the primary subject of the master thesis of Leonard Präve that he prepared under my supervision ("Functional Cooperativity in a UraA Dimer", Goethe University Frankfurt am Main, Oct 2019). In summary, it was not possible to determine a melting temperature for the apo state. Pre-incubation of the protein with 1 mM uracil yet seemed to be able to stabilize the protein in such a way that a melting temperature could be obtained. For this, the protein had to be immediately used after SEC, while wild type UraA can be stored at 4°C for several days without a change in DSF melting profile. This indicated not only an instability of the mutant compared to wild type UraA, but also a retained ability to bind uracil. A similar destabilizing effect, preventing an accurate T_m determination, was observed for UraA(H245A) (**Figure S7**). For the UraA(E290A) mutant, however, a melting temperature of 46.9 ± 0.2°C could be determined, which is 2.5 ± 0.3°C lower than the one of wild type UraA (49.4 ± 0.2°C, section 4.3.1). Nevertheless, also UraA(E290A) was compromised in protein quality compared to wild type UraA, indicated by aggregation in SEC and impurities in the final protein sample (**Figure S8A**) Notably, compared to the wild type protein, the E290A mutant exhibited a different binding profile (**Figure 34**). Both cytosine and 6-methyluracil, which have a pronounced stabilizing effect on wild type UraA, did not stabilize the mutant. Additionally, no melting temperature at all could be determined when UraA(E290A) was pre-incubated with 5-fluorouracil which indicates a destabilizing effect of this compound (**Figure S8B-D**). In summary, the differing substrate specificity suggests an alteration in the overall binding site which might for example be caused by different affinities towards the different substrate analogs compared to wild type UraA, although this fails to explain the destabilization observed in the presence of 5FU. The instability of the alanine substrate binding site mutants might explain why Lu *et al.* observed no binding of uracil in SPA measurements, while herein, using thermal shift assays and in general higher substrate concentrations, substrate binding seemed to be not entirely suppressed. The instability of the mutants also prevented further analysis of the mutants' affinity towards the substrates. Affinity measurements on the basis of uptake assays in whole cells proved inconclusive due to pronounced background uracil uptake activity described in section 4.2. Therefore, it remains unclear whether the reduced transport activity of the E241A, H245A or E290A mutants of UraA is caused by an effect on the transport rate, or rather by a reduction in affinity for the substrate, or by a combination of both.
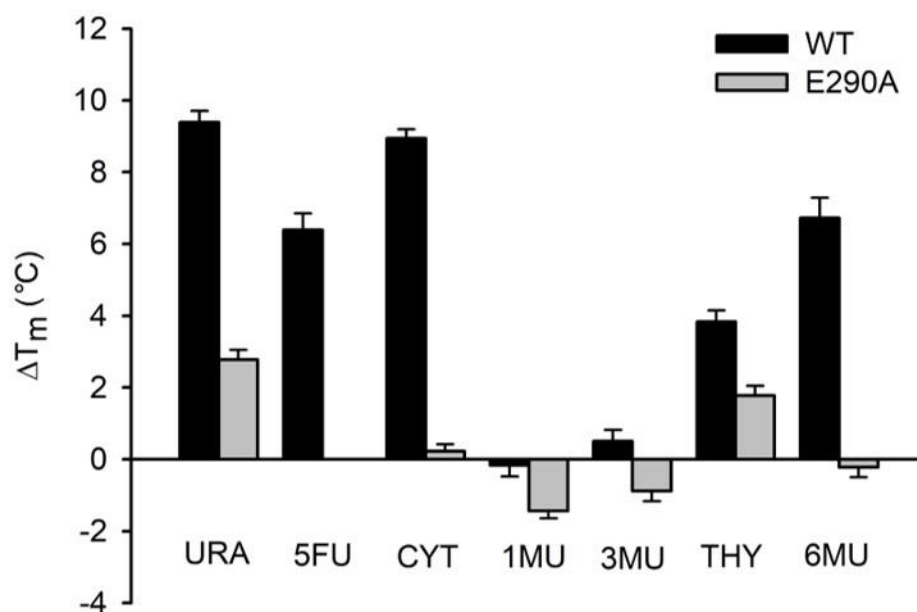
Figure 34: Comparison of substrate binding profile of wild type UraA and UraA(E290A). 18 μM (0.5 mg/mL) SEC-pure protein (either wild type UraA (black) or UraA(E290A) (grey)) were pre-incubated with 1 mM of the indicated substrates and subjected to DSF. ΔT_m was obtained by subtracting the T_m of protein in absence of substrate. No melting temperature of UraA(E290A) in presence of 5FU could be determined due to a destabilizing effect (**Figure S8D**). Melting temperatures were determined in triplicates, errors shown are derived from standard deviations and error propagation. URA, uracil; 5FU, 5-fluorouracil; CYT, cytosine; 1MU, 1-methyluracil; 3MU, 3-methyluracil; THY, thymine; 6MU, 6-methyluracil.

A similar destabilizing effect was as well observed for the respective mutants of 51ThXi. Both 51ThXi(E256A) and 51ThXi(H260A) were purified to homogeneity, but no melting temperature could be obtained in DSF in the absence of a stabilizing compound (**Figure 35**). Nevertheless, in presence of several uracil analogs reliable melting temperatures could be determined, indicating binding of this substrates to the mutants; DSF therefore still allows for a quantitative analysis of their binding specificities. As no T_m for the protein in absence of substrate is available, absolute melting temperatures are given in **Figure 36.** All substrates for which no T_m is given were tested nevertheless but found to not stabilize the protein in such a way that a melting temperature could be obtained, suggesting that they are binding with a low affinity or not at all. For 51ThXi(E303A), low fluorescence was obtained in presence of any of the substrates tested, suggesting that none of them - including uracil - were able to stabilize the protein significantly (**Figure S9**).
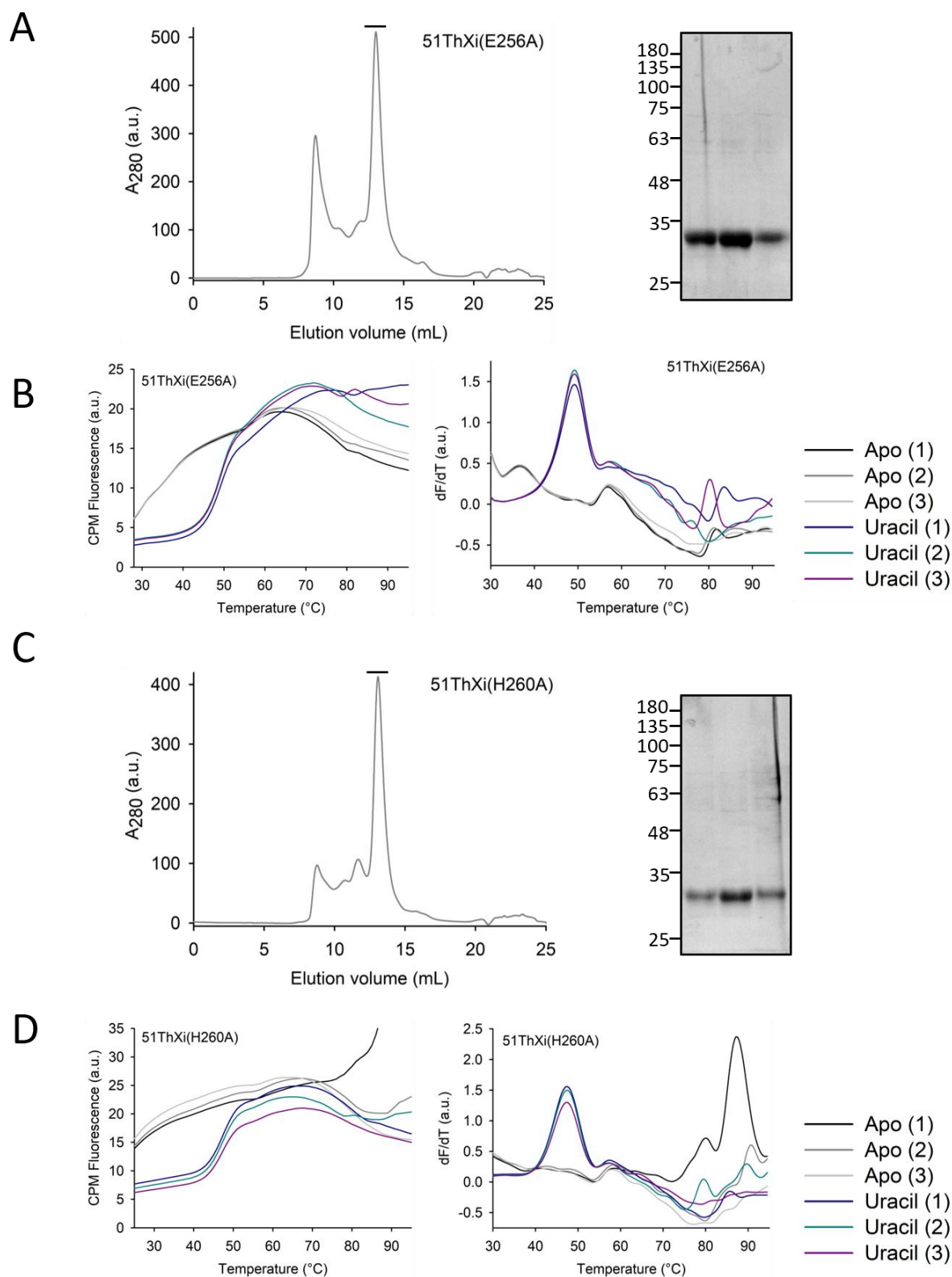
Figure 35: Purification and thermal melting of 51ThXi(E256A) and 51ThXi(H260A). (A) Size-exclusion chromatogram of 51ThXi(E256A) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) Thermal melting of 51ThXi(E256A) in absence of substrate (black and grey lines) and in presence of 1 mM uracil (colored lines). 18 µM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 seconds incubation per step and CPM dye fluorescence was recorded. Shown are raw CPM fluorescence (left) and dF/dT (right). (C) and (D) Same as in (A) and (B), but for 51ThXi(H260A).
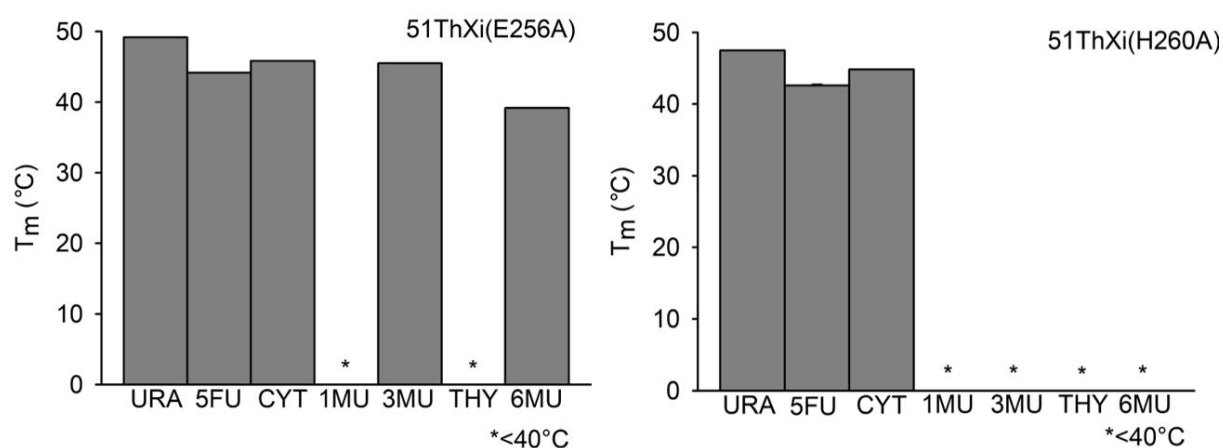
Figure 36: Binding profiles of 51ThXi(E256A) and 51ThXi(H260A). Binding of different uracil analogs to 51ThXi(E256A) and 51ThXi(H260A) was tested using DSF. As it was not possible to determine a T_m in the absence of substrate, absolute melting temperatures are given here. For several uracil analogs, indicated here with an asterisk (*), melting resembled the one obtained in absence of substrate, indicating that these compounds were not able to stabilize the mutants. URA, uracil; 5FU, 5-fluorouracil; CYT, cytosine; 1MU, 1-methyluracil; 3MU, 3-methyluracil; THY, thymine; 6MU, 6-methyluracil.

Interestingly, while the H260A mutant showed the same specificity as wild type 51ThXi (stabilization by uracil, 5-fluorouracil, and cytosine), 51ThXi(E256A) exhibited an extended spectrum of substrate binding: next to uracil, 5-fluorouracil and cytosine, 51ThXi(E256A) was also stabilized by 6-methyluracil, and, notably, 3-methyluracil (**Figure 36**). For wild type 51ThXi (as well as for UraA and 55PyPi), E256 was suggested to form electrostatic interactions with N3 of uracil (section 4.3.1). Replacing E256 by alanine creates space in the substrate binding site into which the additional methyl group of 3MU can fit. This once again demonstrates that E256 is not essential for substrate binding.

To further substantiate the destabilizing effect of mutating E256, H260 or E303 of 51ThXi to alanine, the melting temperature in absence of substrate, that was not accessible by DSF, was determined by an alternative approach: the IMAC-pure protein was incubated for 5 min at different temperatures and the protein content in the supernatant was analyzed using SEC (**Figure 37**). The melting temperatures obtained for wild type 51ThXi in absence and presence of uracil are in good agreement with the ones determined by DSF. As anticipated, both E256A and H260A mutants were destabilized in the apo state compared to the wild type protein by ~7°C and ~4°C, respectively (**Figure 37B**). For both, incubation with uracil caused an upshift of the melting temperatures to ~45°C (**Figure 37C**). This suggests that both mutants are able to bind uracil. The 51ThXi(E303A) mutant, on the other hand, showed a melting temperature close to the one of wild type 51ThXi, consistent with the data obtained for UraA(E290A) (**Figure S8B,C**). Interestingly, the melting temperature of 51ThXi(E303A) was drastically reduced in the presence of uracil (-13.0 ± 0.6°C) (**Figure 37C**). In summary, the SEC data confirms the crucial role of E256 and H260 in 51ThXi for protein stability while not being absolutely essential for uracil binding.
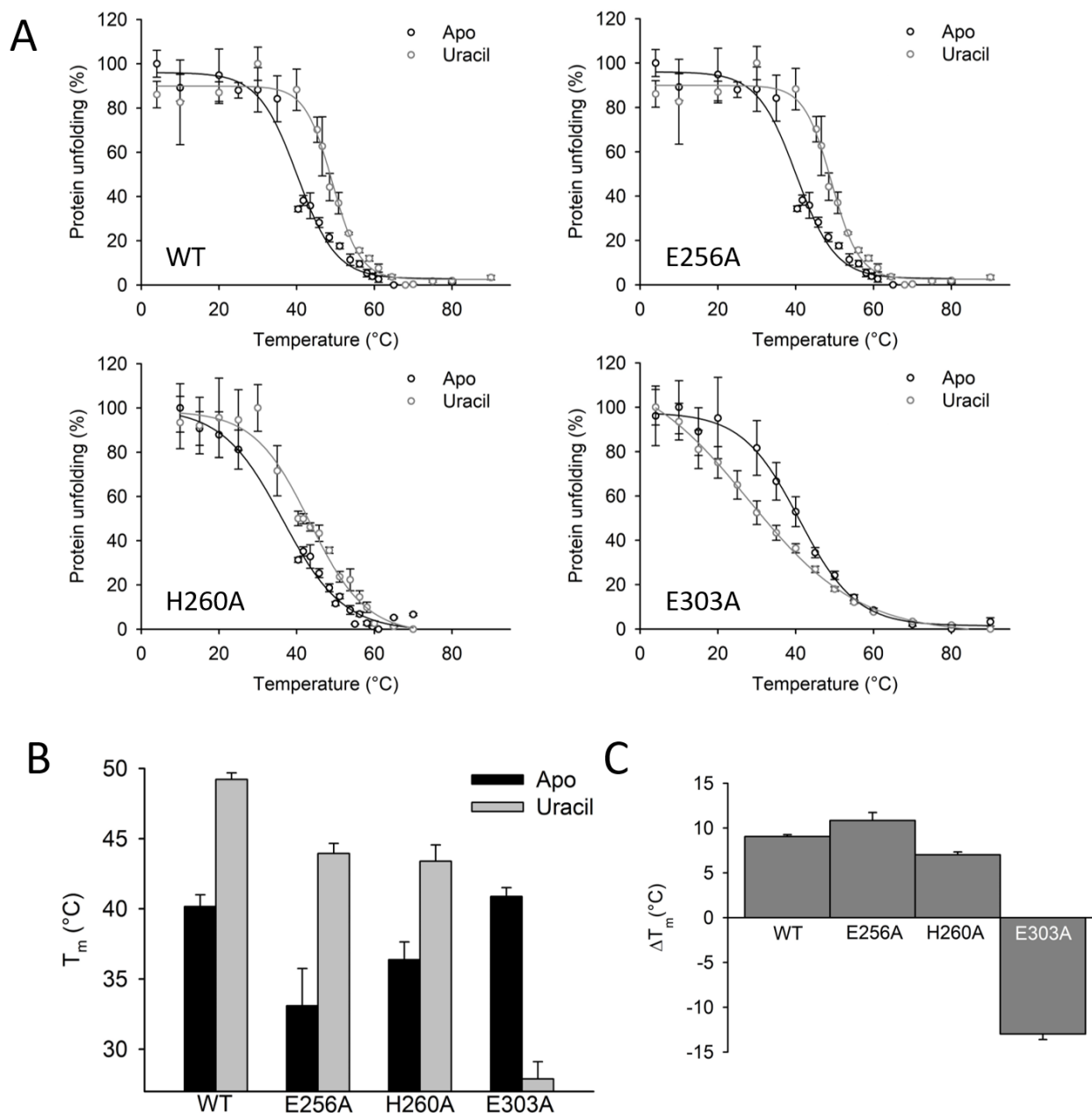
Figure 37: Determination of melting temperatures of 51ThXi substrate binding site mutants using SEC. (A) 18 µM (0.5 mg/mL) IMAC-pure wild protein was incubated at different temperatures for 5 min and protein content in the supernatant was analyzed using SEC. Melting temperatures were determined without additional substrate (black) or in presence of 1 mM uracil (grey). Peak areas were determined on triplicates using SigmaPlot 10, averaged and normalized. Sigmoidal peak fitting was performed using SigmaPlot 10. (B) Melting temperatures obtained in (A) in absence (black) and presence of uracil (grey). (C) The difference in T_m in presence and absence of uracil (ΔT_m) demonstrates the stabilizing effect on WT 51ThXi, 51ThXi(E256A) and 51ThXi(H260A) and the destabilization of 51ThXi(E303A).
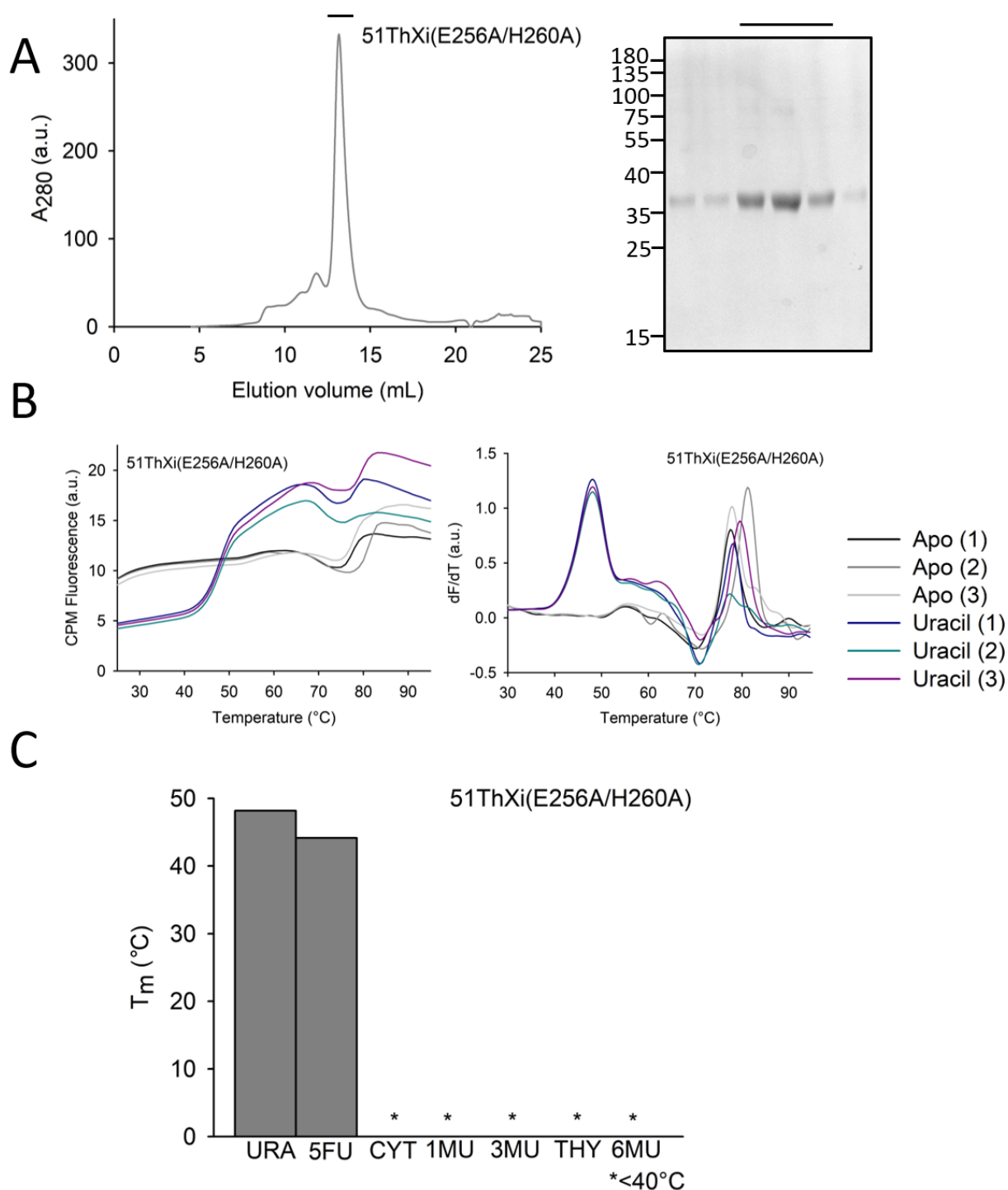
Figure 38: Uracil binding to 51ThXi(E256A/H260A). (A) Size-exclusion chromatogram of 51ThXi(E256A/H260A) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) DSF measurement on 51ThXi(E256A/H260A). 18 µM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. For determination of T_m, the maximum of the fluorescence derivative was used (dF/dT). (C) Binding of uracil analogs to 51ThXi(E256A/H260A) determined by DSF. Substrate analogs were used at a final concentration of 1 mM. For several uracil analogs, indicated here with an asterisk (*), melting resembled the one obtained in absence of substrate, indicating that these compounds were not able to stabilize the mutant. Melting temperatures were measured in triplicates, errors are standard deviations. URA: uracil, 5FU: 5-fluorouracil, CYT: cytosine, 1MU: 1-methyluracil, 3MU: 3-methyluracil, THY: thymine, 6MU: 6-methyluracil.

As both E256A and H260A as separate mutants were found to not be impaired *per se* in substrate binding, next a 51ThXi(E256A/H260A) double mutant was generated and binding and transport were measured (**Figure 38**). For this double mutant, a monodisperse peak of pure protein could be obtained (**Figure 38A**). In DSF measurement, the melting curve of the apo state indicated a similar instability compared to wild type protein as observed for the E256A and H260A single mutants, and uracil still seemed to be able to bind to the protein (**Figure 38B**). The melting temperature obtained in the presence of 1 mM uracil for the double mutant was similar as the one for the single mutants (48.2 ± 0.0°C for E256A/H260A, 49.2 ± 0.0°C for E256A, 47.5 ± 0.0°C for H260A). In uracil uptake experiments, an impairment in transport similar to the one of the E256A and H260A single mutants was observed (**Figure S10**).

While stabilization in presence of uracil and 5-fluorouracil was seen, 51ThXi(E256A/H260A) was no longer able to bind cytosine (**Figure 38C**), while this compound was found to stabilize both E256A and H260A single mutants (**Figure 36**). Cytosine and uracil differ in an amino group at C4 of the pyrimidine ring that is a ketone group in uracil. In UraA, the C4-ketone group of uracil was suggested to interact with the backbone amide nitrogen of F73 (F82 in 51ThXi) (**Figure 39**) (Lu *et al.*, 2011). With an amino group at that position as found in cytosine this favorable interaction is disrupted. However, currently all considerations concerning the location of uracil side groups in the substrate binding site are based on the assumption that the orientation of the ligand is the same in all of the mutants. Both E256A and H260 single mutants miss another interaction with the ligand compared to wild type protein (E256 with N3, H260 with the ketone group at C2) which both presumably contribute to high-affinity binding. Although in the E256A/H260 double mutant both of these interactions are missing, uracil binding can still be detected, while cytosine binding probably is prevented due to loss of the additionally favorable interaction of the F82 amide nitrogen with this ligand.
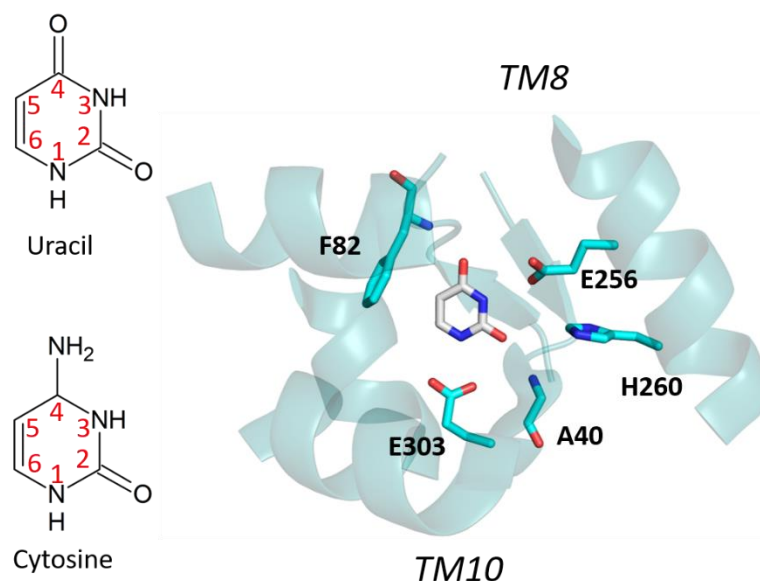
Figure 39: **Structural difference of uracil and cytosine**. Uracil (IUPAC name: pyrimidine-2,4(1*H*,3*H*)-dione) and cytosine (IUPAC name: 4-Aminopyrimidin-2(1*H*)-one) differ in one side functional group located at C4 of the pyrimidine ring that is a ketone group in uracil, but an amino group in cytosine. As the substrate binding site residues located close to the substrate in the UraA structures (3QE7 and 5XLS) are identical in 51ThXi, uracil is expected to bind in a similar orientation in both proteins as indicated on the right. This places the functional group on C4 close to the amide nitrogen of F82.

Conspicuously, E256 and H260 are located right next to each other on TM8 and close enough to interact. If one of the residues is removed, this might create an unbalanced charge in the substrate binding site that might explain the instability of the E256A and H260A mutants compared to wild type 51ThXi. Consequently, removing both charges as in the E256A/H260A double mutant might be able to restore the stability of the protein. However, this effect was not observed here. A 51ThXi(E256H/H260E) mutant, in which both residues are swapped, showed poor protein quality and was not only affected in transport, but also no uracil binding was observed (**Figure S11**). This indicates that next to E256 and H260, more residues, possibly including E303, are involved in an interaction network that stabilizes the substrate binding site or the substrate binding domain as a whole.

4.3.2.3. Substrate binding site glutamates in 51ThXi are irreplaceable for transport

The alanine mutant of the conserved glutamate on TM8 in 51ThXi (51ThXi(E256A)) was largely inactive in whole cell-based uracil uptakes. In contrast, uracil stabilized the purified protein in DSF experiments, indicating that uracil was still able to bind, although a lower affinity cannot be excluded. In contrast, no binding or transport of uracil was obtained for the 51ThXi(E303A) mutant. In order to further elucidate the role of the conserved glutamates in the substrate binding side of 51ThXi, additional, more conservative mutants were generated. In particular, E256 was mutated to S, D and Q, E303 to D and

Q. In uracil uptake assays, none of the mutant showed wild type-like transport activity, indicating that both E256 and E303 of 51ThXi are absolutely irreplaceable for uracil transport (**Figure 40**).
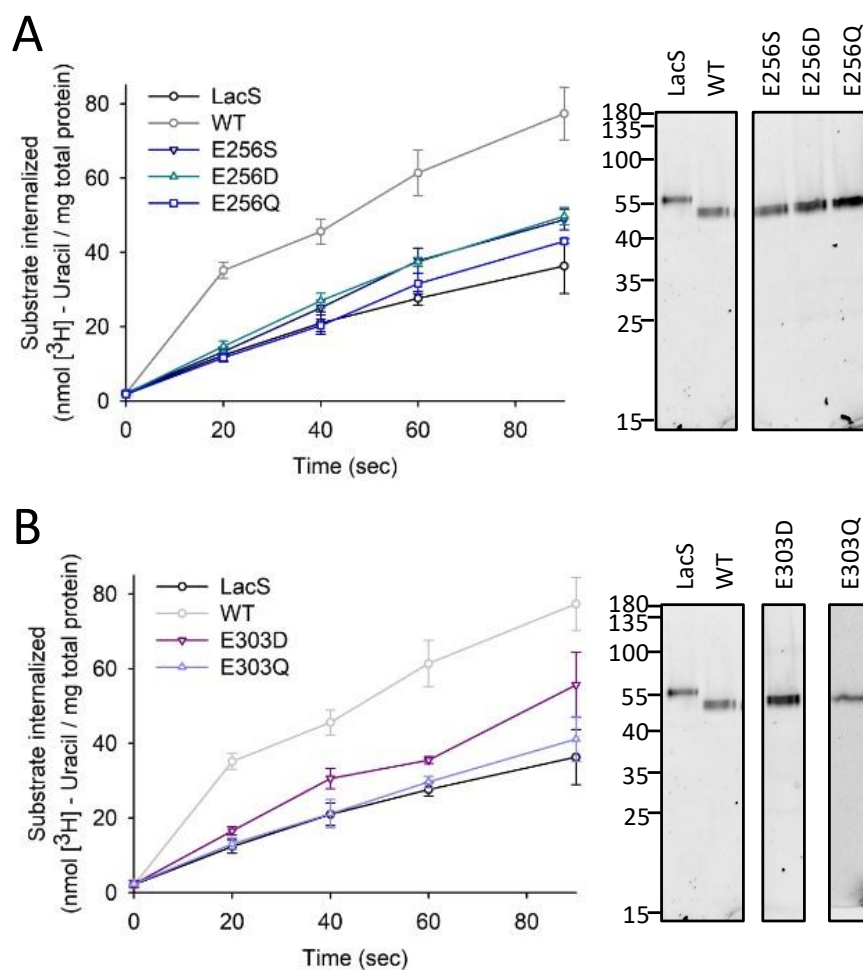
Figure 40: Uracil uptake of mutants of E256 and E303 of 51ThXi. (A) Uracil uptake in *E. coli* BW25113(*ΔuraA*) overexpressing GFP fusions of LacS as negative control (black), wild type 51ThXi (grey), or 51ThXi mutants E256S (dark blue), E256D (teal) or E256Q (blue). (B) Uracil uptake in *E. coli* BW25113(*ΔuraA*) overexpressing GFP fusions of LacS as negative control (black), wild type 51ThXi (grey), or 51ThXi mutants E303D (dark pink) or E303Q (powder blue). All uptakes were measured in triplicates, errors shown are from standard deviations. Expression levels of uptake samples were monitored by 12% SDS-PAGE and *in gel* GFP fluorescence on whole cell lysates.
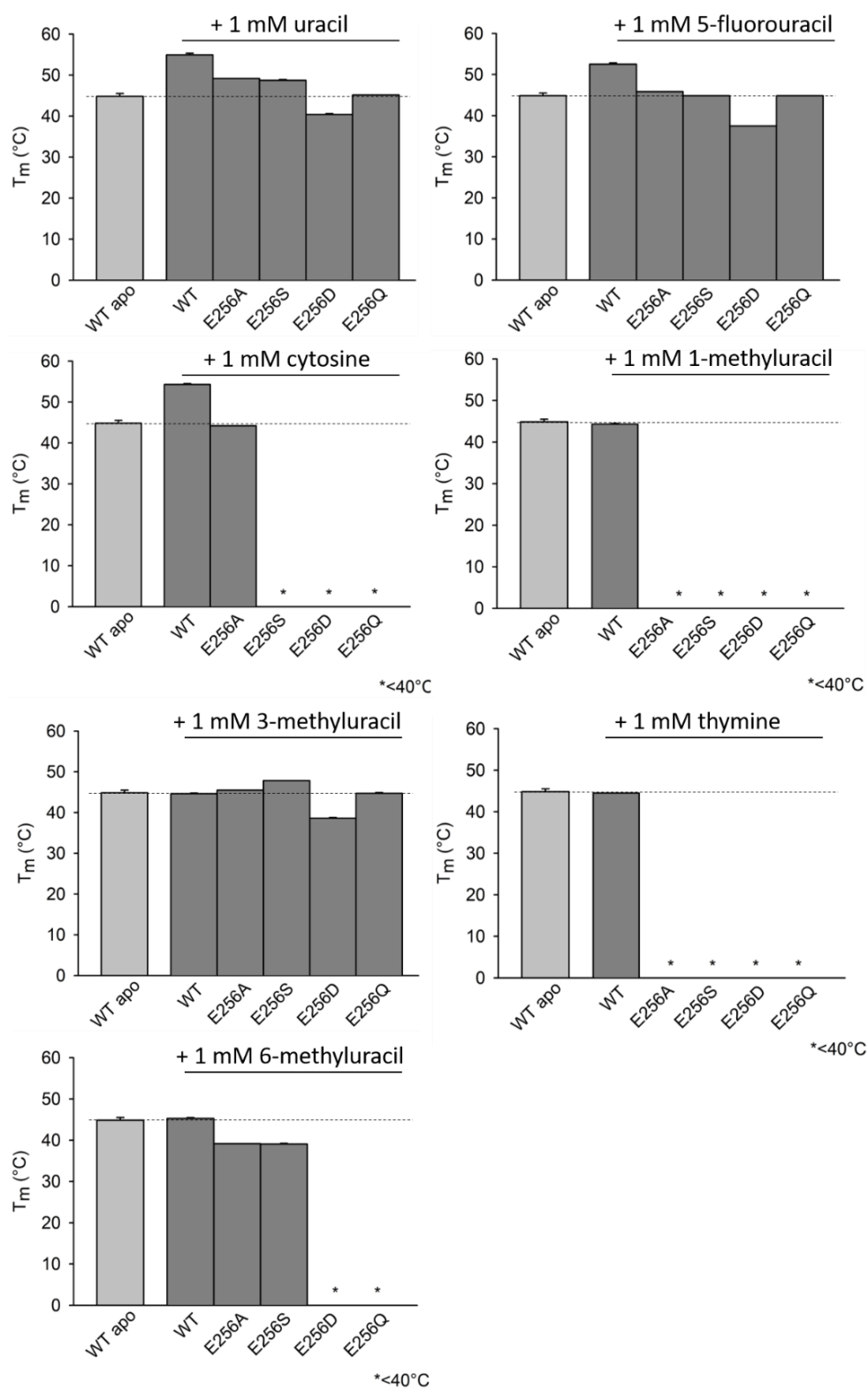
Figure 41: Binding specificities of E256 mutants of 51ThXi. 18 µM (0.5 mg/mL) SEC-pure protein were incubated with 1 mM of different uracil analogs and DSF was performed to determine a melting temperature (T_m). For comparison, the T_m of wild type 51ThXi obtained in the absence of substrate is given. As the T_m of all E256 mutants of 51ThXi in absence of substrate cannot be assessed by DSF, absolute melting temperatures are given here. In presence of several uracil analogs, no distinct T_m could be obtained, indicating that these are not able to bind to the protein or only with low affinity. In turn, if a distinct T_m can be determined, this suggests that this compound is binding to the mutant. All DSF experiments were performed in triplicates, errors shown represent standard deviations.

All mutants were purified and subjected to DSF measurements (**Figure S11**-**Figure S16**). Following what was observed for the E303A mutant, no binding of any uracil analog could be detected for neither E303D nor E303Q of 51ThXi. This indicates that not only the charge, but the precise residue at that position is relevant for protein functionality. In line with that, this glutamate was absolutely conserved within the 500 UraA homologs that were automatically picked for the ConSurf analysis shown in **Figure 32**. As is discussed in detail in section 4.3.3, in other SLC23 transporters, in particular those specific for purines such as xanthine and urate, this residue is found to be a glutamine. Due to the low solubility of most purines in water, these compounds are not suitable for use in DSF experiments as performed here.

Concerning the binding of uracil analogs, the mutants of E256 largely reproduce the results obtained for the E256A mutant (**Figure 41**). As already for E256A, no melting temperature of protein without substrate could be determined due to pronounced destabilization of the mutants compared to wild type 51ThXi. All four mutants were stabilized by uracil, 5-fluororuacil, and 3-methyluracil. In contrast, only E256A appeared to be able to bind cytosine. 6-methyluracil only stabilized the E256A and E256S mutants, but not E256Q or E256D. Interestingly, the stabilizing effect of the uracil compounds was the lowest for E256D, although this mutant conserves the charge of the glutamate found in wild type 51ThXi. In summary, removing the TM8 glutamate in any case reduced the stability of 51ThXi, but depending on which residue was introduced different degrees of destabilization were observed. Together with diverse binding behaviors of the different mutants, indicating a different binding site architecture or a certain modulation of the substrate binding site, the data suggests a delicate network of (probably electrostatic) interactions within the substrate binding site of 51ThXi. Disturbing this network by introducing mutations can lead to protein variants that are still able to bind substrate, but in no case were observed to preserve substrate transport.

4.3.2.4. The TM8 histidine can be replaced by aspartate

The two glutamates found in the substrate binding site of 51ThXi, that are conserved in UraA and 55PyPi, were shown in section 4.3.2.3 to be absolutely essential for substrate transport. Replacing H245 of UraA or H260 of 51ThXi, respectively, by alanine also caused a reduction of uracil uptake (section 4.3.2.1). Nevertheless, the histidine is not as thoroughly conserved as the two glutamates, as in about ~50% of homologs an aspartate is found instead (**Figure 32**). 55PyPi, the UraA homolog from *Pyramidobacter piscolens*, also harbors an aspartate on TM8. In uptake assays, cells overexpressing 55PyPi showed robust internalization of uracil. Interestingly, the D268A mutant was not inactive but rather exhibited wild type-like activity (**Figure 33C**). Consistently, the UraA(H245D) and 51ThXi(H260D) mutants were active in cell-based uracil uptake assays (**Figure 42**). Purified UraA(H245D) has a melting

temperature close to the one of wild type UraA (WT, 49.4 ± 0.2°C; UraA(H245D), 50.9°± 0.2°C) and shows a similar binding profile (**Figure S17**). The histidine on TM8 is found in proximity to the TM8 glutamate, which suggests that it might be important in compensating the negative charge, especially in the light of the presence of the other glutamate on TM10. In this line of thought, replacing histidine by aspartate appears controversial. In this case, water molecules might be involved in attenuation of charges and stabilization of the substrate binding site. Involvement of water molecules was already suggested to play a role in uracil binding by UraA in mediating hydrogen bond formation between H245 and the ligand (Lu *et al.*, 2011). In the occluded structure of UraA, two water molecules were modelled in the substrate binding site, one between S72, E241 and O4 of uracil, and the other between E241 and H245 (Yu *et al.*, 2017). An extension or remodeling of this water network might provide an explanation to why also aspartate is allowed instead of histidine in the substrate binding site of SLC23 uracil transporters.
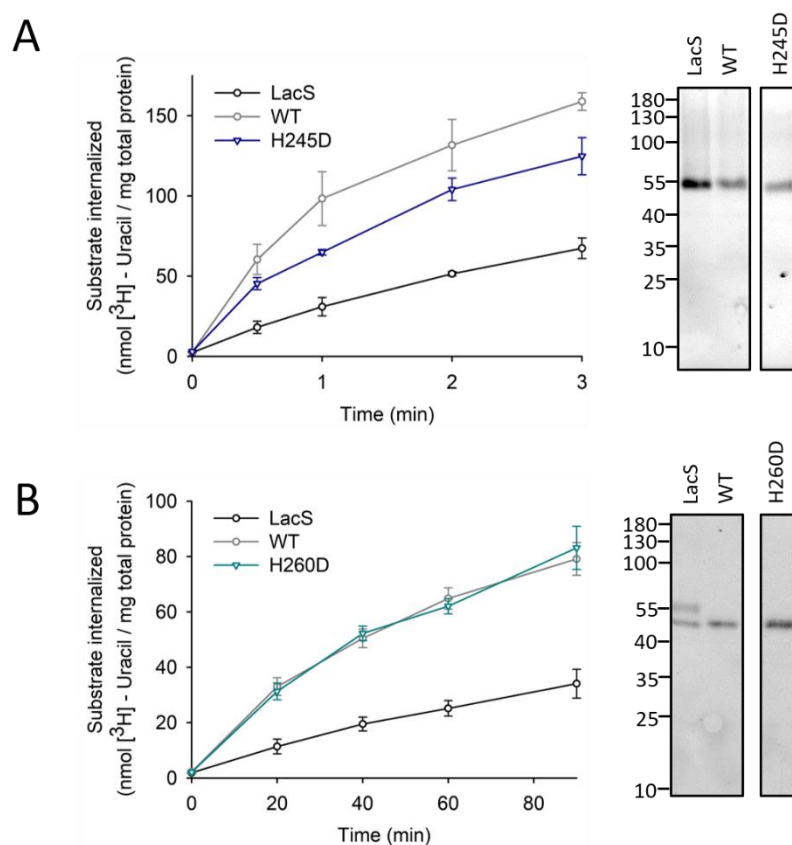
Figure 42: Mutating the histidine on TM8 to aspartate does not compromise activity of UraA and 51ThXi. (A) Uracil uptake in *E .coli* BW25113(*ΔuraA*) overexpressing GFP fusions of LacS as negative control (black), wild type UraA (grey), or UraA(H245D) (dark blue). Cells were incubated at 30°C with 2 µM partially radiolabeled uracil. Uptakes were measured three times, errors shown are standard deviations. Expression levels of uptake samples were monitored by 12% SDS-PAGE and *in gel* GFP fluorescence on whole cell lysates. (B) Uracil uptake in *E. coli* BW25113(*ΔuraA*) overexpressing GFP fusions of LacS as negative control (black), wild type 51ThXi (grey), or 51ThXi(H260D) (teal). Cells were incubated at 25°C with 1 µM partially radiolabeled uracil. Uptakes were measured three times, errors shown are standard deviations. Expression levels of uptake samples were monitored by 12% SDS-PAGE and *in gel* GFP fluorescence on whole cell lysates.

4.3.3. Changing the substrate specificity of SLC23 transporters

4.3.3.1. Differing binding site residues are suggested to govern the substrate specificities of SLC23 transporters

In section 4.3.2, the relevance of conserved substrate binding site residues for binding and transport in SLC23 uracil transporters was explored in detail. Both E241/E256 and H245/H260 located on TM8 in UraA/51ThXi were demonstrated to play a role in protein stability rather than being essential for substrate binding. The charged or polar substrate binding site residues were suggested to form a network of interactions, probably with the involvement of water molecules. Multiple binding site residues within the binding site contribute to (high-affinity) uracil binding, and removal of E241/E256 or H245/H260 still allows for substrate recognition by the transporter, although transport was markedly reduced. At the moment it is not clear whether a reduced affinity for uracil is responsible for this effect on the transport rate.
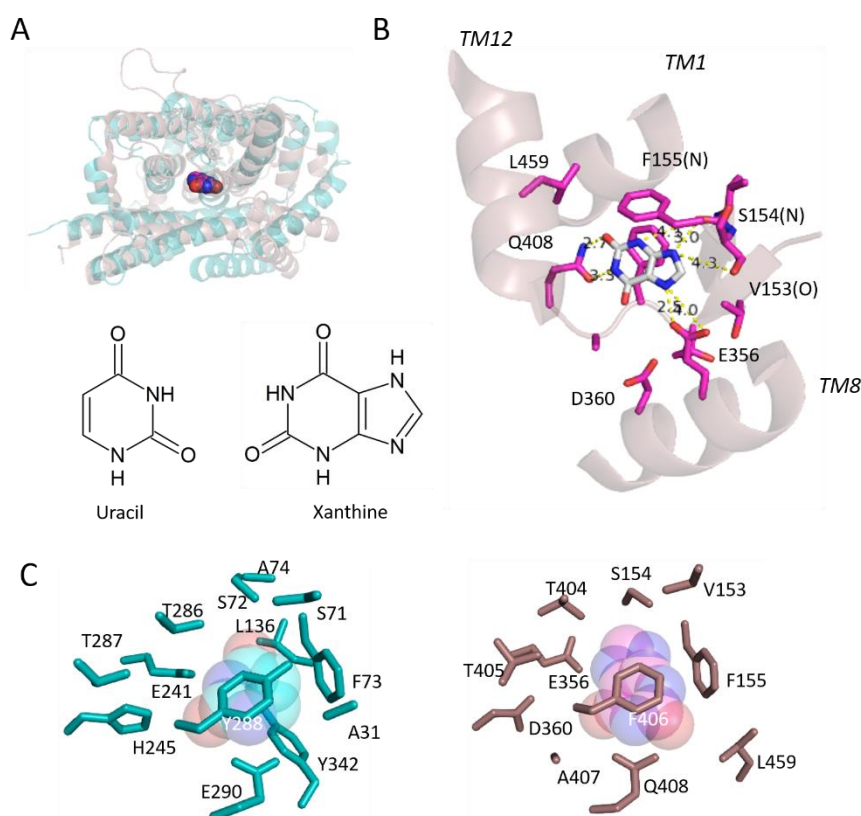
Figure 43: The UapA substrate binding site. (A) Superposition of the UraA (teal, PDB: 3QE7) and UapA (pink, PDB: 5I6C) structures. UapA was co-crystallized with xanthine, shown in spheres, that is located at a similar position as uracil in the UraA structure at the interface between the two domains. (B) Residues within 5 Å around xanthine (residues in dark pink sticks, ligand in light grey sticks). For simplification, only helices 1, 8 and 12 are indicated. Distances between opposite charges were measured using the built-in PyMOL tool. (C) Binding site residues within 5 Å of uracil in UraA (left side) and of xanthine in UapA (right side) shown in a similar orientation for better comparison. The ligands are indicated as transparent spheres, side chains of residues in sticks are annotated. At the position of H245 in UraA (TM8), in UapA D360 is located, E290 in UraA is Q408 in UapA.

Most of the substrate binding site residues are preserved throughout all SLC23 members, not only those specific for uracil. Less conserved residues were therefore suggested to determine the specificity of the transporters (Frillingos, 2012; Papakostas and Frillingos, 2012; Karena *et al.*, 2015; Chaliotis *et al.*, 2018). UapA, the other structurally resolved SLC23 transporter, was also crystallized in a substrate-bound conformation. Xanthine is binding at a similar position as uracil in UraA, at the interface between the substrate binding and scaffold domains (**Figure 43A**). The substrate binding site of UapA contains many residues that are found in corresponding positions in UraA. Exceptions are D360 on TM8, which is in place of H245 in UraA, as well as Q408 instead of E290 (**Figure 43B,C**). The two xanthine transporters of *E. coli*, XanQ and XanP, which have been intensively studied regarding their function, also harbor aspartate and glutamine residues on equivalent positions (Karatza and Frillingos, 2005; Karatza *et al.*, 2006; Papakostas *et al.*, 2008; Karena and Frillingos, 2011) (**Figure S18**). It has therefore been suggested that these residues determine the substrate specificities of the different nucleobase transporters. It should be noted that RutG, that exhibits a dual specificity for both uracil and xanthine, harbors a histidine residue on TM8 and a glutamate on TM10, resembling the UraA binding site.

To test whether 51ThXi could be changed from an uracil to a xanthine transporter, the respective 51ThXi(H260D/E303Q) mutant was generated. Xanthine uptake was performed in *E. coli* BW25113(*ΔxanP*) which lacks only one of the two xanthine transporters of *E. coli*. This strain shows lower background than the *ΔxanQ* strain and consequently a higher signal-to-noise ratio (not shown) and was also used in xanthine uptake experiments that are reported in literature.

The 51ThXi(H260D/E303Q) mutant did neither show uptake of uracil, nor xanthine (**Figure 44**). As already mentioned, the substrate binding site of *E. coli* RutG that shows a dual specificity for uracil and xanthine, resembles the one of UraA (histidine on TM8, glutamate on TM10). This already suggests that the substrate specificity is not only controlled by these two residues. Because RutG is missing F73, which is close to the substrate in UraA and probably interacting with O4 of uracil (although via the backbone amide nitrogen), it was suggested that F73 is restricting the substrate specificity of UraA (Botou *et al.*, 2018). The UraA(F73A) mutant was shown to bind, but not transport, thymine, which is not competing for the wild type UraA binding site in cell-based assays. No transport or binding of xanthine could be detected for UraA(F73A) in the same study. Nevertheless, the respective F82A mutation was included for a triple 51ThXi mutant (F82A/H260D/E303Q) in order to account for this observation made for UraA. However, also this triple mutant was incapable of uracil or xanthine transport (**Figure 44**).
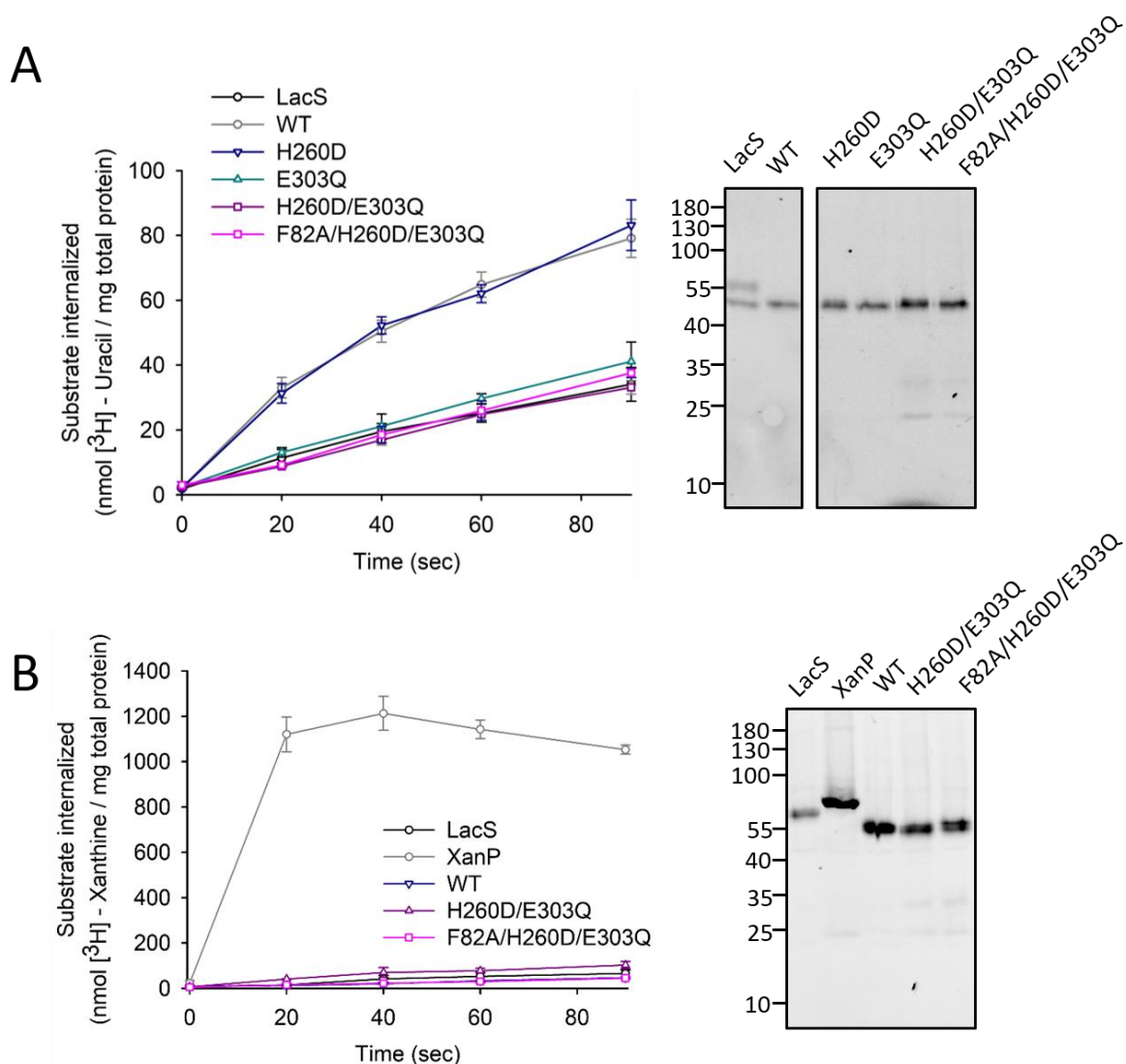
Figure 44: Uracil and xanthine uptake of 51ThXi(H260D/E303) and F82A(H260D/E303Q). (A) Uracil uptake in *E. coli* BW25113(*ΔuraA*) overexpressing GFP fusions of LacS as negative control (black), wild type 51ThXi (grey) or mutants of 51ThXi (H260D, dark blue; E303Q, teal; H260D/E303Q, dark pink; F82A/H260D/E303Q, pink). Cells were incubated at 25°C in presence of 1 µM uracil. (B) Uracil uptake in *E. coli* BW25113(*ΔxanP*) overexpressing GFP fusions of LacS as negative control (black), XanP as positive control (grey), wild type 51ThXi (dark blue) or mutants of 51ThXi (H260D/E303Q, dark pink; F82A/H260D/E303Q, pink). Cells were incubated at 20°C in presence of 4 µM xanthine. Uptakes were measured three times, errors shown represent standard deviations. Expression levels of uptake samples were monitored by 12% SDS-PAGE and *in gel* GFP fluorescence on whole cell lysates.

4.3.3.2. RutG substrate specificity

The UraA homolog RutG of *E. coli* was reported to transport both uracil and xanthine (Botou *et al.*, 2018). In order to gain more insight into the determinants of RutG substrate specificity, a competition assay in whole cells was performed using different uracil derivatives (**Figure 45**). Reproducibly, purification of RutG gave only low yields, and no melting temperature could be determined using DSF (supplement **Figure S19**). In whole cells, similar uracil analogs compete for the substrate binding site of RutG as for the one of UraA (5-fluorouracil, cytosine). Additionally, reduced uptake of uracil was

obtained in the presence of excess thymine, which is in line with the reduced uracil uptake in presence of excess thymine observed by Botou *et al.*. No competition was observed for xanthine, which is in contrast to the results reported by Botou *et al.* An explanation for this might be the low solubility of xanthine water which only allowed to use an 500x excess over uracil.
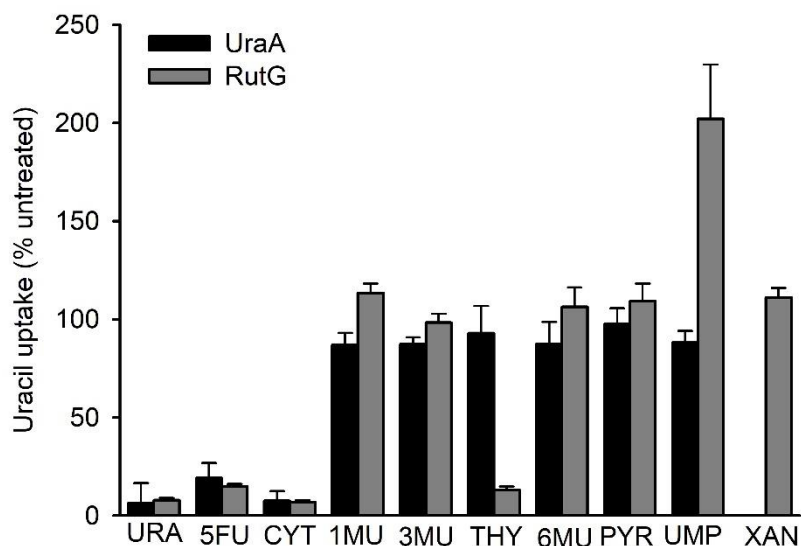
Figure 45: Substrate specificity of RutG. Uracil uptake in *E. coli* BW25113 *ΔuraA* overexpressing GFP fusions of UraA (black) or RutG (grey) was determined in presence of different uracil analogs. Cells were incubated at 30°C with 0.1 µM uracil and 100 µM competitor for 10 sec in both cases. For xanthine, 50 µM final concentration were used due to lower solubility in water. Shown is the remaining uracil uptake related to uptake by cells without any addition of competitor. Uptake was measured in triplicates, errors are derived from error propagation of standard deviations. URA: uracil, 5FU: 5-fluorouracil, CYT: cytosine, 1MU: 1-methyluracil, 3MU: 3-methyluracil, THY: thymine, 6MU: 6-methyluracil, PYR: pyrimidine, UMP: uridine monophosphate, XAN: xanthine.

According to the sequence alignment, the residues of the substrate binding sites of UraA and RutG are, apart from F73 in UraA that is A88 in RutG, identical (**Figure S18**). Botou *et al.* suggest that A88 in RutG is responsible for the larger substrate specificity of RutG. As already mentioned above, the UraA(F73A) mutant was shown in whole-cell based competition assays to bind thymine, but xanthine binding was still prohibited (Botou *et al.*, 2018). Herein, thymine binding to purified wild type UraA was demonstrated using DSF (section 4.3.1). It should be noted that the known xanthine transporters XanQ, XanP and UapA also harbor a phenylalanine at the position corresponding to F73 in UraA, which makes it implausible that A88 in RutG is the major determinant of substrate specificity of this transporter as was suggested. As the substrate binding site residues of UraA/51ThXi and RutG yet are, with the exception of F73/A88, absolutely identical, it is at the moment unclear what governs the larger substrate specificity of RutG compared to UraA. The so far available data suggests that determinants for substrate specificity might also be located outside the substrate binding site.

4.3.3.3. Determination of substrate specificities of novel SLC23 homologs

Neither changing the substrate binding site residues of the uracil transporter 51ThXi to the ones found in the xanthine transporter UapA, XanQ and XanP (section 4.3.3.1), nor a more detailed analysis of the substrate specificity of the dual-specificity transporter RutG (section 4.3.3.2) revealed new information on what governs the substrate specificity of SLC23 uracil versus xanthine transporters. All the data available to date has been derived from only a small set of homologs, two of them on a structural and functional basis (UraA and UapA), while the other only have been studied regarding their function. To increase the number of homologs with known specificity, the substrate preference of several new SLC23 homologs was determined. For this, several homologs originating from the same organism were used, reasoning that these would presumably exhibit different specificities. In total, 20 homologs were tested for uptake of both uracil and xanthine. 8 of them were identified as uracil transporters, 8 were specific for xanthine, for 3 none of the two were transported, and one homolog from *Yersinia enterocolitica* transported both uracil and xanthine (**Figure 46**).
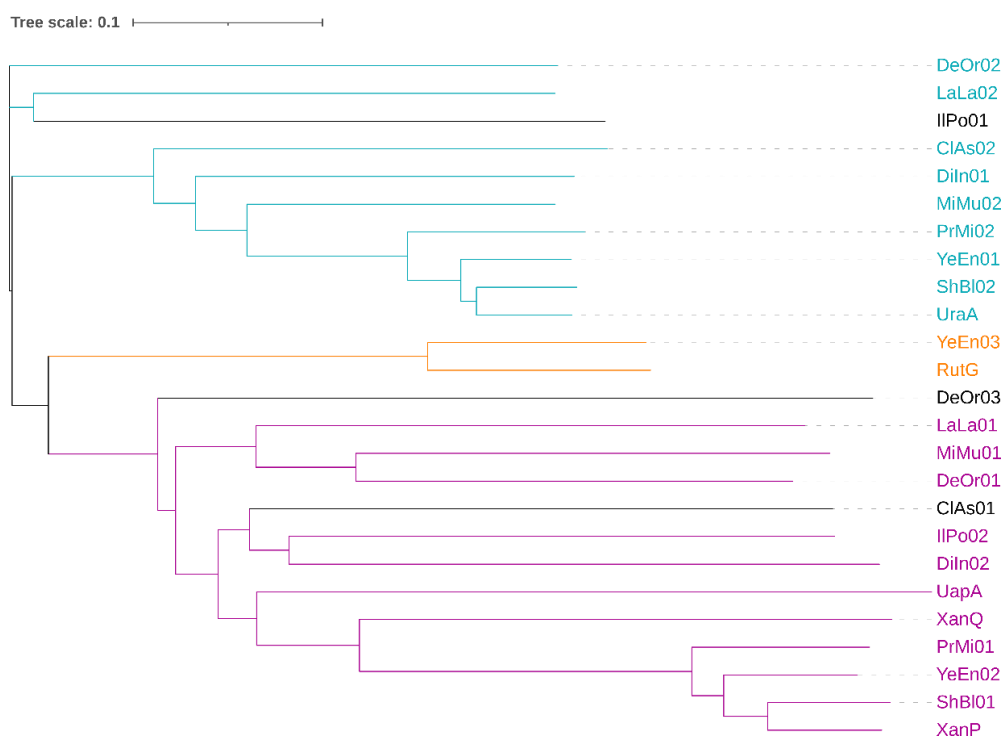
Figure 46: **Substrate specificities of novel SLC23 homologs**. 20 new UraA homologs were tested for uptake of uracil and xanthine in *E. coli* BW25113(*ΔuraA*) or *E. coli* BW25113(*ΔxanP*), respectively. Alignment of homologous sequences as well as sequences of known uracil and xanthine transporters (UraA, RutG, XanQ and XanP of *E. coli* and UapA from *Aspergillus nidulans*) was generated using Clustal Ω, the phylogenetic tree was generated using iTOL (https://itol.embl.de/). Homologs are annotated with an abbreviation for the organism they were derived from as well as an arbitrary number (DeOr: *Desulfosporosinus orientis,* LaLa: *Lactococcus lactis*, IlPo: *Ilyobacter polytropus*, ClAs: *Clostridium asparagiforme*, DiIn: *Dialister invisus*, MiMu: *Mitsuokella multacida*, PrMi: *Proteus mirabilis*, YeEn: *Yersinia enterolitica*, ShBl: *Shimwellia blattae*). Homologs that transported uracil are colored in teal, the ones transporting xanthine in dark pink, and the two that showed uptake for both (RutG and YeEn03) in orange. For three of them, indicated in black (IlPo01, DeOr03 and ClAs01), no uptake of neither compound was detected.

The sequences of the homologs and their identities to known uracil/xanthine transporters as well as the uptake data and assessment of expression levels can be found in the supplement (**Table 31**, **Figure S20** and **Figure S21**). According to the phylogenetic tree shown in **Figure 46**, the homologs form three larger clusters, one containing UraA, one RutG, and one the xanthine transporters XanQ, XanP and UapA. The analysis of the substrate preference of the new homologs reveals that indeed the first cluster contains transporters specific for uracil and the second larger one those specific for xanthine. One homolog, closest related in sequence to *E. coli* RutG, was shown to transport both uracil and xanthine.

Strikingly, all homologs that were shown to transport uracil harbor a histidine and a glutamate on TM8 and 10, respectively, resembling the UraA binding site, while the transporters specific for xanthine (almost) exclusively harbor an aspartate on TM8 and glutamine on TM10 (**Figure 47**). An exception is LaLa01 that has a valine on the position of histidine/aspartate in TM8 but was still shown to be specific for xanthine. Replacement of the histidine on TM8 of uracil transporters by aspartate has been shown to not affect substrate binding or transport, although both H245A and H260A mutants of UraA and 51ThXi, respectively, were inactive (section 4.3.2.4). On the other hand, the D268A mutant of 55PyPi showed uracil uptake similar to the one of wild type 55PyPi. The finding that all xanthine transporters identified herein combine an aspartate on TM8 with glutamine on TM10 suggests a general relevance of these residues for the substrate specificity; yet the H260D/E303Q mutant of the uracil transporter 51ThXi did not show any detectable xanthine transport activity (section 4.3.3.1).

From the uracil and xanthine transporters identified here, no further insights can be obtained when only considering the residues close to the substrate. Except for the already discussed differences in TM8 (histidine/aspartate in uracil/xanthine transporters) and TM10 (glutamate/glutamine in uracil/xanthine transporters), all other substrate binding site residues are either closely related, or not conserved (**Figure 47**). In general, the substrate binding site of uracil transporters seems to be more conserved than the one of xanthine-specific transporters. This is also reflected in the overall sequence identity (supplement **Table 31**). In general, transporters specific for uracil share a high sequence identity with UraA. Sequence identities for the xanthine transporters on the other hand range from high (e.g., ShBl01 shares an identity of 87.66% with XanP) to moderate (e.g., sequence identities of DeOr01 and DiIn02 is about 30% to both XanQ and XanP). This might indicate that only small changes on the sequence level are allowed to preserve the functionality of uracil transporters. RutG that transports both uracil and xanthine shares a sequence identity of ~37% with UraA but only ~27 and ~23 with XanQ and XanP, respectively.
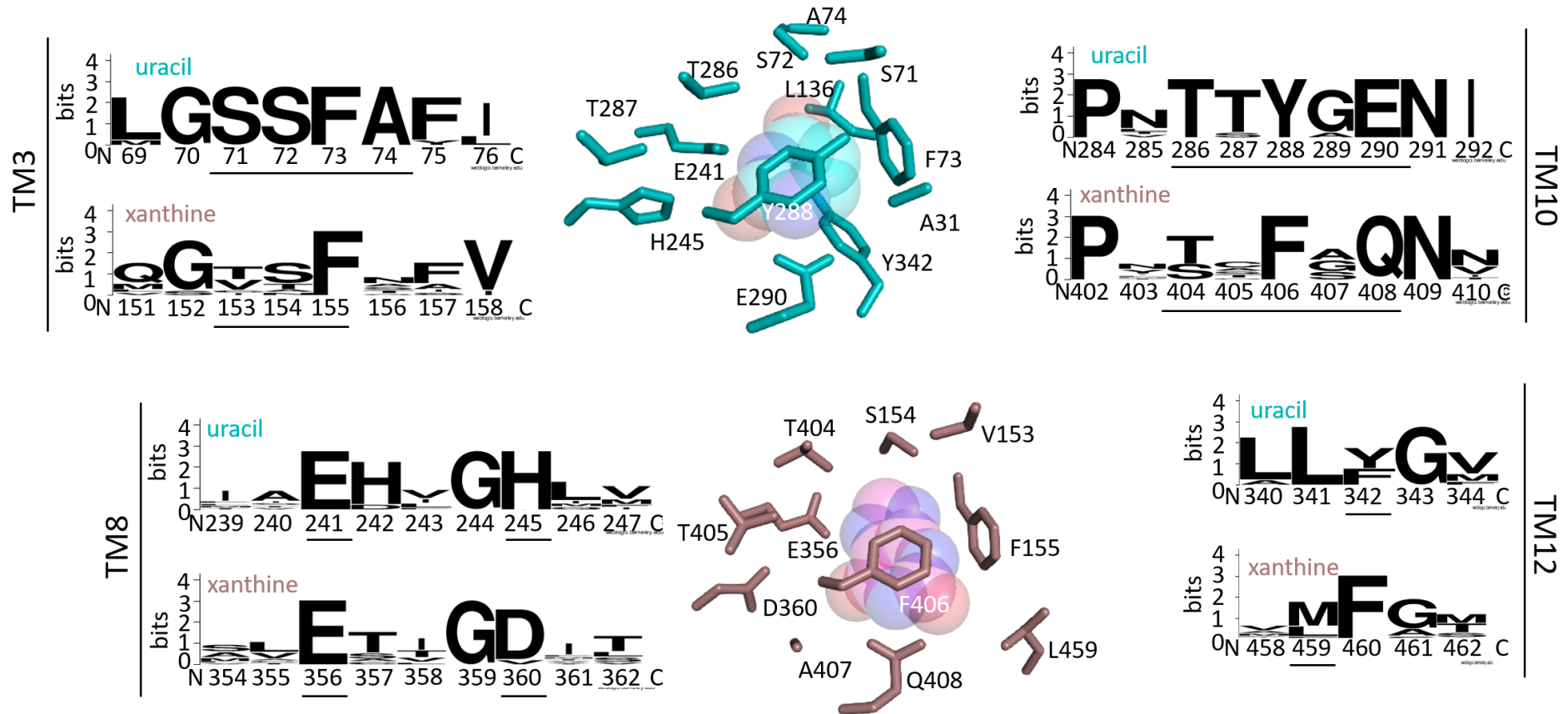
Figure 47: Predicted substrate binding site residues of newly identified uracil and xanthine transporters. To assess similarities and differences that might be responsible for the distinct substrate specificities, sequence alignments of newly identified uracil or xanthine transporters were generated using ClustalΩ and respective residues found at the substrate binding sides of UraA (top, in teal) or UapA (bottom, in dark pink) were analyzed. Sequence alignments of relevant positions are depicted as sequence logos generated using WebLogo (https://weblogo.berkeley.edu/logo.cgi) (Schneider and Stephens, 1990) with residues within 5 Å of the substrate underlined. Numbering for the uracil transporters as in UraA, for the xanthine transporters as in UapA.

Results

The number of transporters for which either uracil or xanthine has been identified as a substrate is still considerably small, although it greatly expands the number of so far functionally characterized SLC23 nucleobase transporters. In order to assess whether the homologs are representative for SLC23 uracil or xanthine transporters, the conservation of their sequences was compared to a ConSurf analysis of relevant residues on 500 automatically picked homologs of UraA or UapA, respectively (**Figure S23** and **Figure S24**). The sequence conservation of the uracil transporter homologs that have been functionally characterized in the here presented study fits extremely well to the results from the ConSurf analysis performed on UraA. More variation was observed for the ones specific for xanthine, but also the ConSurf analysis on UapA reveals a less overall conservation of residues close to the substrate binding site.

In summary, the here presented data clearly shows that specificity does not only rely on the presence of the distinct binding site residues (histidine/glutamate on TM8/10 in uracil transporters, aspartate/glutamine in equivalent positions in xanthine transporters). To support this hypothesis, several lines of evidence have been provided: First, mutants of the uracil transporter 51ThXi that resemble the xanthine transporter binding site were found to neither transport xanthine, nor uracil; second, the TM8 histidine in both UraA and 51ThXi could be replaced by aspartate without a general impairment in uracil transport activity; third, the binding sites of the dual-specificity transporters RutG and YeEn03 resemble the UraA binding site, i.e. they harbor a histidine on TM8 and a glutamate on TM10 near the substrate. The identification of specificities of a set of novel SLC23 homologs that showed transport of either uracil or xanthine did not reveal immediate new insights into this.

4.4.   Relevance of the dimeric state of SLC23 transporters

4.4.1.   Oligomeric state of SLC23 transporters

The oligomeric state of both reconstituted UraA and 51ThXi was tested by unspecific glutaraldehyde crosslinking. Glutaraldehyde is commonly used for protein crosslinking as it readily reacts with nucleophiles under a variety of conditions. In proteins, mostly primary amines of lysine residues are involved in this (Migneault *et al.*, 2004). Interestingly, only a small dimer fraction was observed for reconstituted 51ThXi (**Figure 48A**). In contrast, glutaraldehyde crosslinking of UraA in proteoliposomes showed distinct dimerization that was reduced when the liposomes were disrupted by a pre-incubation step with 1% (w/v) SDS (**Figure 48B**).
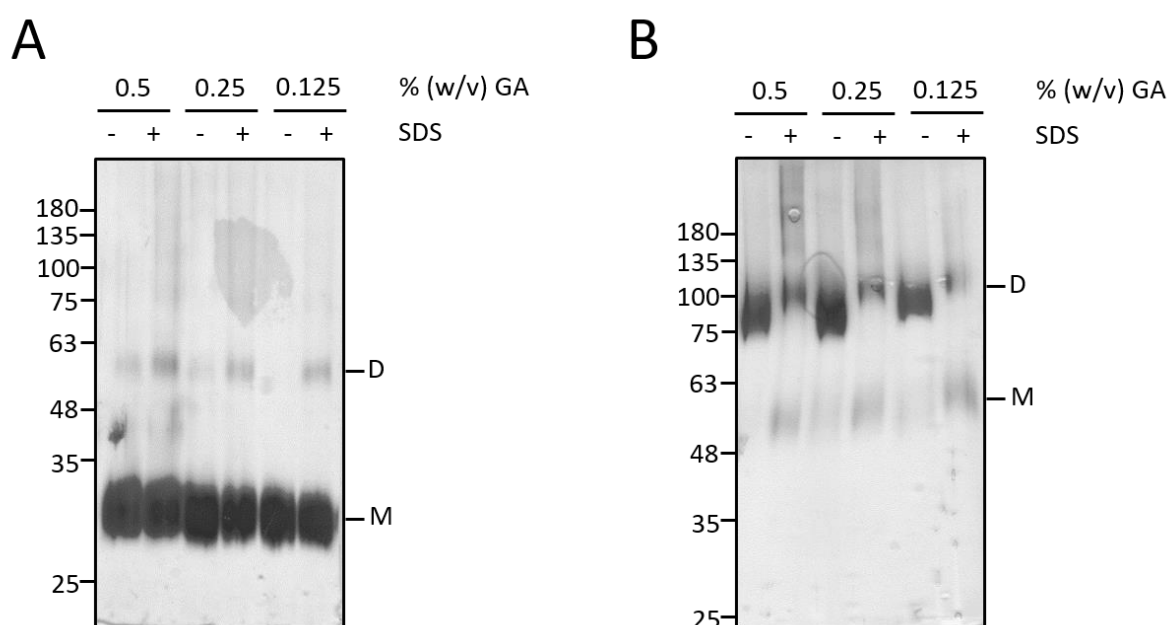
Figure 48: Oligomeric state of reconstituted SLC23 transporters. (A) Unspecific crosslinking using glutaraldehyde (GA) on 51ThXi reconstituted to soy PC 1:50 protein-to-lipid ratio. Proteoliposomes were incubated with different amounts of GA for 20 min at RT, either with or without a pre-incubation step with 1 % (w/v) SDS as indicated by plus (+) and minus (-). Samples were separated on 10% SDS-PAGE and proteins were visualized by silver staining. Monomer and dimer are indicated by M and D, respectively. (B) Same as in (A), but using proteoliposomes of reconstituted UraA.

Since no structure of the 51ThXi protein is available, a Phyre2 model was generated using UraA as template (PDB: 3QE7) (Kelley *et al.*, 2015). In order to obtain a dimeric structure, this Phyre2 model was aligned to the UraA dimer (PDB: 5XLS) using PyMOL. According to the model, four of five lysine residues are located at the substrate binding domain of 51ThXi and none at the scaffold domain (**Figure S25A**). The fifth lysine is located on a C-terminal amino acid stretch not found in UraA and is therefore not included in the model. Of the four lysines, three are located near to each other, which might favor intra- instead of interprotomer crosslinking. The unfavorable location of the lysines in

51ThXi might provide an explanation to why no glutaraldehyde-induced crosslinking of 51ThXi in proteoliposomes was observed. UraA, which readily crosslinked in presence of glutaraldehyde, contains 14 lysine residues per protomer, with a number of them located at the scaffold domain (**Figure S25B**).

To further analyze dimerization of 51ThXi, cysteine mutants for specific inter-protomer crosslinking were constructed based on the 51ThXi dimer model described above. Three positions at the putative interface were chosen that combine to two double-cysteine mutants with suitable distances for crosslinking using copper phenanthroline (< 5Å for the $C_{ß}$-Cß distance) (**Figure 49**) (Hazes and Dijkstra, 1988). According to the model, the $C_{ß}$-$C_{ß}$ distance for the G356C_I387C mutant is 1.8 Å, the one for the G356C_G383C mutant is 7.8 Å, which is higher than the ideal distance. The $C_{ß}$-$C_{ß}$ distance determined for the corresponding positions in UraA, however, is 2.8 Å, which is why this mutant was included for 51ThXi nevertheless.
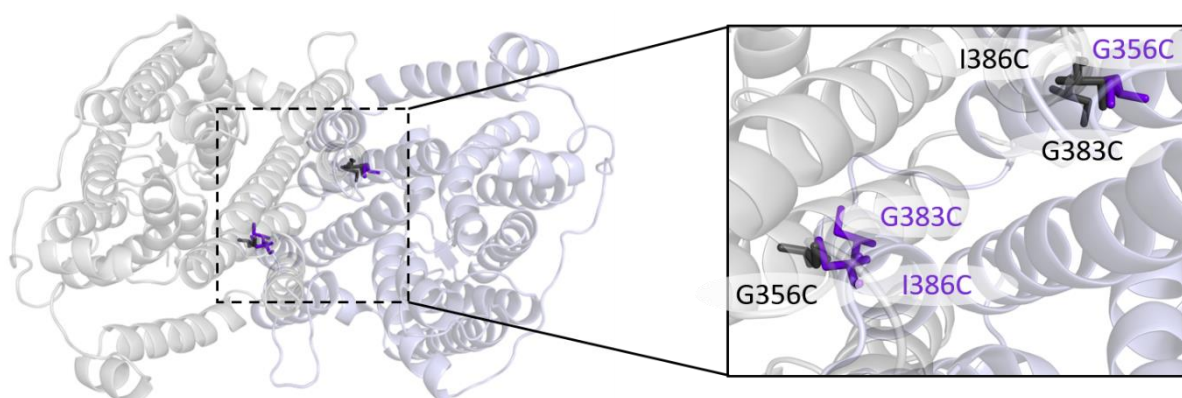
51ThXi	$C_{ß}$ - $C_{ß}$	UraA	$C_{ß}$ - $C_{ß}$
G356C – G356C	19.7 Å	G343C – G343C	18.9 Å
G383C – G383C	18.3 Å	S370C – S370C	21.2 Å
I387C – I387C	19.5 Å	I374C – I374C	17.2 Å
G356C – G383C	7.8 Å	G343C – S370C	2.8 Å
G356C – I387C	1.8 Å	G343C – I374C	4.2 Å

Figure 49: Location of cysteine mutants in 51ThXi to test dimerization. A dimer model of 51ThXi based on the UraA dimer (5XLS) was built using a Phyre2 model (template: 3QE7). To test dimerization of 51ThXi by inter-protomer crosslinking, cysteine residues were introduced at the putative dimer interface with suitable $C_{ß}$-$C_{ß}$ distances as indicated in the table. The protein is shown in top view with one protomer in light grey, the other in light blue, and cysteines indicated as sticks.

Disulfide bridge formation indicating 51ThXi dimerization was first explored in vesicles obtained from whole cells. GFP-fusion proteins of single- and double cysteine mutants of 51ThXi were expressed in *E. coli* MC1061. Cells were lysed in presence of 10 mM dithiothreitol (DTT), which was subsequently removed using BioSpin 6 columns before cell lysates were incubated with copper phenanthroline (CuPhen) to induce cysteine crosslinking. The lysates were also incubated with PEG-5K-maleimide (PEG-5K) to monitor accessibility and reactivity of the cysteines during the assay. Modification of cysteines with PEG-5K results in a prominent upshift of the protein on gel. The samples were subjected to 8% SDS-PAGE and proteins were visualized by *in gel* GFP fluorescence (**Figure 50A**).
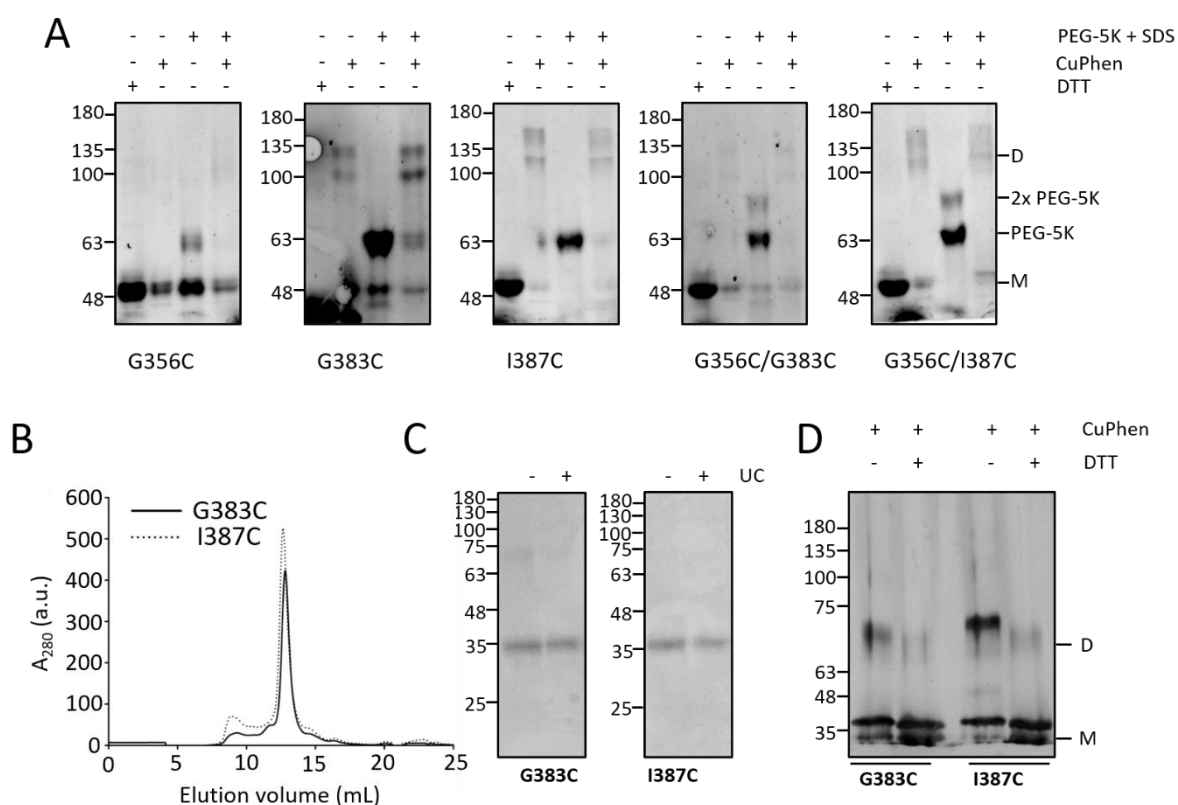
Figure 50: Specific cysteine crosslinking of 51ThXi. (A) Dimerization of 51ThXi in whole cell vesicles. Cysteine mutants of 51ThXi were expressed as GFP fusion proteins. Whole cell lysates were incubated with DTT, copper phenanthroline (CuPhen) and/or PEG-5K-maleimide (PEG-5K) together with 1% (w/v) SDS, as indicated by minus (-) and plus (+). Monomeric and dimeric 51ThXi is indicated by M and D, respectively, as well as one- and two-fold PEGylated protein. Samples were separated on 8% SDS-PAGE and proteins were visualized by *in gel* GFP fluorescence. (B) Size-exclusion chromatogram of 51ThXi(G383C) (solid line) or 51ThXi(I386C) (dashed line) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 5 mM ß-mercaptoethanol, 0.2% (w/v) DM. (C) Reconstitution efficiency of 51ThXi(G383C) and 51ThXi(I387C). Proteoliposomes were incubated with 1% (w/v) DDM and protein amount before and from the supernatant after ultracentrifugation (UC) as indicated with minus (-) and plus (+) were compared on gel. Proteins were separated on 12% SDS-PAGE and visualized by Coomassie staining. (D) Crosslinking of 51ThXi (G383C) and 51ThXi(I387C) in proteoliposomes. Proteoliposomes were incubated with 40x CuPhen (molar ratio over protein amount) with or without additional 10 mM DTT as indicated by minus (-) and plus (+). Monomers and dimers of 51ThXi are indicated by M and D, respectively. Proteins were subjected to 8% SDS-PAGE and visualized by silver staining.

Dimer formation as suggested by a protein signal at ~100 kDa on gel was observed for the G356C/I387C mutant of 51ThXi and also, weakly, for the G356C/G383C mutant. Interestingly, also the G383C and I387C single cysteine mutants appeared to crosslink, although the predicted $C_ß$-$C_ß$ distances are more than 18 Å apart. This indicates either an inaccuracy of the 51ThXi dimer model, or a certain flexibility of the dimer interface. Despite that, the results imply that 51ThXi forms a dimer in the cellular membrane.

Because also the single cysteine mutants G383C and I387C showed crosslinking in whole cell vesicles, these mutants were chosen for purification and reconstitution to soy PC (**Figure 50B, C**). Both proteins were purified in the presence of 5 mM ß-mercaptoethanol. After reconstitution, proteoliposomes were washed three times to remove the reducing agent. Both mutants showed high degrees of crosslinking in proteoliposomes when being incubated with CuPhen, indicating that 51ThXi also forms dimers in the soy PC liposomes (**Figure 50D**). The dimer formation could be reversed by adding 10 mM DTT.

The oligomeric state of SLC23 has not been well studied so far, but both UraA and UapA have been crystallized as homodimers. Since all elevator proteins that have been identified to date were shown to form higher oligomers, the oligomeric state has been suggested to be relevant for the elevator transport mechanism in general (Holzhüter and Geertsma, 2020). The proteoliposome-based activity assay for the uracil transporter 51ThXi provided the opportunity to study the relevance of dimer formation for substrate transportation in the SLC23 family in more detail.

4.4.2. Interfering with the dimer interface interferes with transport activity

As stated above, all elevator proteins that have been identified so far were found to form dimers or trimers. In this, as one defining characteristic of elevator proteins, every protomer harbors its own complete substrate binding site and translocation pathway. From this angle, oligomerization appears not to be a necessity for substrate transport. Accordingly, a monomerized version of the sodium-proton antiporter NhaA, an elevator protein from *E. coli*, retained wild type like activity (Rimon *et al.*, 2007). Monomers of UraA, in contrast, showed little activity in whole cell-based radioactive uracil uptake assays while substrate binding was unaffected (Yu *et al.*, 2017).

Yu *et al.* created UraA monomers by introducing bulky tryptophans at the dimer interface, an approach that had already been successfully used on a chloride-proton antiporter of the ClC family (Robertson *et al.*, 2010). Introducing tryptophans at the interface may not only create sterical clashes, but also attract lipids that further prevent dimerization ("warts-and-hook" strategy). In order to create monomers of 51ThXi in a similar fashion, based on the dimer model of 51ThXi, two double tryptophan

mutants were constructed (A153W/I387W and L379W/I387W) (**Figure 51A**). Both mutants were purified by IMAC and SEC and efficiently reconstituted to soy PC (**Figure 51B**).
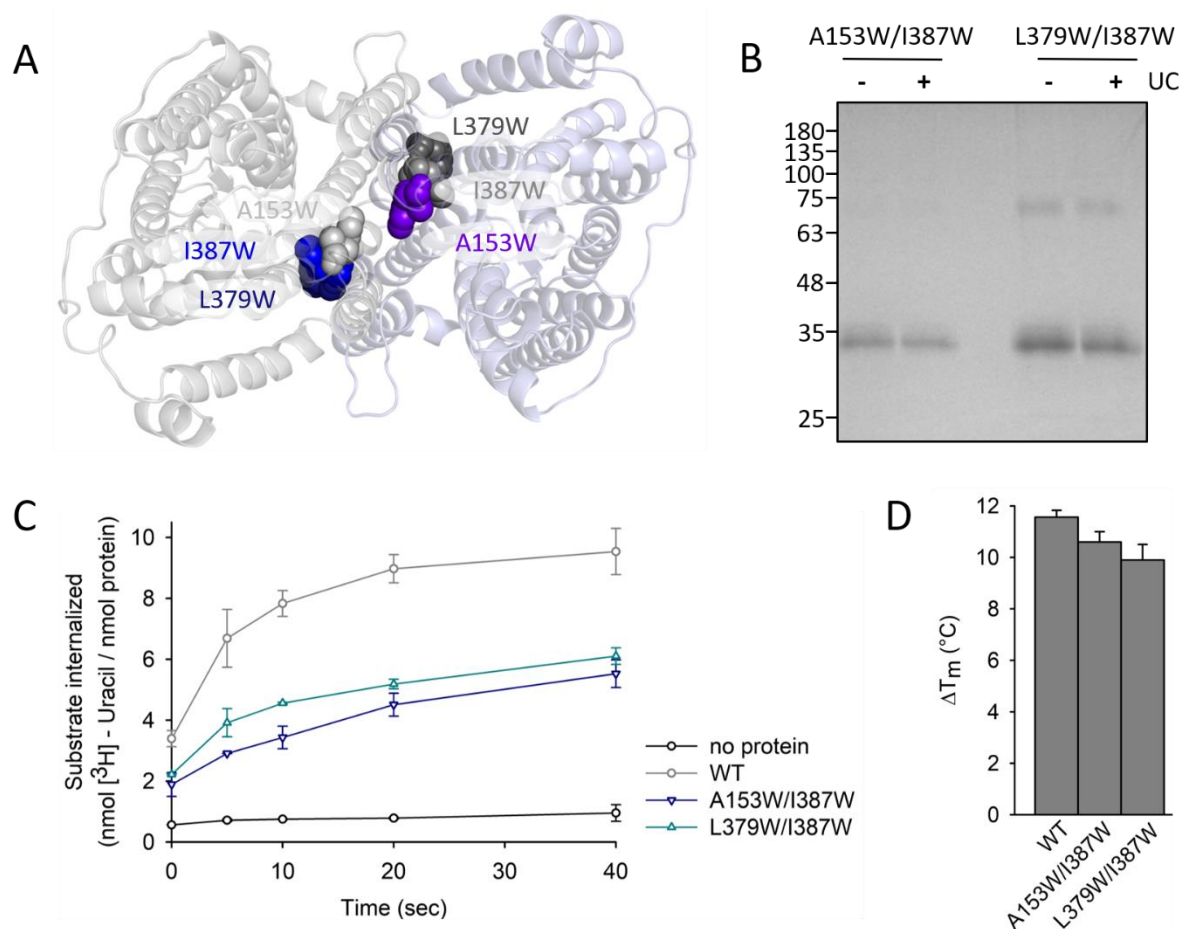
Figure 51: Effect of tryptophan mutants at the 51ThXi dimer interface. (A) Tryptophans were introduced at the dimer interface of 51ThXi to interfere with dimerization. The dimer model is shown in top view, with one protomer in light grey, the other in light blue. Tryptophans are shown in spheres and annotated. (B) Efficiency of reconstitution of 51ThXi(A153W/I387W) and 51ThXi(L379W/I387W) to soy PC. Proteoliposomes were solubilized with 1% (w/v) DDM and protein amounts before and in the supernatant after ultracentrifugation (UC) as indicated by minus (-) and plus (+) were compared on gel. Proteins were separated on 12% SDS-PAGE and visualized by Coomassie staining. (C) Uracil uptake in proteoliposomes of 51ThXi(A153W/I387W) (dark blue) or 51ThXi(L379W/I387W) compared to liposomes containing no protein (black) or wild type 51ThXi (grey). Uptake of 1 µM partially radiolabeled uracil was monitored at 10°C in triplicates, errors represent standard deviations. (D) Stabilization of wild type 51ThXi, 51ThXi(A153W/I387W) or 51ThXi(L379W/I387W) in presence of uracil in standard DSF experiments. Difference in melting temperature (ΔT_m) was obtained by subtracting T_m of purified protein obtained in absence of any substrate from the one obtained when protein was pre-incubated with 1 mM uracil. Melting temperatures were determined in triplicates, errors shown were obtained from error propagation of standard deviations.

Uptakes in proteoliposomes show a reduced activity of both mutants compared to wild type 51ThXi, but higher substrate internalization than the empty liposome control (**Figure 51C**). Both mutants showed a similar melting temperature as wild type protein and the T_m was shifted to a similar degree in the presence of 1 mM uracil, indicating that the overall protein fold was not compromised (**Figure 51D**). A similar intermediate activity was also observed in whole cell-based uptake assays in

E. coli BW25113(*ΔuraA*) overexpressing the 51ThXi double tryptophan mutants, supporting the results from the *in vitro* assay (**Figure 52**)
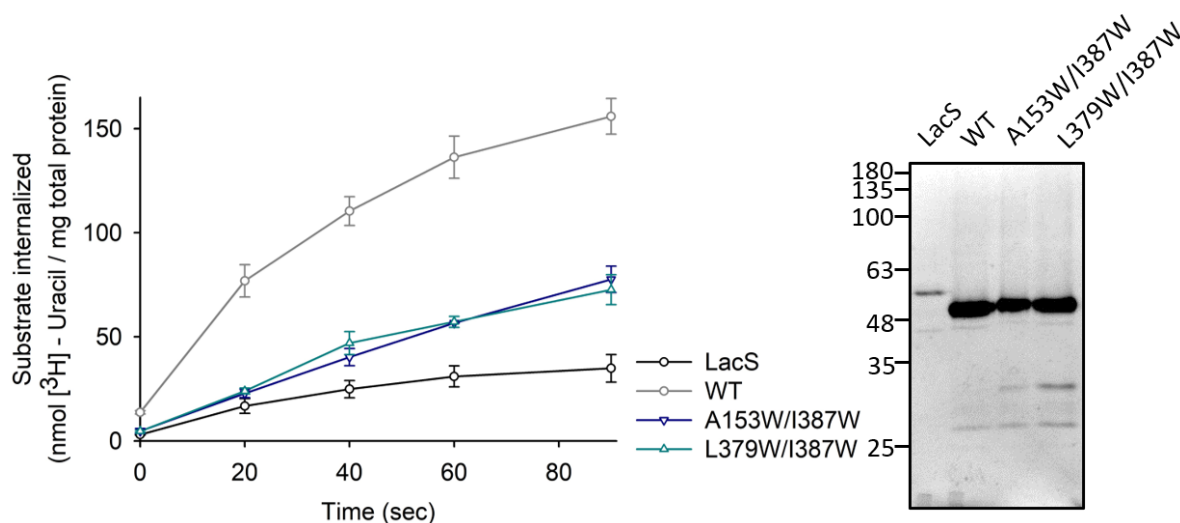
Figure 52: Intermediate activity of 51ThXi tryptophan mutants in *E. coli* BW25113(*ΔuraA)*. Uracil uptake in *E. coli* BW25113(*ΔuraA*) expressing LacS as negative control (black), wild type 51ThXi (grey) or the two mutants of 51ThXi harboring tryptophan mutants at the dimer interface (51ThXi(A153W/I387W), dark blue, or 51ThXi(L379W/I387W), teal). Uptake experiments were performed in technical triplicates, errors shown represent standard deviations. Expression levels were monitored by *in gel* GFP fluorescence on whole cell lysates separated on 12% SDS-PAGE.

Unfortunately, it was not possible to determine whether or to what extent the tryptophan mutations led to a monomerization of 51ThXi. As mentioned above, unspecific GA crosslinking of 51ThXi in proteoliposomes only showed a small amount of dimer fraction for wild type 51ThXi and did not provide more information on the oligomeric state of the tryptophan mutants (**Figure S26**). For the UraA monomerization mutants, the degree of monomerization was judged based on SEC, though this may not necessarily reflect the oligomeric state in the membrane (Yu *et al.*, 2017). A similar approach was not feasible in the case of 51ThXi as it appeared exclusively monomeric on SEC in contrast to wild type UraA, for which usually also a dimer peak with higher elution volume was observed (**Figure 17**). The monomeric state of wild type 51ThXi in detergent was confirmed using SEC-MALS (performed by Dr. Rupert Abele) (**Figure S27**). In a last attempt, the interface tryptophan mutations were combined with the cysteine mutations that were used to demonstrate dimerization of 51ThXi in proteoliposomes (51ThXi(G383C) and (G356C/G383C)) (**Figure 50**). Cysteine crosslinking in whole cell vesicles was tested for the resulting A153W/G383C/I387W, A153W/G356C/G383C/I387W, L379W/G383C/I387W and G356C/L379W/G383C/I387W mutants of 51ThXi. However, a similar degree of crosslinking of mutants with and without the additional tryptophan mutations was observed (**Figure S28**). This crosslinking data suggests that the interface of the protein might rather be distorted than completely disrupted. Since protein stability and uracil binding of the tryptophan mutants appear unaffected as seen in DSF

measurements (**Figure 51D**), the reduced transport activity in both proteoliposomes and whole cells thus would be a result of this altered dimer interface.

The cysteine mutants created for crosslinking of the 51ThXi dimer were also used to test the effect of constricting the interface on transport activity. Similar "stapling mutants" of the elevator-like Na$^+$/succinate transporter vcINDY showed similar or slightly reduced activity (2-5x) compared to wild type protein, depending on which position the crosslink between the scaffold domains was introduced (Mulligan *et al.*, 2016). In elevator proteins, substrate translocation is postulated to being mediated mostly by movement of the substrate binding domain against the scaffold domain, which remains relatively static. Consequently, constricting the dimer interface is supposed to have little effect on transport activity. For 51ThXi, two double cysteine mutants were generated (G356C/G383C and G356C/I387C, section 4.4.1). Because also the single-cysteine mutants G383C and I387C of 51ThXi showed crosslinking in whole cell vesicles and proteoliposomes, these mutants were tested for uracil uptake in proteoliposomes (**Figure 50B, C**; **Figure 53**).
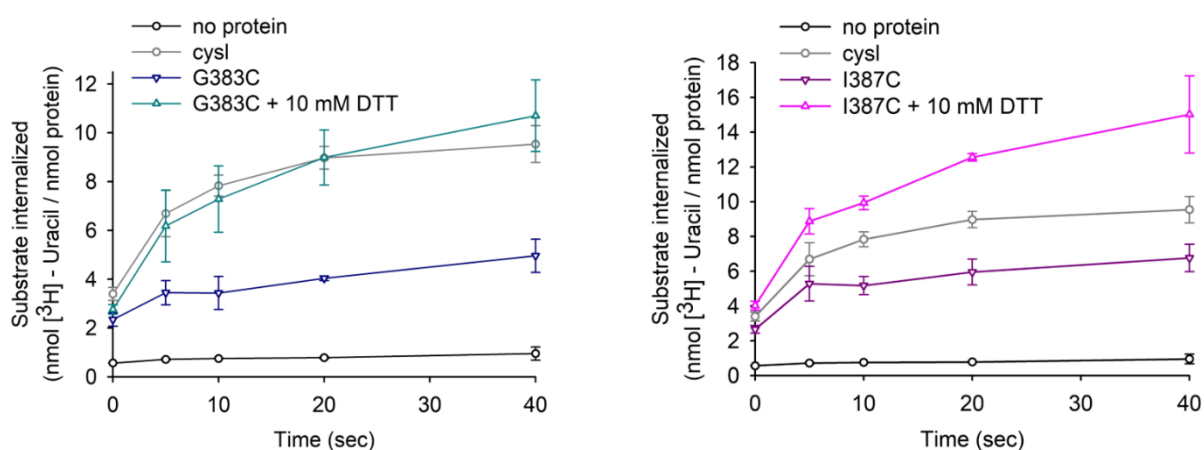
Figure 53: Restricting the dimer interface of 51ThXi might affect transportation. Uptake of 1 µM uracil into liposomes containing no protein (black), or proteoliposomes of 51ThXi(cysl) (grey) or the stapling mutants (51ThXi(G383C), dark blue, and 51ThXi(I387C), dark pink). To test the reversibility of the crosslinking on protein activity, uptakes were also recorded for proteoliposomes that have been pre-incubated with 10 mM DTT (shown in teal and pink, respectively). Uptakes were measured in triplicates, errors shown represent standard deviations.

For both mutants, protein of high purity was obtained that reconstituted with high efficiency, indicating good protein quality (**Figure 50B, C**). In both cases, incubation with copper phenanthroline efficiently induced dimer formation (**Figure 50D**). In *in vitro* uracil uptake assays, a reduction of activity was observed that was restored to wild type levels (or higher) in presence of reducing agent (**Figure 53C**). An elevator mechanism which is characterized by a movement of the substrate binding domain moves against the scaffold domain should not be affected by the constriction of the dimer interface as posed by the crosslink that was introduced here. The reduction of transport observed here

might therefore result from a distortion of the dimer interface in a way which precludes an efficient movement of the substrate binding domain against the scaffold domain. Consequently, the results presented here do not necessarily negate an elevator-like movement as it is suggested for SLC23 transporters. In the two alternative double mutants G356C/G383C and G356C/I387C the cysteines are positioned, according to our model, at more suitable distances leading to less distortion of the dimer interface. For both mutants, efficiencies of reconstitution to soy PC were extraordinarily low in two independent trials, resulting in low uptake signal that could not be evaluated (not shown). It should be noted that also the I374C mutant of UraA (corresponding to I387C of 51ThXi) showed crosslinking in whole cell vesicles, although the C_β-C_β distance according to the dimeric UraA structure is 17.2 Å (**Figure S29**). This suggests a certain flexibility of the interface of UraA, and 51ThXi, in general. The here presented results suggests that a constriction of the dimer interface still allows for transport, but less efficient, which might also indicate a dynamic behavior of the scaffold domain during the transport cycle.

4.4.3. Cooperativity in the 51ThXi dimer

The relevance of dimerization for transport by UraA was already investigated, using uracil uptake in a whole-cell context as readout for protein activity (Yu *et al.*, 2017). Monomerized versions of UraA showed a reduced transport activity compared to dimers. On the other hand, artificial dimers composed of active protein and inactive mutants showed a similar activity as the wild type dimer. In conclusion, dimerization appeared to be required for transport, but only one of two monomers had to be active. However, the reported study has several shortcomings that prevent a clear interpretation of the data. As already discussed above, the degree of UraA monomerization was judged based on the observation of two peaks in SEC, which not necessarily reflects the oligomerization state of the protein in the membrane environment. Additionally, to determine precise protein activities on the basis of uracil uptake in whole cells is difficult, as described thoroughly in this thesis. Herein, similar experiments were therefore performed using artificial dimers of 51ThXi in combination with a proteoliposome-based uptake assay to assess the relevance of 51ThXi dimerization for transport.

A concatemer of 51ThXi, in which two protomers were linked with a 9 amino acid long stretch of glycine and serine residues (51ThXi_(GSS)$_3$_51ThXi) could be expressed as a GFP fusion protein with high quality and showed uptake in a cell-based assay similar to the one of wild type 51ThXi (**Figure 54**). Next, different fusions of wild type 51ThXi and 51ThXi mutants were generated and tested for uracil uptake, in particular using the E256A mutant (similar to the experiments of Yu *et al.* using the UraA(E241A) mutant) and a triple alanine mutant missing all the charged substrate binding site residues (E256A/H260A/E303A).
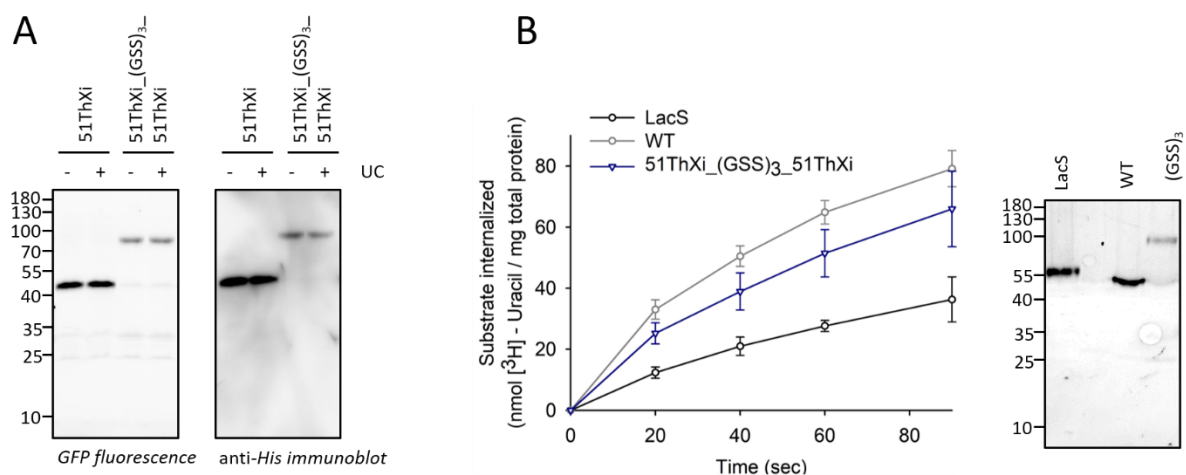
Figure 54: Expression and activity of a 51ThXi concatemer. (A) Wild type 51ThXi and a concatemer, in which two protomers were linked with a nine amino acid glycine-serine stretch (51ThXi_(GSS)$_3$_51ThXi), were expressed as GFP fusion proteins in *E. coli* MC1061. Whole cell lysates were incubated with 1% (w/v) DDM and protein amounts before (-) and after (+) ultracentrifugation (UC) were compared on 12% SDS-PAGE. Proteins were visualized by *in gel* GFP fluorescence and *anti*-His Western blotting. (B) Uracil uptake in *E. coli* BW25113(*ΔuraA*) overexpressing GFP fusions of LacS (black), 51ThXi (grey) or the 51ThXi concatemer (blue). Cells were incubated at 25°C with 1 µM partially radiolabeled uracil. Uptake experiments were performed three times, errors are from standard deviations. Expression of proteins was monitored by *in gel* GFP fluorescence on whole cell lysates.

As discussed in detail in section 4.3, the effect of replacing the TM8 glutamate with alanine on substrate translocation is not entirely clear. For this reason, next to E256A the triple mutant E256A/H260A/E303A of 51ThXi was used in the experiments described here, which lacks all three of the residues mostly involved in substrate coordination. However, as it was not possible to purify this mutant due to pronounced instability (not shown), no DSF measurements could be performed to explicitly demonstrate its incapability of uracil binding.
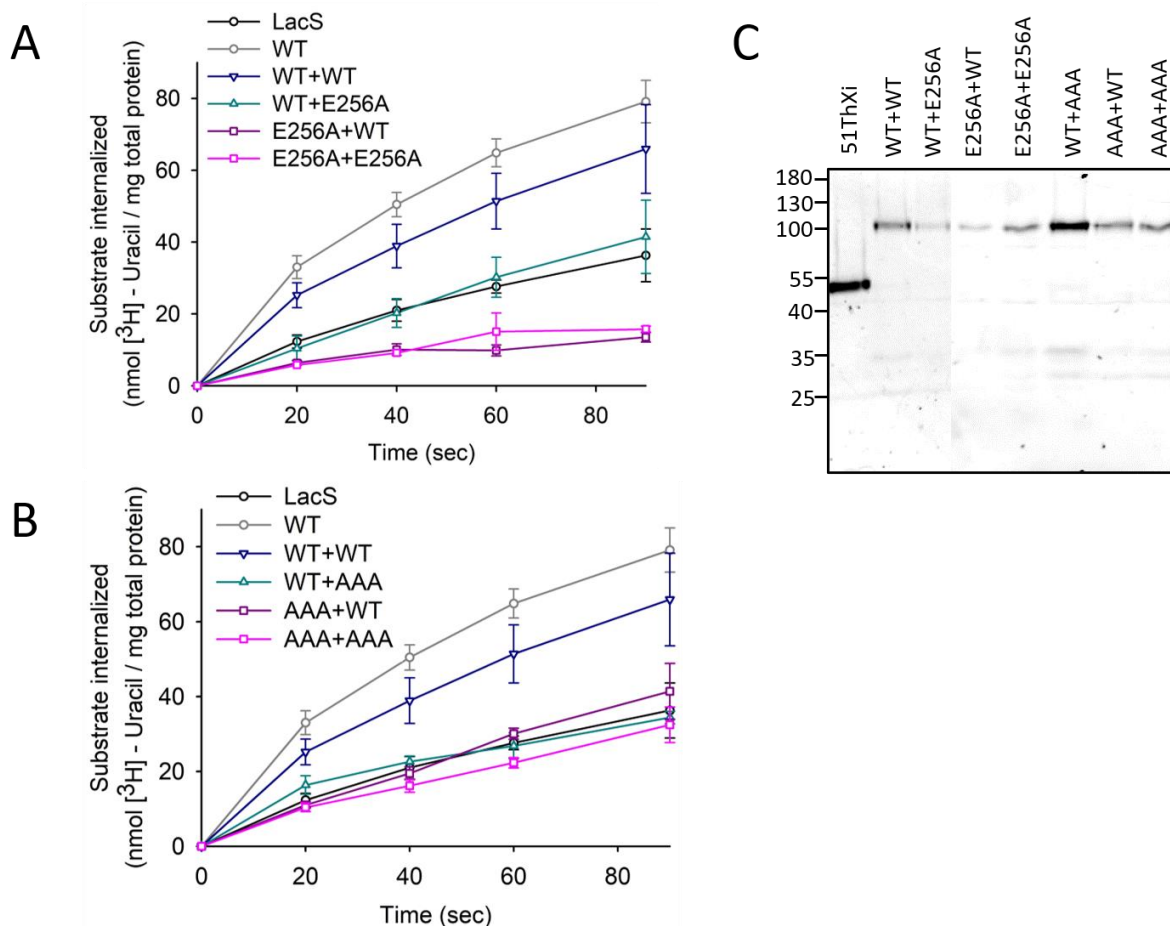
Figure 55: Uptake of different constitutive dimers of 51ThXi. (A) Uracil uptake in *E. coli* BW25113(*ΔuraA*) expressing GFP fusions of different concatemers of 51ThXi, combining wild type 51ThXi with the E256A mutant. Uptake of a concatemer containing two times wild type 51ThXi (WT+WT, blue) was compared to constitutive dimers carrying the mutation in the second protomer (WT+E256A, teal), in the first protomer (E256A+WT, dark pink) or in both protomers (E256A+E256A, pink). Cells were incubated at 25°C in buffer supplemented with 1 μM uracil. Uptakes were measured three times, errors are standard deviations. (B) Same as in (A), but using the E256A/H260A/E303A ("AAA") mutant of 51ThXi instead of E256A. (C) Expression of samples used for uptakes shown in (A) and (B) was monitored by *in gel* GFP fluorescence of whole cell lysates subjected to 12% SDS-PAGE.

All concatemers of 51ThXi that carried either one of the mutations in one of two protomers showed uracil uptake similar, or lower, than the LacS negative control or the dimer composed of two mutated protomers (**Figure 55**). The finding that two mutants (E256A+WT and E256A+E256A) show lower uptake than the LacS negative control probably reflects the intrinsic difficulties of the whole cell-based uracil uptake assays caused by the high background. Also the expression levels of the proteins are variable, which further complicates the interpretation of the uptake experiments. Nevertheless, the uptakes provide a first hint to a negative dominant effect of an inactive mutant over the active protein, as was already described for the SLC23 transporter UapA (Alguel *et al.*, 2016). To further substantiate this finding, cooperativity within the 51ThXi dimer was also explored on basis of a proteoliposome-based assay, using mixes of wild type 51ThXi and the 51ThXi(E256A) mutant in different ratios. In contrast to the remaining (low) uracil internalization that was observed in whole cells overexpressing

the 51ThXi(E256A) mutant compared to the negative control (**Figure 33B**), substrate internalization in proteoliposomes of reconstituted 51ThXi(E256A) was extremely reduced (**Figure 56A**).
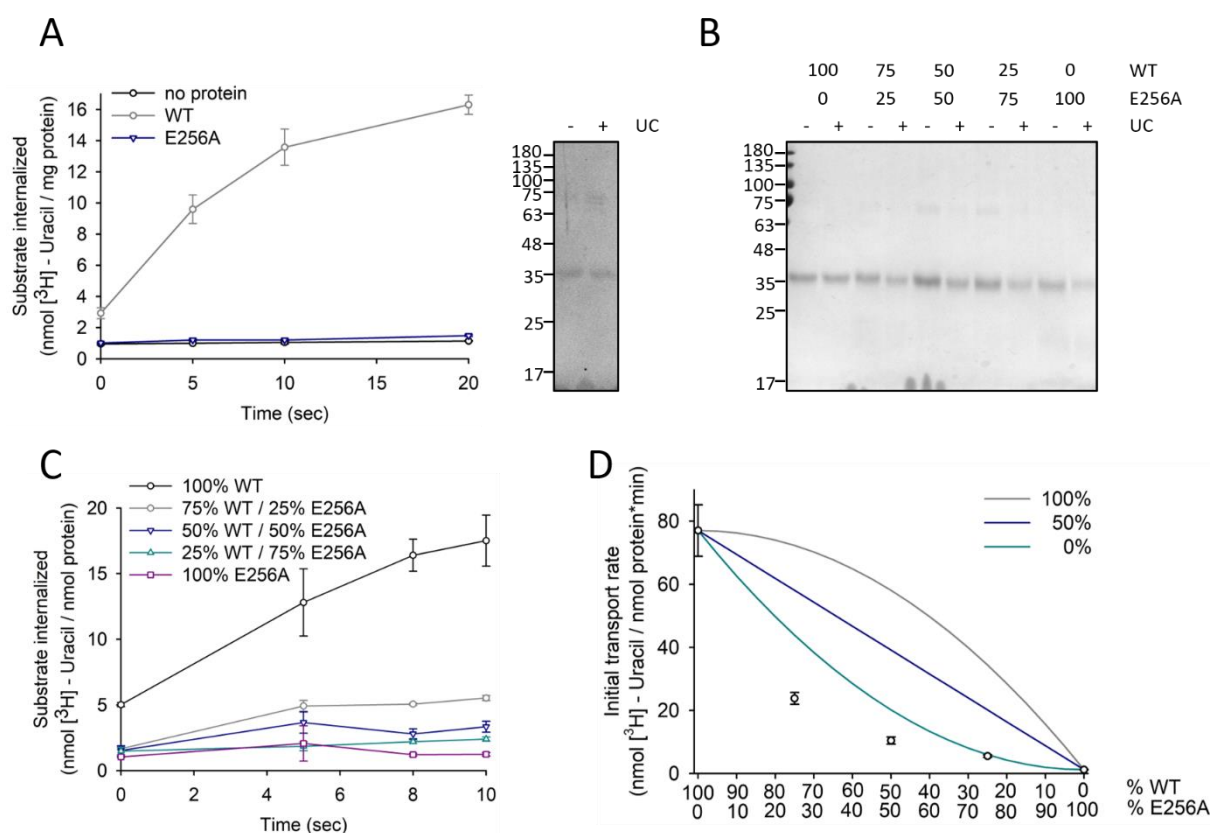
Figure 56: Cooperativity in the 51ThXi dimer. (A) Uracil uptake in proteoliposomes of 51ThXi wild type (grey) and the E256A mutant (blue) compared to liposomes containing no protein (black). Proteoliposomes were incubated at 10°C in presence of 1 µM uracil. Uptakes were performed in triplicates, errors represent standard deviations. The efficiency of E256A reconstitution was assessed by incubating the proteoliposomes with 1% (w/v) DDM and comparing the protein amount before (-) and after (+) and ultracentrifugation (UC) step. Samples were subjected to 12% SDS-PAGE and proteins were visualized by Coomassie staining. (B) Wild type 51ThXi and the E256A mutant were mixed in different ratios and reconstituted 1:50 protein-to-lipid to soy PC. Reconstitution efficiencies were determined for all samples as described in (A). (C) Uracil uptake in proteoliposomes of wild type 51ThXi and 51ThXi(E256A) mixed in different ratios. Proteoliposomes were incubated at 10°C in buffer containing 1 µM partially radiolabeled uracil. Uptakes were recorded three times, errors are standard variations. (D) Initial transport rates of uracil uptake of proteoliposomes of 51ThXi wild type and the E256A mutant mixed in different ratios obtained from uptakes shown in (C) by linear fitting using SigmaPlot 10. Errors shown are derived from error propagation of standard variations. Grey, blue and teal lines indicate the anticipated curves assuming an activity of heterodimers corresponding to 100%, 50% or 0% of the one of wild type homodimers.

The difference of apparent activity of the E256A mutant in whole cells compared to in proteoliposomes can have several reasons. First, the activity observed in the whole cell-based assay might not be caused by substrate transport by the mutant at all, but from a different background of uracil uptake in cells expressing the 51ThXi mutant compared to the LacS negative control. However, also 51ThXi mutants were observed for which uptake was very similar to uptake in cells expressing LacS (e.g., **Figure 44A**), meaning the difference in uptake between cells expressing LacS and 51ThXi mutants presumable is not

inherently caused by 51ThXi overexpression. Another reason might be the different lipidic environment of the cellular membrane and the proteoliposomes; as outlined in the discussion section, lipids might play an important role in promoting an efficient movement of the substrate binding domain against the scaffold domain in elevator proteins. In this, also the membrane curvature might be relevant. Lastly, the action of uracil phosphoribosyltransferase in the cell enhances substrate internalization by maintaining a steep substrate gradient, meaning a residual transport activity probably gives rise to a higher signal than in proteoliposomes.

Although the E256A mutant was shown to have a low stability, it reconstituted well, both on its own and mixed in different ratios with wild type 51ThXi (**Figure 56B**). To account for the lower stability of the mutant, all steps of reconstitution were performed at 4°C, whereas for the reconstitution of other proteins some steps were carried out at room temperature. The total amount of reconstituted protein was similar for all the protein mixes, i.e., different reconstitution efficiencies for the two protein variants that might bias the uptake data can practically be excluded. Next, uracil uptake in proteoliposomes was recorded to determine initial transport rates (**Figure 56C, D**). Notably, the activities that were observed were lower than what would be expected from the individual activity of the wild type protein. For example, in a sample in which 50% of wild type protein are mixed with 50% of inactive mutant, 50% of activity would be recorded in case the activity of the wild type was not influenced by the activity of the inactive mutant. In conclusion, a similar negative dominant effect of the inactive mutant over the active mutant was observed as already seen in the whole cell uptake assay using 51ThXi concatemers.

In order to gain more evidence for the negative cooperativity observed for 51ThXi, it was also tried to perform uptake experiments in proteoliposomes using the concatemer. In contrast to the mixing experiment described above, using a concatemer implies that only heterodimers are present. The concatemer consisting of two wild type protomers could be purified and was well incorporated in the proteoliposomes as judged by the standard reconstitution efficiency test (**Figure 57A, B**). Interestingly, no activity was observed in the proteoliposome-based assay (**Figure 57C**). Thermal melting experiments on the purified concatemer suggest that the protein was well-folded and that uracil binding was uncompromised (**Figure 57D**).
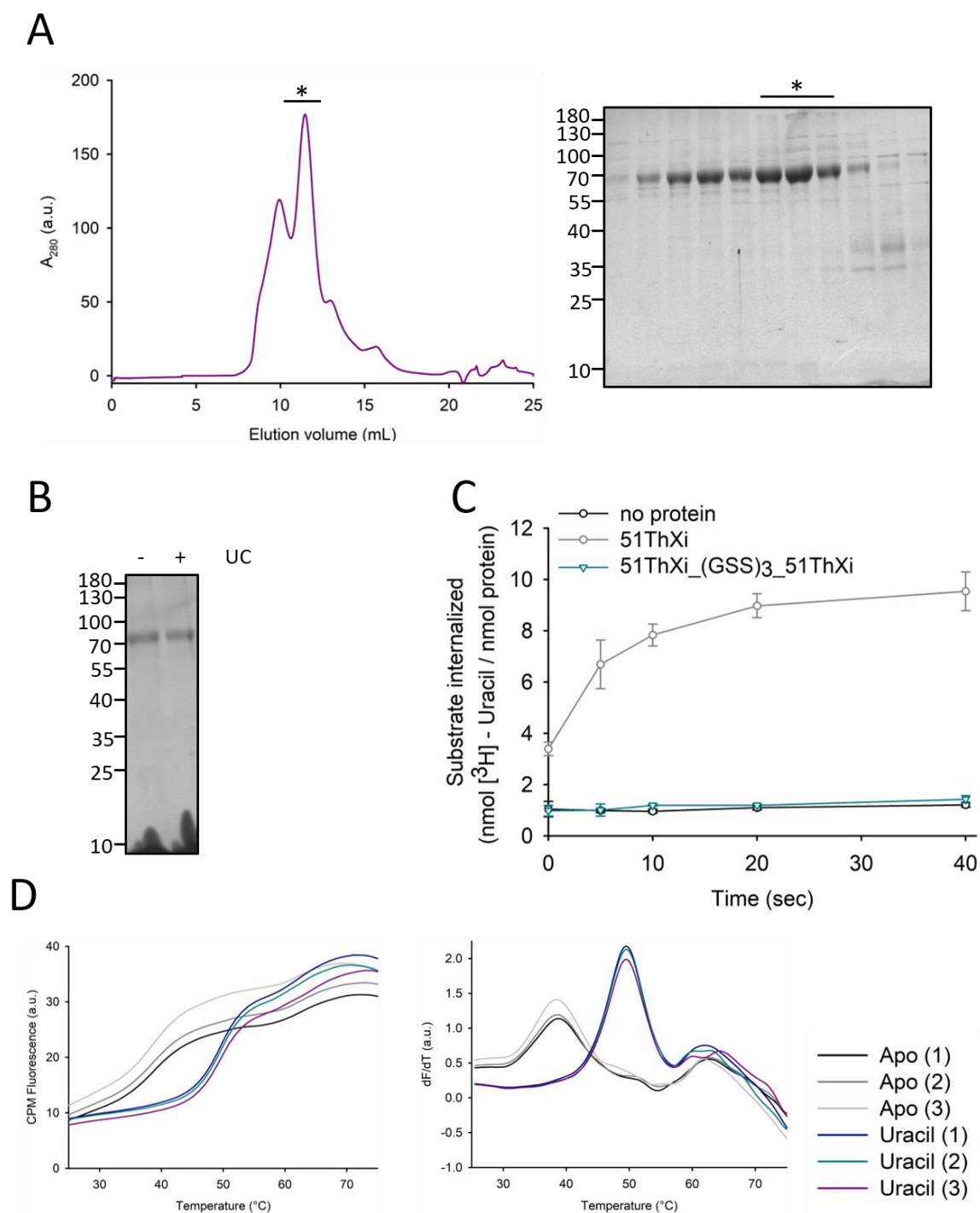
Figure 57: Purification, reconstitution and uptake of the concatemer of 51ThXi. (A) Size-exclusion chromatogram of 51ThXi_(GSS)$_3$_51ThXi on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were controlled by SDS-PAGE and Coomassie staining. Fractions used for reconstitution are indicated by an asterisk (*). (B) Efficiency of reconstitution of the 51ThXi concatemer soy PC. Proteoliposomes were incubated with 1% (w/v) DDM and protein amounts before (-) and after (+) ultracentrifugation (UC) were compared by SDS-PAGE and Coomassie staining. (C) Uracil uptake in proteoliposomes of the 51ThXi concatemer (teal) compared to wild type 51ThXi (grey) and liposomes containing no protein (black). (D) DSF measurements on purified 51ThXi_(GSS)$_3$_51ThXi without additional substrate (black and grey lines) and in presence of 1 mM uracil (colored lines). 0.5 mg/mL protein (9 µM) were incubated from 25°C to 95°C in steps of 1°C with 20 seconds incubation per step and fluorescence of CPM dye was recorded. The melting temperature is found at the maximum of the first derivative of the CPM fluorescence (dF/dT).

The inactivity of the concatemer in proteoliposomes is in stark contrast to the results of uracil uptake in whole cells, where the protein exhibited wild type-like activity (**Figure 54**). Although the protein appeared to be well-incorporated into the proteoliposomes, it might not be possible to exclude that only one protomer is inserted into the lipid bilayer. Resolubilization from the membrane might not lead to complete aggregation and protein would consequently not be removed in the ultracentrifugation step. The not-incorporated protomer would probably inhibit free movement of the incorporated protomer, resulting in a net inactivity. Another possibility might be again the lipid environment; in whole cells, the constitutive dimer apparently is unrestricted in efficient substrate transport, but in proteoliposomes a necessary dynamic of the protomers, for example some degree of movement of the protomers against each other, might be missing. The DSF measurements support the finding that the concatemer in general is well-folded, as the melting temperature curves closely resemble the ones obtained for wild type 51ThXi. As the wild type protein was shown to be monomeric in detergent by SEC-MALS, as mentioned in section 4.4.2, this could be taken as a hint that dimerization does not additionally stabilize the protein. This would however assume that the dimer interfaces of the concatemer and wild type 51ThXi in the membrane resemble each other.

4.5. Relevance of the dimeric state of SLC26 transporters

The SLC23 family of nucleobase and ascorbate transporters and the multifunctional SLC26 family of anion transporters share the same architecture of the membrane-embedded protein part. SLC26 proteins, however, contain an additional C-terminal, cytosolic domain, the so-called STAS domain. Multiple functions have been suggested for this additional domain, as outlined in the introduction. To date, the precise role of the STAS domain for the transport cycle of SLC26 anion transporters has not been fully elucidated. For several proteins it has been shown that mutants lacking the STAS domain exhibit lower rates of substrate transport, e.g., the plant SLC26 transporter Sultr1;2 from *Arabidopsis thaliana* (Shibagaki and Grossman, 2004) or the structurally resolved SLC26Dg from the prokaryotic organism *Deinococcus geothermalis* (Geertsma *et al.*, 2015). Similar to the elevator proteins of the SLC4 and SLC23 families, the substrate translocation pathway, though, is confined to the membrane part of SLC26 transporters. The results presented here on the role of the STAS domain in substrate translocation of SLC26 transporters are partially based on unpublished results summarized in the PhD thesis of Dr. Yung-Ning Chang ("Structural and functional interfaces of SLC26 transporters", Goethe University Frankfurt am Main, 2018). This includes in particular the identification of another prokaryotic SLC26 homolog from *Sulfitobacter indolifex*, termed SLC26Si, as a proton-coupled transporter of oxalate. Also in SLC26Si, removal of the STAS domain decreased the rate of substrate transport, similar as observed for SLC26Dg, without changing the affinity for the substrate. In chimeric constructs of SLC26Dg and SLC26Si, fusing the STAS domain of one homolog to the transmembrane domain of the other and *vice versa*, transport capacity was not restored, although both chimeras were well expressed and reconstituted. This suggests specific interaction of STAS and transmembrane domain. The dimer interface of SLC26 transporters, as was shown on a dimer model of SLC26Dg based on EPR constraints but later on also in cryo-EM structures of murine SLC26a9 and human SLC26A9, is built from only a small part of the transmembrane domain, mostly involving TM14 (Chang *et al.*, 2019; Walter *et al.*, 2019; Chi *et al.*, 2020). STAS-less SLC26Dg, however, still dimerizes in lipidic environment. EPR measurements revealed small changes in distances between residues of TM14, suggesting small modulations of the dimer interface in response to STAS domain removal. How this altered dimer interface contributes to a lower rate of substrate transport was so far not investigated.

4.5.1. Overall architecture of the SLC26Dg dimer

As already mentioned above, recent cryo-EM structures of eukaryotic SLC26 transporters revealed the overall architecture of dimers of SLC26 family members. Before, only the crystal structure of monomeric SLC26Dg had been available, as well as a model of the dimer of the transmembrane domain of SLC26Dg, based on constraints from EPR (Chang *et al.*, 2019; Walter *et al.*, 2019; Chi *et al.*, 2020).

In the full-length dimer structures of SLC26 proteins, the STAS domain of one protomer is placed close to the transmembrane domain of the other protomer (**Figure 58A**). The dimer interface is built from interactions of both transmembrane and STAS domains. As was shown in an extensive crosslinking study on SLC26Dg, SLC26Si and was well as rat prestin (SLC26a5), the transmembrane domain mostly contributes over interaction of TM14 of the two protomers to the dimer interface (Chang *et al.*, 2019). Later on, this was confirmed by the two new structures of SLC26a9/SLC26A9 (**Figure 58B**).
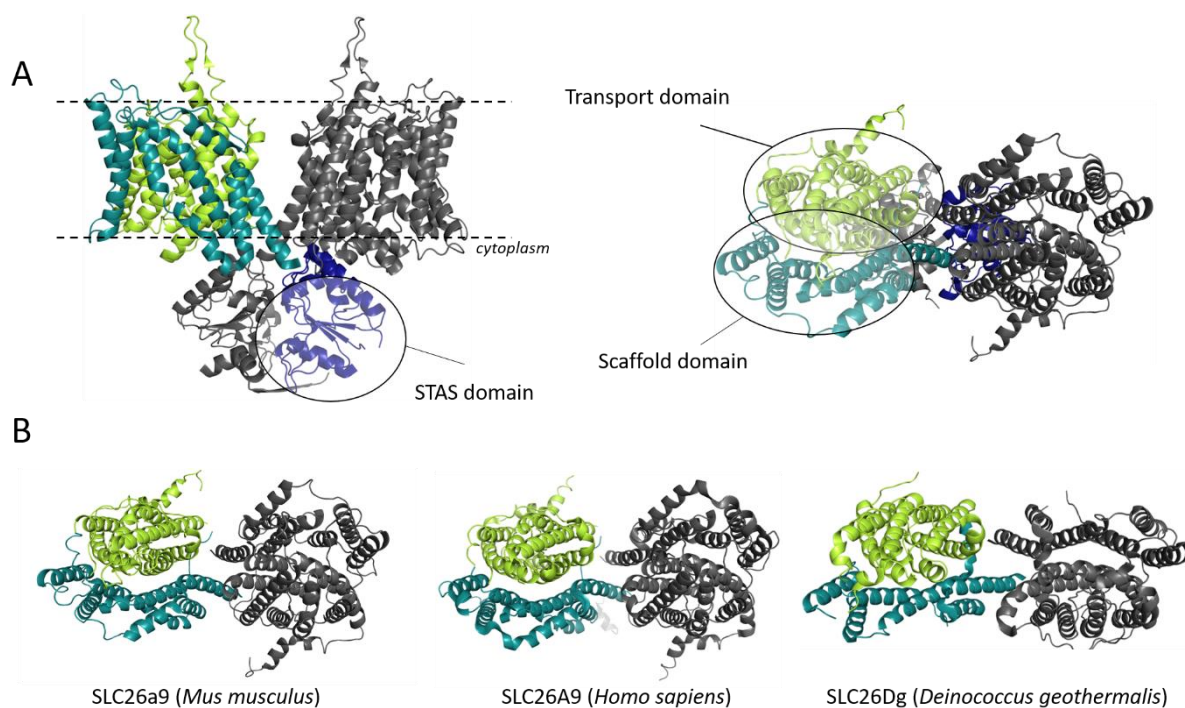
Figure 58: Overall structure of dimers of SLC26 transporters. (A) In the structure of SLC26a9, the STAS domains are placed in vicinity to the transmembrane domain of the opposing protomer. The protein is depicted in the side view (left) with one protomer in grey, the other in green and blue. On the right side, the top view is shown. The transport domain is colored in light green, the scaffold domain in teal, and the STAS domain in dark blue. The approximate barriers of the membrane are indicated by dashed lines. (B) The transmembrane domains of SLC26 transporters contribute with an interaction of TM14 to the overall dimer interface. This is conserved in SLC26a9 (PDB: 6RTC), SLC26A9 (PDB: 7CH1) and SLC26Dg (PDB: 5DA0). For better visualization, STAS domains are omitted here.

The model of the SLC26Dg dimer determined using distance constraints obtained by EPR is confined to the transmembrane part and does not include the STAS domains. Herein, in order to test whether the dimer architecture of SLC26Dg resembles the one of SLC26a9, cysteine crosslinking mutants were generated. For that, the transmembrane part and STAS domain from the monomeric SLC26Dg crystal structure were separately aligned to the SLC26a9 structure using PyMOL. It should be noted that the structure of SLC26A9 from *Homo sapiens* was not available at the time these experiments were performed, so structural alignments were exclusively based on using the dimer of SLC26a9 from *Mus musculus*. SLC26Dg shares a sequence identity of 14.81% with SLC26a9 as determined using AlignMe.

The transmembrane domain of SLC26Dg aligns with an RMSD of 3.2 Å to the one of SLC26a9 (over 360 residues), the STAS domains of both proteins also align with an RMSD of 3.2 Å (over 88 residues).

Different double cysteine mutants were generated for SLC26Dg, combining mutations on the STAS domain with mutations on the transmembrane domain that are, according to the dimer model, in suitable distances for formation of disulfide bridges (**Figure 59**).

Pos 1	Location	Pos 2	Location	Distance
K121	Scaffold	S419	STAS	2 Å
K390	Scaffold	V481	STAS	4.7 Å
Q247	Transport	H421	STAS	6.7 Å
R412	STAS	A475	STAS	5.8 Å
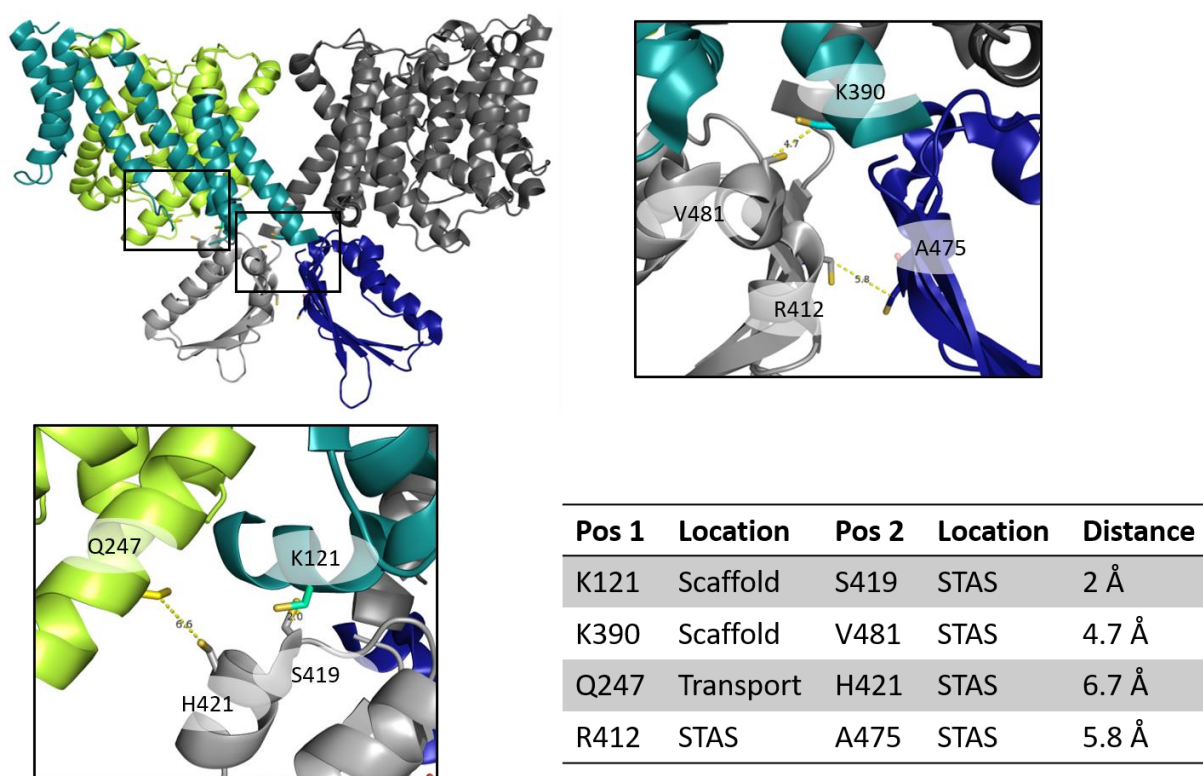
Figure 59: Cysteine mutants of SLC26Dg to determine the orientation of the STAS domain in relation to the transmembrane domain. The dimer model of SLC26Dg, based on structural alignment on the dimeric structure of SLC26a9 (PDB: 6RTC), was used to generate double cysteine mutants supposed to form crosslinks between the different domains of the protein, as indicated in the table. One protomer is shown in grey, the other in color. The scaffold domain is indicated in light green, the transport domain in teal and the STAS domain in dark blue. Distances were measured using the built-in PyMOL tool.

The mutants were expressed as GFP fusion proteins in *E. coli* and crosslinking induced by copper phenanthroline as well as cysteine accessibility were analyzed (**Figure 60**). All the mutants carrying one cysteine on either scaffold or transport domain and one on the STAS domain showed dimer formation on gel, meaning that a crosslink was formed between the STAS domain of one protomer and the transmembrane domain of the other protomer. This suggests that indeed the overall architecture of SLC26Dg resembles the one of SLC26a9. Interestingly, for the double cysteine mutant R412C/A475C, designed to form a disulfide bridge between the two STAS domains, no dimer formation was observed (not shown). In contrast, the SLC26Dg(A475C) single mutant showed dimerization. This indicates that the precise positioning of the STAS domains might differ to a certain extent from the model.
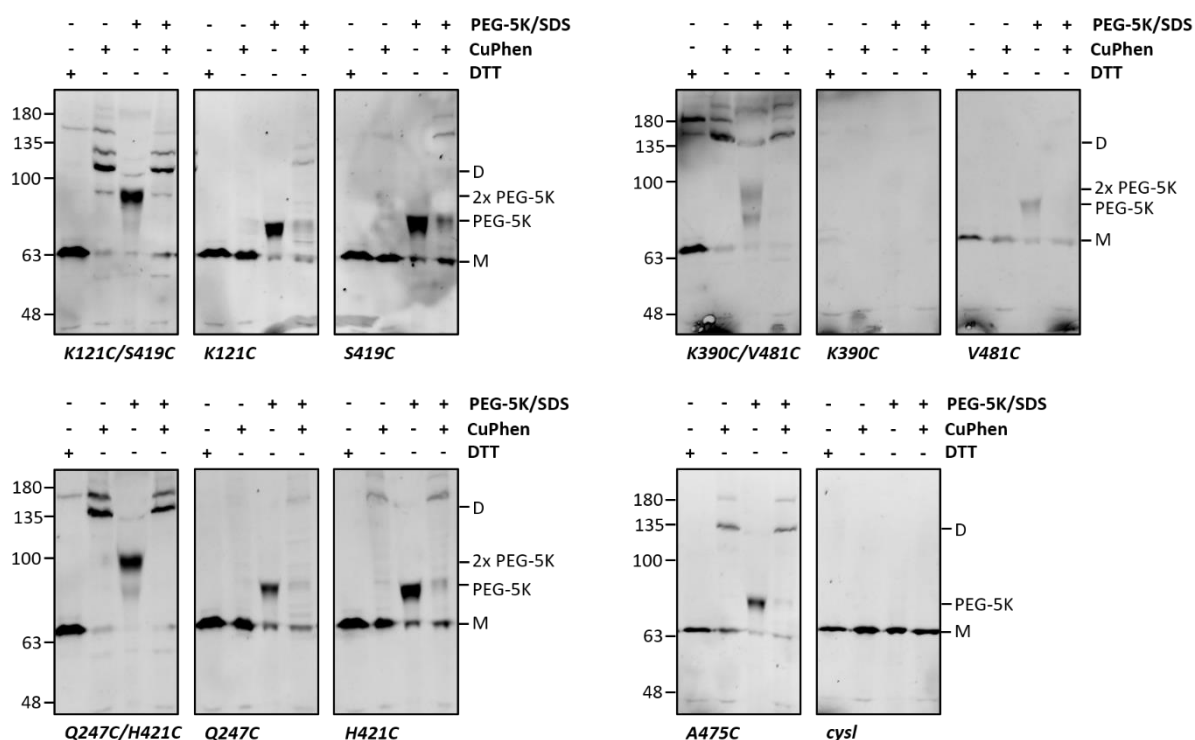
Figure 60: Oxidative crosslinking of cysteine mutants of SLC26Dg. Single- and double cysteine mutants of SLC26Dg were expressed as GFP fusion proteins. Cells were lysed in presence of 10 mM DTT that was subsequently removed using Biospin 6 columns. Samples were incubated with 1 mM CuPhen and/or 1 mM PEG-5K + 1 % (w/v) SDS, as indicated by minus (-) and plus (+). Samples were separated on 8% SDS-PAGE and proteins were visualized by *in gel* GFP fluorescence. Monomer and dimer are indicated by M and D, respectively, as well as one- and two-fold PEG-labeled protein. Expression and crosslinking analysis were performed three times, representative results are shown.

Dimer formation was also explored in proteoliposomes, using PELDOR EPR spectroscopy on reconstituted SLC26Dg mutants. Sample preparation (purification, spin labeling, reconstitution) were performed by myself, EPR spectra were recorded and data was evaluated by Dr. Benesh Joseph. For PELDOR measurements, three single cysteine mutants of SLC26Dg were constructed. Since SLC26Dg is a monomer in detergent but a dimer on the membrane, heterodimers in proteoliposomes can be obtained by mixing the mutants before reconstitution. This was preferred over double cysteine mutants to reduce the number of spin labels and thus the background. Suitable spin-label distances range from 1.5 to 10 nm (Jeschke, 2018). One residue on each transport and scaffold domain (V129 and L248, respectively) were paired with the same position on the STAS domain (L459) (**Figure 61A**). The mutations do not affect the transport activity of SLC26Dg (**Figure S30**). The spin-label efficiency (62% - 91%) was determined on the basis of CW spectra (recorded and evaluated by Dr. Benesh Joseph) before mixing and reconstituting 1:20 protein-to-lipid (weight-to-weight) to soy PC. For both heterodimer samples, smaller distances around 2 nm were detected that indicate spin label interaction of V129R1-L459R1 and L248R1-L459R1 in line with the SLC26Dg dimer model (**Figure 61B,C**).
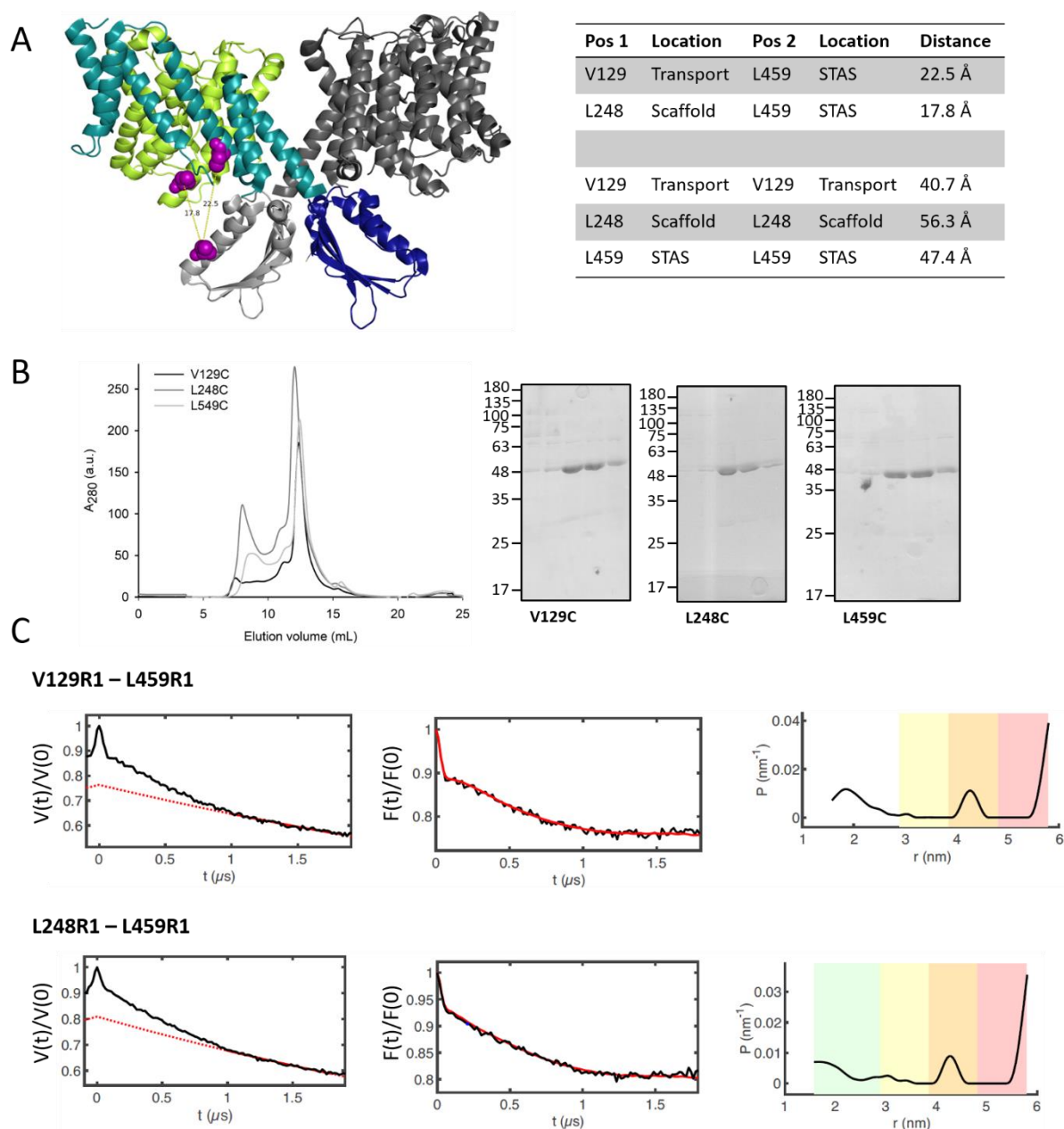
Figure 61: PELDOR measurement on SLC26Dg to confirm the dimer model. (A) Three cysteine mutants of SLC26Dg were generated for spin-labeling and PELDOR measurements. Cysteines are indicated in purple spheres, one protomer of SLC26Dg is shown in grey, the other in color. For the colored protomer, the transport domain is in light green, the scaffold domain in teal and the STAS domain in dark blue. Based on the model, the C_{β}-C_{β} distance of V129C (located on the transport domain) and L459C (located on the STAS domain) was determined as 22.5 Å using PyMOL, the one for L248C (located on the scaffold domain) and L459C is 17.8 Å. Additional distances are indicated in the table. (B) Size-exclusion chromatogram of SLC26Dg_V129C, L248C and L459C on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 5 mM ß-mercaptoethanol, 0.2% (w/v) DM. The proteins were overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Purity of fractions used for reconstitution to soy PC and PELDOR measurements was monitored by SDS-PAGE and Coomassie staining depicted on the right. (C) PELDOR measurements on spin-labeled SLC26Dg double mutants V129R1-L459R1 and L248R1-L459R1. On the left the original time traces are shown, in the middle the data after background subtraction, and the right panel displays the corresponding distance distributions. Data was analyzed and figures were prepared by Dr. Benesh Joseph.

4.5.2. Mutating the linker between transmembrane and STAS domain affects transport activity of SLC26 proteins

The transmembrane domain and the STAS domain of SLC26 transporters are connected by a short linker sequence (**Figure 62A**). In SLC26Dg, this linker is comprised of 5 amino acids (^{391}VSQLS^{395}). It was demonstrated that mutating this linker to GSGGS (mutations underlined), thus presumably making it more flexible, drastically reduces the transport activity of SLC26Dg (Chang and Geertsma, unpublished, and **Figure 63E**). The linker is variable in size (usually 5 to 7 amino acids, with longer stretches found in proteins from eukaryotes) and sequence, although there seems to be a preference for a positive charge located in the middle of it (**Figure 62B**). The role for this positive charge is unknown. Herein, the reduced activity of the SLC26Dg linker mutant (designated as SLC26Dg^{GS-L} from now on) was confirmed and extended to a similar glycine-serine linker mutant in SLC26Si (termed SLC26Si^{GS-L}). Since no structure of SLC26Si is available to date, the location of the linker was judged by sequence alignment to SLC26Dg. The putative linker sequence of SLC26Si is ^{436}TSRPH^{440}, which was mutated to GSGGS (mutations underlined) (**Figure 62C**).
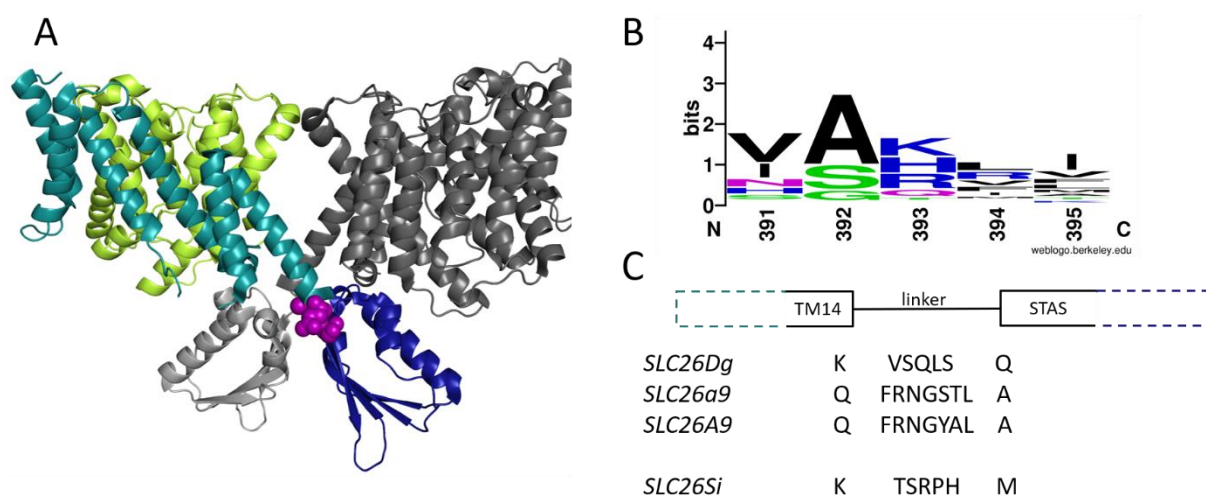
Figure 62: The linker connecting transmembrane and STAS domain in SLC26 transporters. (A) Model of dimeric SLC26Dg with one protomer in grey, the other in color (transport domain in light grey, scaffold domain in teal, STAS domain in dark blue). The linker connecting TM14 of the transmembrane domain and the first ß-sheet of the STAS domain is shown in purple spheres. (B) Sequence conservation of 500 automatically selected SLC26Dg homologs (sequence identity from 42.06% to 86.97%) was analyzed using the ConSurf server. Conservation of residues aligning to the linker sequence of SLC26Dg (residues 391-395) were depicted using WebLogo. (C) Linker composition of structurally resolved SLC26 transporters (SLC26Dg, PDB: 5DA0; SLC26a9, PDB: 6RTC; SLC26A9, PDB: 7CH1) and the putative one of SLC26Si, based on sequence alignment to SLC26Dg using AlignMe.
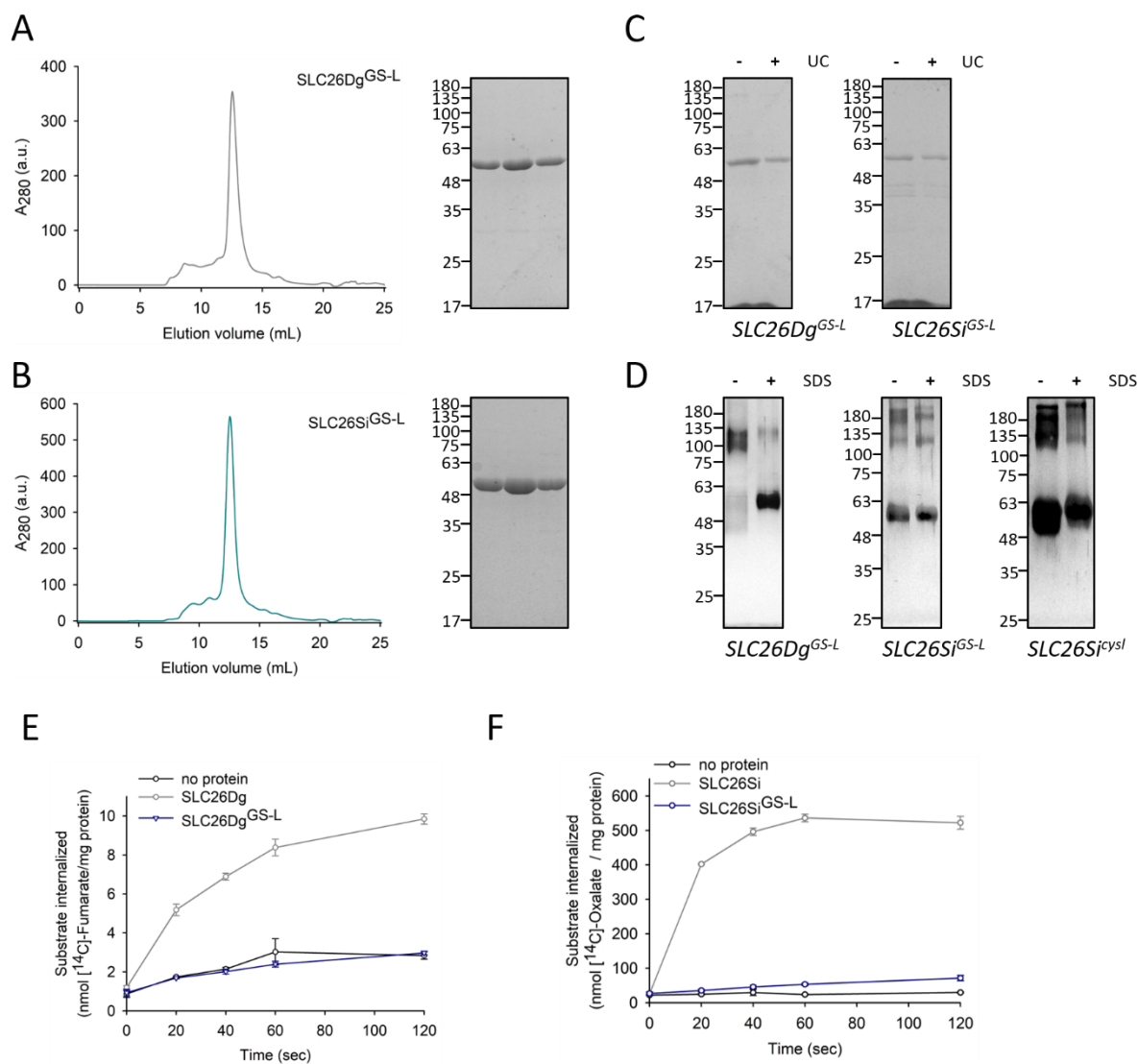
Figure 63: Effect of linker mutations on activity of SLC26Dg and SLC26Si. (A) Size-exclusion chromatogram of SLC26Dg^{GS-L} on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Purity of fractions used for reconstitution was monitored by SDS-PAGE and Coomassie staining. (B) Same as in (A), but for the SLC26Si linker mutant (SLC26Si^{GS-L}). (C) SLC26Dg^{GS-L} and SLC26Si^{GS-L} were reconstituted in a weight-to-weight ratio of 1:50 (protein-to-lipid) to soy PC. The reconstitution efficiency was estimated by resolubilization of proteins from the liposomes using 1% (w/v) DDM and comparing protein amount before (-) and after (+) ultracentrifugation (UC) on SDS-PAGE followed by Coomassie staining. (D) For determination of the oligomeric state by unspecific crosslinking, proteoliposomes were incubated with 0.5% (w/v) glutaraldehyde either without (-) or including a pre-incubation step (+) with 1% (w/v) SDS. Samples were separated on 10% SDS-PAGE and proteins were visualized by silver staining. (E) Uptake of radiolabeled fumarate in proteoliposomes containing no protein (black), wild type SLC26Dg (grey) or SLC26Dg^{GS-L} (blue). Uptake of 50 µM fumarate was monitored at 30°C in presence of a proton gradient and an inward positive membrane potential. (F) Uptake of radiolabeled oxalate in proteoliposomes containing no protein (black), wild type SLC26Si (grey) or SLC26Si^{GS-L} (blue). Uptake of 1 mM oxalate was monitored at 30°C in presence of a proton gradient and an inward positive membrane potential. All uptakes were performed in triplicates, errors shown are from standard deviations.

Both linker mutants of SLC26Dg and SLC26Si eluted as monodisperse peaks on SEC and reconstituted with high efficiency to soy PC (**Figure 63A, B, C**). Dimerization of SLC26Dg^{GS-L} in proteoliposomes was unaffected, as suggested by unspecific crosslinking using glutaraldehyde (**Figure 63D**). SLC26Si^{GS-L}, on

the other hand, did not show distinct dimerization in a similar assay as most of the proteins were not crosslinked. Furthermore, the pre-incubation step with SDS, that effectively disrupts the dimer of SLC26Dg^{GS-L}, did not have a pronounced effect on SLC26Si^{GS-L}. However, similar behavior was observed for proteoliposomes of a SLC26Si control (SLC26Si^{cysl}) and is therefore presumably not a specific effect of the linker mutation in SLC26Si.

For SLC26Dg^{GS-L}, a severe reduction of fumarate uptake into proteoliposomes was observed, in line with what was previously reported (PhD thesis of Dr. Yung-Ning Chang and **Figure 63E**). A similar effect was observed for oxalate uptake by the linker mutant of SLC26Si (**Figure 63F**). Although the wild type linker sequences of SLC26Dg and SLC26Si differ (^{391}VSQLS^{395}for SLC26Dg and ^{436}TSRPH^{440} for SLC26Si), mutating them to a more flexible GS-linker (GSGGS) reduces the activity for both proteins. As mentioned above, the STAS domain is not an absolute requirement for dimerization, but necessary for efficient substrate transportation by the transmembrane domain. The data presented here indicate that the specific positioning of the STAS domain with respect to the transmembrane domain might play a role in substrate translocation. In the linker mutants the connection between transmembrane and STAS domain is (presumably) more flexible and this might negatively influence the transport rate of the SLC26 transporters.

In order to investigate to which extent the position of the STAS domain is influenced in the linker mutants, EPR distances were determined, combining the cysteine mutants used to validate the SLC26Dg dimer model (section 4.5.1) with the linker mutation. The resulting SLC26Dg^{GS-L}(V129C), SLC26Dg^{GS-L}(L248C) and SLC26Dg^{GS-L}(L459C) protein variants were purified, spin-labeled, mixed and reconstituted to soy PC as described above. The obtained distances, surprisingly, were only slightly larger than the one in wild type background (**Figure 64**), suggesting that the position of the STAS domain is not changed drastically. The influence of the linker mutation on the overall dimer interface remains to be investigated. Removing the STAS domain from SLC26Dg (and SLC26Si) inactivates the protein, although dimer formation is still observed. As stated above, the STAS domain contributes to the dimer interface of SLC26 proteins to a large extent. It appears that this dimer interface is necessary for transport activity of the protein, presumably in facilitating an efficient movement of the transport domain against the scaffold domain. Changes to the dimer interface, for example caused by removal of the STAS domain or by a more flexible attachment of the STAS domain to the transmembrane domain, might impair this movement.
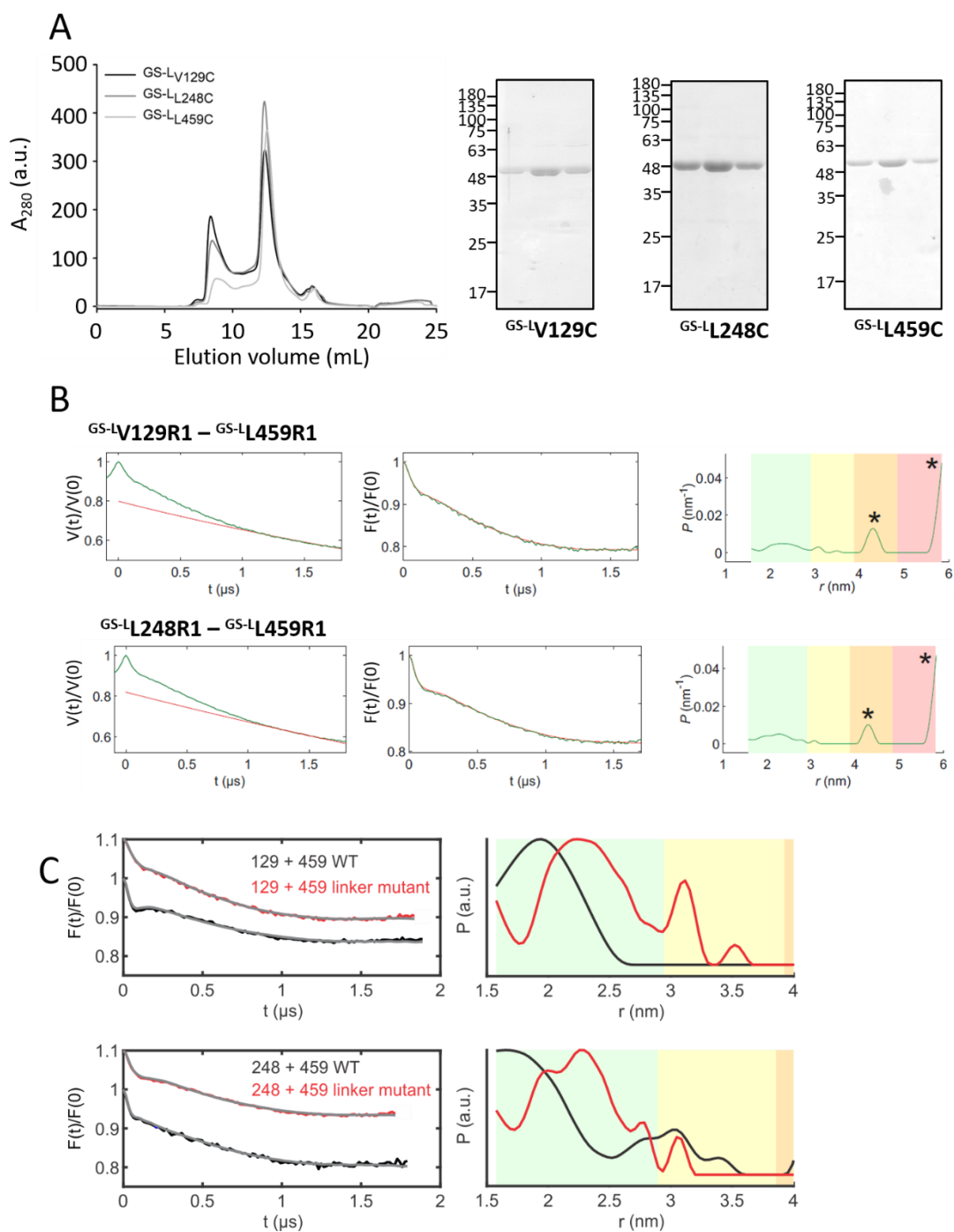
Figure 64: PELDOR measurements on the linker mutant of SLC26Dg. (A) SLC26Dg mutants for PELDOR measurements (V129C, L248C and L459C) were combined with the GS-linker mutant of the linker between transmembrane and STAS domain (391GSGGS395). Size-exclusion chromatogram of SLC26Dg^{GS-L}_V129C, L248C and L459C on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 5 mM ß-mercaptoethanol, 0.2% (w/v) DM. The proteins were overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Purity of fractions used for reconstitution to soy PC and PELDOR measurements was monitored by SDS-PAGE and Coomassie staining depicted on the right. (C) PELDOR measurements on spin-labeled SLC26Dg^{GS-L} double mutants V129R1-L459R1 and L248R1-L459R1. On the left the original time traces are shown, in the middle the data after background subtraction, and the right panel displays the corresponding distance distributions. Data was analyzed and figures were prepared by Dr. Benesh Joseph. (C) Comparison of PELDOR measurements on spin-labeled SLC26Dg double mutants with ("linker mutant", in red) and without ("WT", in black) additional mutation of the linker connecting transmembrane and STAS domain. Data was analyzed and figures were prepared by Dr. Benesh Joseph.

4.5.3. *Cooperativity in SLC26Si*

The functional dependency of the two protomers in the SLC26Dg dimer has been investigated by mixing wild type SLC26Dg with an inactive, inward-locked mutant in different ratios and measuring the uptake activity (Chang *et al.*, 2019). Interestingly, the activity of the wild type protein combined with the immobilized mutant was higher than what would be expected from the activity of wild type homodimers only, suggesting a functional crosstalk between the protomers and a relevance of dimerization. It was suggested that the scaffold domains forming the dimer interface act as an anchor for an efficient movement of the transport domain for substrate translocation. In this context, the immobilized protomer would work as an extended scaffold, making the transport domain movement more efficient and thus leading to higher activity of the wild type protomer. To increase the knowledge on cooperative behavior of SLC26 proteins, a similar experiment was performed using SLC26Si, as will be described in the following.

The immobilized version of SLC26Dg was generated by introducing two cysteines at the interface of transport and scaffold domain that lock the protein in the inward-open conformation upon formation of a disulfide bond. The inward-locked mutant can be identified by a shift on SDS-PAGE, as it exhibits different electrophoretic mobilities in the oxidized and reduced state (Chang *et al.*, 2019). This was used to identify a similar inward-locked mutant of SLC26Si. The double cysteine mutant corresponding to the inward-locked mutant of SLC26Dg (I45C/A142C) is L49C/S153C in SLC26Si (**Figure 65A**). For this mutant, a similar difference in electrophoretic mobility was observed under oxidizing and reducing conditions as was described for SLC26Dg (**Figure 65B,C**). Furthermore, crosslinking efficiently reduced the activity of the protein, which was restored in the presence of DTT (**Figure 65D**). This indicates that the crosslink indeed locks the protein in a transport-incompetent state.
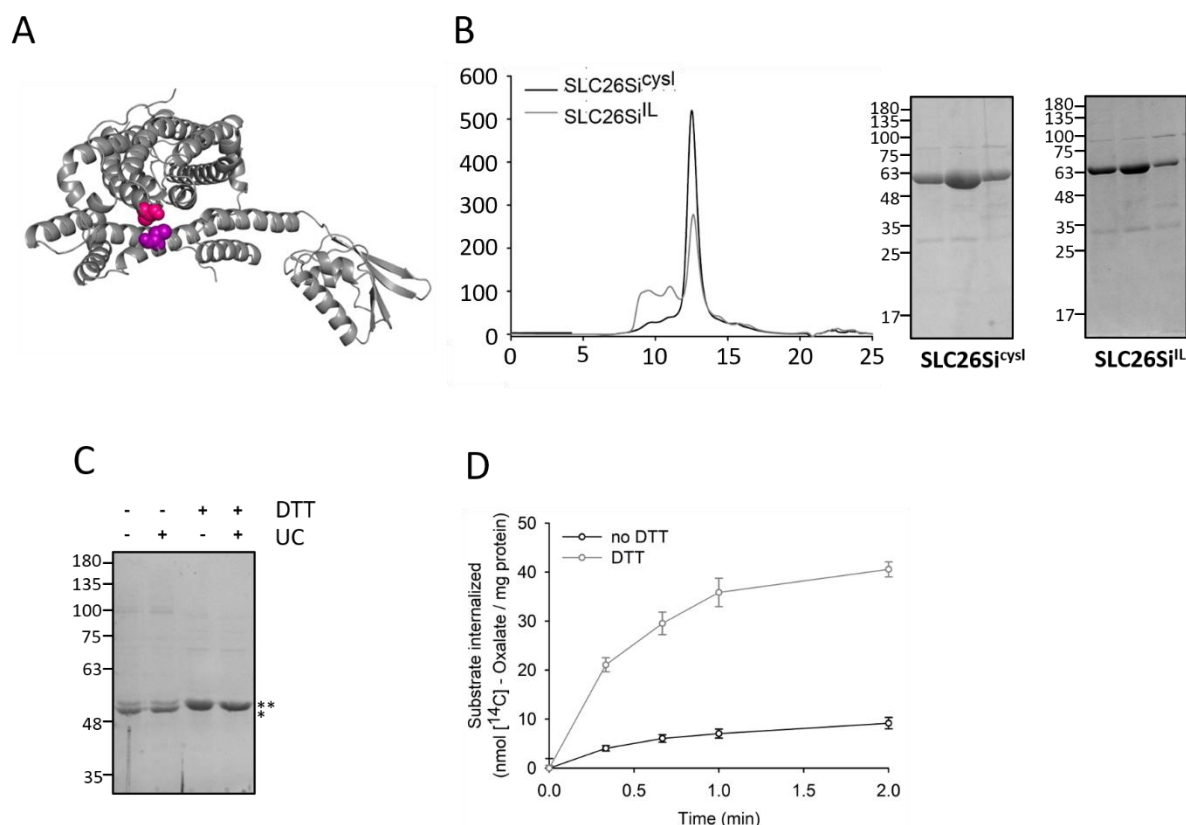
Figure 65: Purification and reconstitution of SLC26Si^IL. (A) Inward-locked mutant of SLC26Si. A model of SLC26Si was generated using Phyre2. Positions of cysteines introduced at the interface of transport and scaffold domain to generate an immobilized version (L49C/S153C) are indicated by colored spheres. (B) Size-exclusion chromatogram of cysteine-free SLC26Si (SLC26Si^cysl, black line) and the inward-locked mutant of SLC26Si (SLC26Si^IL, grey line) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 5 mM ß-mercaptoethanol, 0.2% (w/v) DM. Purity of fractions used for mixing and reconstitution to soy PC was monitored by SDS-PAGE and Coomassie staining. (C) SLC26Si^IL was incubated with CuPhen for oxidative crosslinking and reconstituted to soy PC. The crosslinked fraction of SLC26Si^IL is indicated by one asterisk (*), the non-crosslinked fraction by two asterisks (**). Reversibility of crosslinking was monitored by incubation with 10 mM DDT. The reconstitution efficiency was estimated by re-solubilizing the protein from the proteoliposomes using 1% (w/v) DDM and comparing protein amounts before (-) and after (+) ultracentrifugation (UC). Samples were applied to 10% SDS-PAGE and proteins were visualized by Coomassie staining. (D) Uptake of oxalate in proteoliposomes of SLC26Si^IL. Protein was crosslinked before reconstitution. Uptake of 1 mM oxalate was monitored at 30°C in presence of a pH gradient and an inward positive membrane potential. Uptake of crosslinked protein ("no DTT", black line) was compared to proteoliposomes pre-incubated with 10 mM DTT ("DTT", grey line). Uptakes were performed in triplicates, errors shown are from standard deviations

To investigate the interprotomer cooperativity of SLC26Si, cysteine-free SLC26Si (SLC26Si^cysl) as active protomer and SLC26Si^IL as immobilized, inactive protomer were purified, crosslinked, mixed in different ratios and reconstituted to soy PC (**Figure 66A**). For all the mixes, oxalate uptake was measured in the linear phase of transport (**Figure 66B**). This was used to determine initial transport rates by linear fitting using SigmaPlot 10 (**Figure 66C**). In contrast to what was observed for SLC26Dg, in which the two protomers appeared to exhibit a positive cooperativity, for SLC26Si the activity that was obtained from mixing active with inactive protomer was lower than expected from independent transport by the single subunits. This indicates that SLC26Si rather displays negative cooperativity,

similar to what was observed for the SLC23 transporter 51ThXi (section 4.4.3). In this case, immobilization of one protomer might impair an efficient movement of the transport domain of the opposing protomer. For SLC26Si, this means that both protomers need to be mobile for efficient substrate transport.
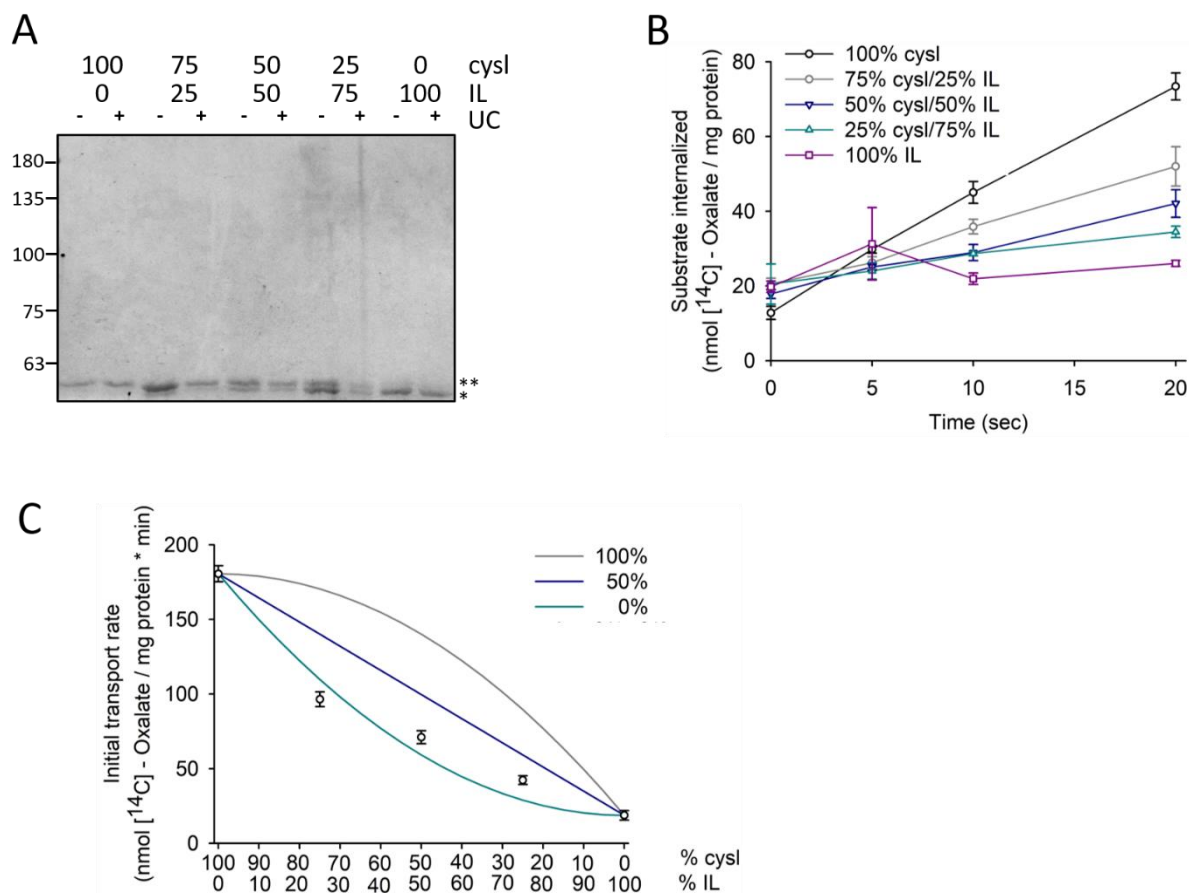
Figure 66: Cooperativity in SLC26Si. (A) SLC26Si^{cysl} and SLC26Si^{IL} were mixed in different ratios and reconstituted 1:50 protein-to-lipid to soy PC. Reconstitution efficiency was estimated by re-solubilizing the protein from the proteoliposomes using 1% (w/v) DDM and comparing protein amounts before (-) and after (+) ultracentrifugation (UC). Samples were separated on 8% SDS-PAGE and visualized by Coomassie staining. (B) Uptake of oxalate into proteoliposomes of SLC26Si^{cysl} and SLC26Si^{IL} mixed in different ratios. Uptake of 1 mM oxalate was measured at 30°C in presence of a pH gradient and an inside positive membrane potential in triplicates, error bars represent standard deviations. (C) Initial transport rates of oxalate uptake in proteoliposomes of SLC26Si^{cysl} and SLC26Si^{IL} mixed in different ratios. Linear fits to initial uptake shown in (F) were performed using SigmaPlot 10, omitting severe outliers. Grey, blue and teal lines indicate the anticipated curves assuming an activity of heterodimers corresponding to 100%, 50% or 0% of the one of wild type homodimers. Error bars shown are from error propagations of standard variations. The whole procedure from purification to determination of initial transport rates was performed twice with similar outcome

5. Discussion

5.1. *The relevance of UPRT activity for uracil uptake in E. coli*

In this study, based on a combination of cell- and proteoliposome-based assays the transport activity of the UraA homolog 51ThXi from *Thalassospira xiamenensis* was investigated in detail. Transport assays on reconstituted 51ThXi revealed that the protein mediates uracil translocation independently of secondary gradients (section 4.2.5). This is in stark contrast to the proton-coupling that has been reported for the *E. coli* uracil transporters UraA and RutG as well as other SLC23 exhibiting other substrate specificities (Goudela *et al.*, 2005; Lu *et al.*, 2011; Papakostas *et al.*, 2013; Botou *et al.*, 2018). Proton-coupled transport has, to our knowledge, so far been only investigated *in vivo* comparing nucleobase uptake in absence and presence of protonophore. A drastic reduction of uracil transport into cells overexpressing 51ThXi when being treated with CCCP has also been observed herein (section 4.2.1). In order to explain the apparent contradiction of results obtained from *in vivo* and *in vitro* uptake experiments, the effect of pH on uracil phosphoribosyltransferase (UPRT), the enzyme that catalyzes the immediate downstream reaction of uracil metabolism, has been investigated (section 4.2.9). Prior to that, it was demonstrated that UPRT activity is crucial for uracil uptake into cells (section 4.2.7). We therefore suggest that inactivation of UPRT in response to a low intracellular pH is the underlying cause for the observed decrease in uracil uptake, and not a proton-coupled transport mechanism.

The dependency of uracil uptake on the activity of UPRT was not only found for cells overexpressing UraA, but also for any other uracil transporter tested, pointing towards a more general role of the downstream reaction (section 4.2.7). The product of UPRT reaction, uracil monophosphate (UMP), was neither competing for the UraA substrate binding site nor stabilizing the protein in a thermal shift assay, indicating that it is not a substrate for the transporter (section 4.2.8). A similar mechanism was described for sugar transporters of the GLUT family (SLC2), which transport mostly glucose that is rapidly phosphorylated by glucokinase or hexokinase. The thiamine transport PnuT from *Shewanella woodyi* (PnuT$_{SW}$) was recently shown to work by a facilitative diffusion mechanism (Jaehme *et al.*, 2018). Genes for members of the Pnu-type family often map together with genes for kinases that are supposed to phosphorylate the transported vitamins (Jaehme and Slotboom, 2015). Both GLUTs and PnuT were demonstrated to work by facilitative diffusion, without coupling to a secondary gradient. Rapid metabolization of the transported molecule maintains the substrate gradient, which alone is required to promote transport. The here presented results suggest that UraA, together with UPRT, works in similar operation mode. While uracil transport over the membrane is mediated by facilitative diffusion, the subsequent conversion of uracil to UMP catalyzed by UPRT maintains the substrate

gradient over the membrane and results in net accumulation of uracil, in the modified form of UMP. The genes for both proteins are found in one bicistronic operon. Genes for homologs of UraA mapping together with UPRT homologs were also identified in other organisms (section 4.2.8, supplement **Table 30**). As uracil is a central compound of cellular metabolic pathways, uracil uptake by facilitated diffusion in combination with intracellular modification would probably ensure rapid intake when needed, similar as described for sugar uptake of GLUTs.

How expression of *upp* and *uraA* is regulated, is not well understood. It was suggested that *upp* expression is down-regulated when pyrimidines are available in excess, in particular in response to high UTP levels. When the intracellular concentration of UTP is low, an alternative transcription site is used, leading to a more productive transcription of the *upp* gene (Tu and Turnbough, 1997). When the *upp* gene was identified, the probable promoter region containing a -10 and -35 consensus sequence was described as well (Andersen *et al.*, 1992). Later experiments of Andersen *et al.* also demonstrate that *upp* and *uraA* are expressed from the same promoter (Andersen *et al.*, 1995). Taken together, it can be assumed that expression of UPRT and UraA is tightly coupled in *E. coli*, consistent with the mechanism described above, in which rapid uracil consumption by UPRT maintains the substrate gradient over the membrane to enable efficient facilitate diffusion mediated by UraA.

The high background uracil uptake in *E. coli* precludes a detailed analysis of uracil transporter activity, as was demonstrated in detail herein (section 4.2.3). The other so far described endogenous uracil transporter in *E. coli*, RutG, is part of a whole *rut* operon comprised of seven genes (RutA to RutG) under the control of a regulator called RutR. RutR was suggested to regulate a number of genes involved in synthesis and degradation of pyrimidines (Shimada *et al.*, 2007; Kim *et al.*, 2010). RutR in turn is controlled by NtrC (nitrogen regulatory protein C) (Loh *et al.*, 2006). Consequently, it has been suggested that the protein expressed from the *rut* operon would be involved in uracil degradation in order to produce nitrogen. Determination of enzymatic functions since then are in line with this hypothesis (Kim *et al.*, 2010; Knapik *et al.*, 2012b, 2012a). Transcription of the *rut* genes is extremely enhanced under conditions of nitrogen starving. Although the *rut* operon is also found in other proteobacteria besides *E. coli*, many of them lack the gene for the RutG uracil transporter (Kim *et al.*, 2010). This points to a minor role of RutG for uracil intake under normal physiological conditions. This is supported by our finding that uracil background uptake was not markedly reduced in a *E. coli* BW25113(*ΔuraA ΔrutG*) strain compared to *E. coli* BW25113(*ΔuraA*) harboring endogenous *rutG* (section 4.2.3).

Usually, *E. coli* does not rely on uracil uptake for the built-up of nucleotides, as uracil (or rather UMP) can also be synthesized *de novo* from arginine (Kantrowitz, 2012). However, it has been shown that

Discussion

exogenous uracil and other nucleobases inhibit the *de novo* synthesis of pyrimidines and purines (Bagnara and Finch, 1974; Christopherson and Finch, 1978). Analysis of cellular extracts demonstrated that intracellular levels of PRPP, that is used by uracil phosphoribosyltransferase to synthesize UMP from uracil, rapidly drop (within 2 min) when uracil is added to the growth medium. Simultaneously, incorporation of radiolabeled uracil into UMP, UDP and UTP increases. It appears likely that salvaging of exogenous uracil by conversion to UMP is favored over the *de novo* synthesis of pyrimidines. The rapid consumption of PRPP in response to adding uracil to the growth medium that has been reported (Bagnara and Finch, 1974) supports our hypothesis that, depending on the overall energy level of the cell, not the transport of uracil by UraA might be the rate-limiting step for overall uracil uptake, but the downstream phosphoribosylation by UPRT. This implies that the so far reported kinetic parameters of UraA, that were determined by classic Michaelis-Menten kinetics, might be inaccurate. In case UPRT activity is rate-limiting, the apparent v_{max} of overall uracil uptake will correspond to the v_{max} of UPRT activity, while the v_{max} of UraA might be higher. Consequently, the K_m of UraA for uracil would also be higher. It should be noted, however, that the reported K_m from whole-cell based uracil transport assays was close to the K_d obtained in SPA experiments on purified protein ($K_m = 0.49 \pm 0.07$ µM and 0.57 ± 0.11 µM, $K_d = 0.41 \pm 0.07$ µM and 143.2 ± 7.5 nM in two different studies (Lu *et al.*, 2011; Yu *et al.*, 2017)).

The so far reported functional analysis on UraA neglects the importance of UPRT activity for *in vivo* uracil uptake. As a consequence, a proton-coupling mechanism has been proposed for UraA and other SLC23 family members. The work presented here indicates that abolishment of uracil uptake into cells in presence of CCCP is not sufficient to show proton-coupling of SLC23 nucleobase transporters. Rather, the inactivation of UPRT due to low internal pH caused by the protonophore results in lower apparent uracil transport rates. Downstream salvaging enzymes, especially phosphoribosyltransferases, exist also for other substrates of SLC23 proteins, for example adenine, guanine or xanthine phosphoribosyltransferases (Maynes *et al.*, 2000; Schramm and Grubmeyer, 2004; Del Arco and Fernández-Lucas, 2017). For some of them, their genes are also located in the same operon as the one of the corresponding transport proteins. Early on, it has been suggested that the activities of adenine and hypoxanthine phosphoribosyltransferases are relevant for the uptake of the respective nucleobases, although the transporters had not yet been identified (Burton, 1977). UPRT activities from different organisms have been reported to be strictly pH-dependent (Linde and Jensen, 1996; Turner *et al.*, 1998; Jensen *et al.*, 2005). In the study presented here, a pH optimum of ~8.5 was determined for *E. coli* UPRT (section 4.2.9). At pH 7.5, the enzymatic activity was reduced to ~40%, and below pH 7.0, no activity at all was observed. For other phosphoribosyltransferases, similar narrow pH-dependent activity profiles were reported. A strong reduction of activity was observed for the

Discussion

E. coli phosphoribosyltransferases specific for guanine-xanthine (Liu and Milman, 1983), orotate (Shimosaka *et al.*, 1985) and adenine (Hochstadt-Ozer and Stadtman, 1971). The guanine-hypoxanthine-xanthine phosphoribosyltransferase of *Schizosaccharomyces pombe*, as well as the homolog specific for adenine from the same organism, showed highest activity at pH 4 to 5, and hardly any activity was detected at pH 7.0 (Houssett and Nagy, 1977). In the same study, a dependency of guanine or adenine uptake rates on activity of the respective phosphoribosyltransferases was observed. Many more examples are reported in literature, but the reason for this pH-dependency of the catalytic activity has not been investigated thoroughly. Sequence alignments of different phosphoribosyltransferases from different organisms revealed a conserved PRPP binding motif, that contains one or two acidic residues which might be involved in this (Del Arco and Fernández-Lucas, 2017). In summary, *in vivo* inactivation of phosphoribosyltransferase or other downstream metabolizing enzymes due to the presence of a protonophore might not be restricted to the uracil-specific transport systems. For UapA, a xanthine-uric acid transporter from *Aspergillus nidulans*, a exchange mechanism of substrate translocation was reported, but only when using the substrate analog allopurinol (Diallinas, 2013). Based on *in vivo* uptake assays in presence of CCCP or DCCP, UapA was defined as a proton-coupled symporter. Transport of allopurinol, however, was not affected when the proton-motive force was abolished, but significantly lower than transport of the natural substrates. Noteworthy, allopurinol is an inhibitor of xanthine oxidase/dehydrogenase, an enzyme catalyzing a downstream reaction of xanthine uptake. Homologs of this enzyme were also reported to be strictly pH-dependent (Greenlee and Handler, 1964; Carro *et al.*, 2009), meaning, a similar inactivation in buffer with low pH in presence of protonophore might happen as we observed for uracil phosphoribosyltransferase. Allopurinol, in inactivating xanthine oxidase, might cause a similar effect, revealing the proton-uncoupled substrate transport mediated by UapA. In total, it seems necessary to re-evaluate the proposed proton-coupling in SLC23 nucleobase transporters.

5.2. Uniporter function of SLC23 nucleobase transporters

51ThXi is, so far, the only SLC23 protein for which an *in vitro* transport assay could be developed. Although both UraA and 55PyPi, a UraA homolog from *Pyramidobacter piscolens*, were well-incorporated into the liposomes, no distinct uptake of radiolabeled uracil could be recorded (section 4.2.5). Cell-based uracil uptake assays, on the other hand, could be established for both of them. The reason for the apparent inactivity of reconstituted UraA and 55PyPi remains elusive. In the inward-open structure of UraA (PDB: 3QE7), a detergent molecule (ß-nonyl-D-glucoside) was found close to the substrate binding site (**Figure 67**). Its acyl chain protrudes into the scaffold domain between TM5 and TM12, while the head group probably interacts with uracil (Lu *et al.*, 2011). It is conceivable that a detergent molecule at this position, interacting with both scaffold and substrate binding domain, interferes with the elevator movement of the protein. For the non-related secondary active transporter LmrP, an inactivating effect of Triton X-100, Triton X-114 and Tween 80 on the reconstituted transporter has been reported, which was not observed when the protein was purified using DDM instead (Putman *et al.*, 1999). To account for a possible effect of detergent on activity, UraA was purified with both DM and DDM and reconstituted, but no activity was observed in both cases (**Figure S1**). In the UraA dimer (PBD: 5XLS), no detergent molecule is found at a similar position. In order to obtain the structure, the protein was purified using a mixture of n-nonyl phosphocholine 9 (Fos-Choline 9) and n-undecyl phosphocholine (Fos-Choline 11) (Yu *et al.*, 2017). The use of non-maltoside detergents for UraA purification may provide a route for establishing a proteoliposome-based activity assay for this protein.
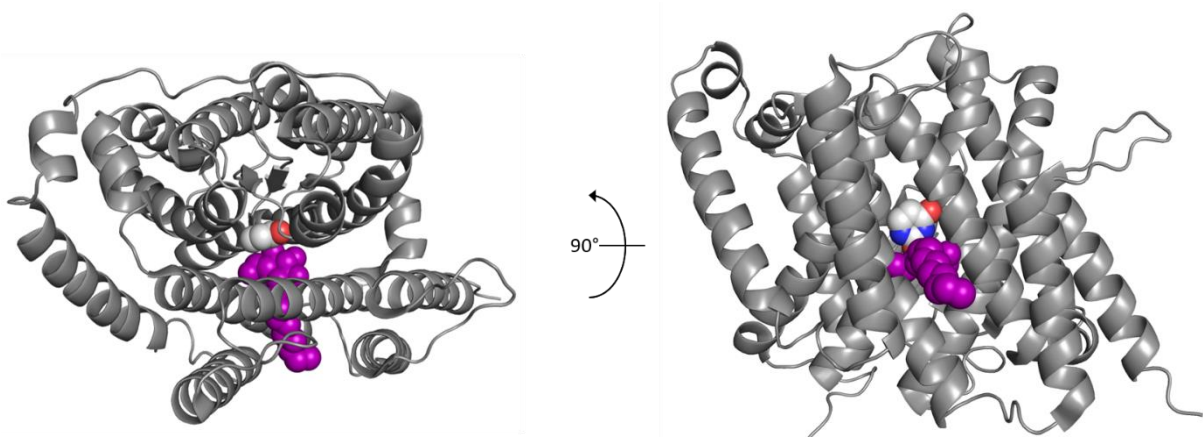
Figure 67: Co-crystallization of UraA with a detergent molecule. In the inward-open structure of UraA (PDB: 3QE7), a ß-nonyl-D-glucoside (ß-NG) molecule is located close to the substrate binding site. The acyl chain of the detergent is protruding between TM5 and 12 into the scaffold domain. UraA is shown in grey from the top view (left) and the side view (right) with ß-NG in purple spheres.

Herein, detailed *in vitro* studies could only be performed on 51ThXi; uniporter function therefore was only demonstrated for this one member of the family. Nevertheless, the results on 51ThXi substrate transport presented in this thesis represent the only conclusive study on proton-coupling of SLC23 transporters.

The proteoliposome-based assay developed for 51ThXi provided the basis for a detailed analysis of the transport mechanism of a member of the SLC23 family that had not been possible so far. 51ThXi was demonstrated to transport not only uracil, but also 5-fluorouracil and cytosine (section 4.3.1). In this, re-localization of the empty transporter appears to be the rate-limiting step, as substrate exchange was much faster than efflux (section 4.2.6).
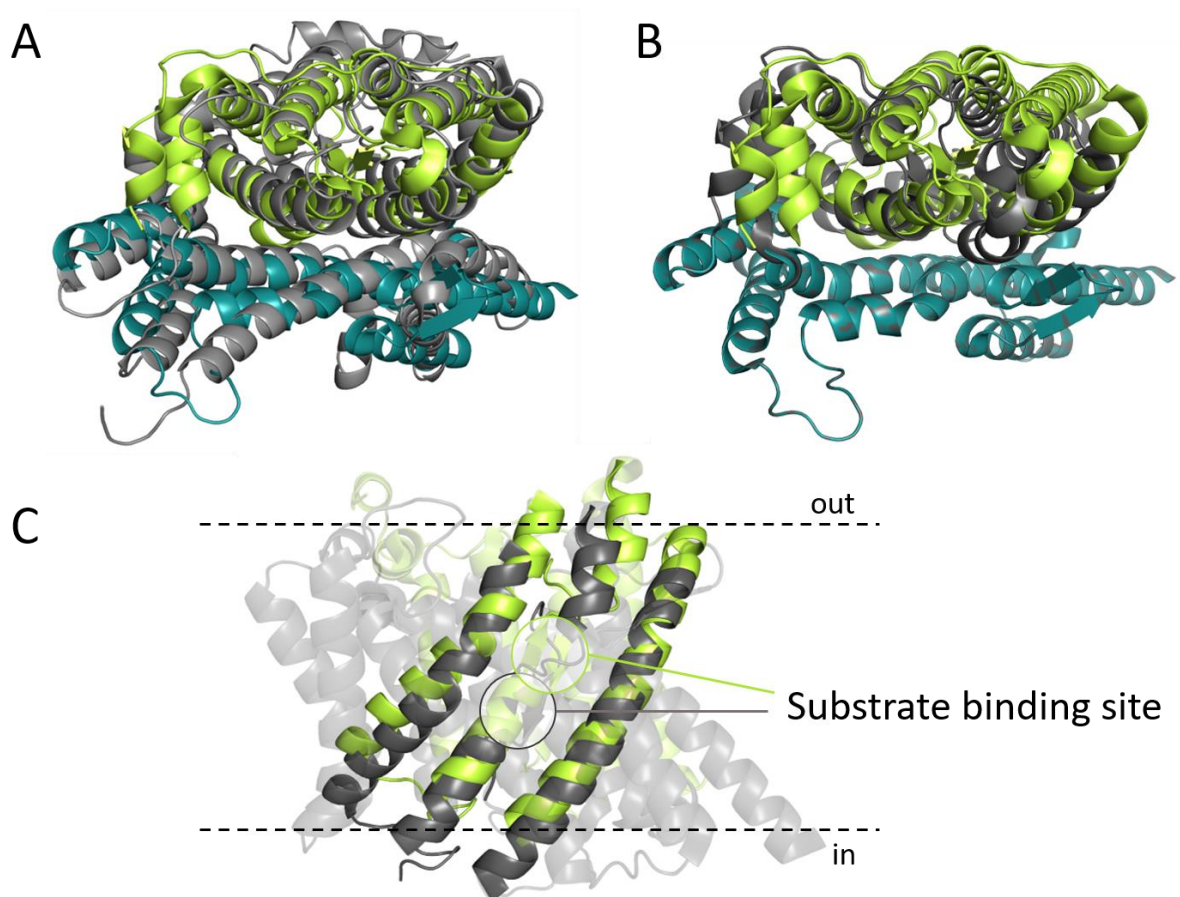
Figure 68: Model of UraA in an outward-open conformation based on the structure of AE1. (A) A model of UraA in an outward-open conformation was generated by separately aligning the transport and scaffolds domains (light green and teal, respectively) to the AE1 structure (PDB: 4YZF) (grey). (B) Superimposition of UraA in the inward-occluded conformation (PDB: 5XLS) and the model of the outward-open conformation in top view. The inward-occluded conformation is shown in grey and the transport and scaffold domains of the model in light green and teal, respectively. (C) Side view of the superimposition of UraA in inward-occluded and outward-open conformation to visualize movement of TM1, 3, 8 and 10 and the substrate binding site. Whole proteins are transparent, the relevant helices are opaque.

Discussion

The members of the 7TMIR protein families (SLC4, SLC23 and SLC26) have been proposed to facilitate substrate translocation *via* an elevator mechanism based on structural alignments of transporters captured in the inward- and outward-open conformation. Comparing the inward-open conformation of SLC26Dg with the outward-open conformation of AE1 suggests a 6 Å vertical translation accompanied by a 30° rotation of the transport domain against the scaffold domain during the conformational change (Chang and Geertsma, 2017). A prerequisite for this is the assumption that the scaffold domains remain relatively static. So far, structures of SLC23 and SLC26 proteins have only been solved in inward-open or inward-occluded conformations, while high-resolution SLC4 structures are available solely from an outward-open state. Furthermore, the SLC23 proteins were crystallized with substrate bound, while all the structures of SLC4 and SLC26 (with the exception of BicA, a cyanobacterial bicarbonate transporter) are without substrate. This as well as the low sequence similarity of the SLC4, SLC23 and SLC26 families complicate, or preclude, a prediction of subtle protein changes accompanying the conformational change. A model of the outward open state of UraA, based on the structure of AE1, suggests an upward movement of the transport domain that appears compatible with an opening of the substrate binding site to the periplasm (**Figure 68**). To generate the model, the scaffold and transport domains of UraA were aligned separately to the outward-open structure of AE1, in a similar approach as was reported for prediction of the elevator movement of SLC26Dg (Chang and Geertsma, 2017).

For many elevator proteins, structural "gating" elements have been described that open and close during the transport cycle (Garaeva and Slotboom, 2020). Usually, these gates consist of helical hairpins that prevent an (efficient) movement of the transport domain against the scaffold domain in the unloaded transporter. Substrate binding is accompanied by a movement of the helical hairpin (gate closure), allowing the elevator movement and substrate translocation to take place. For the 7TMIR carriers, nevertheless, such gating elements were so far not detected. In the case of the ion-coupled transporters of the SLC4 and SLC26 families, transport mechanisms can be envisioned in which ion binding would enable transport domain movement in a similar fashion, for example if electrostatic interactions between transport and scaffold domain would be disturbed that prevent the domain movement in absence of the secondary substrate. For VcINDY and LaINDY, two members of the DASS (divalent anion sodium symporter, SLC13) family, it has been recently suggested that sodium binding causes a net positive charge that hinders the elevator movement. Binding of the negatively charged substrate counteracts this positive charge, allowing the conformational change to take place (Sauer *et al.*, 2020). Such more subtle "gating" mechanisms, however, are difficult to identify without structural information from more than one transporter conformation or substrate-bound states.

Discussion

For the SLC23 nucleobase transporters on the other hand, it appears that the conformational change can take place without additional binding of secondary ions. In this, the empty transporter moves much slower than the substrate-loaded one, as is suggested by the slower efflux than exchange rate in proteoliposomes (section 4.2.6). How substrate binding would accelerate the conformational change is yet unknown. The binding site contains a number of residues that can be protonated, a glutamate and a histidine on TM8 and another glutamate on TM10. A calculation of the theoretical pK_a of these residues (performed by Dr. Ahmadreza Mehdipour) suggest that E241 and H245 of UraA are charged under physiological conditions in both absence and presence of uracil (**Table 29**). The presence of has the largest effect on the pK_a of E290, but also for this residue, the theoretical calculations indicate that the protonation state does not change upon substrate binding. In summary, the pK_a calculations are in line with a proton-independent transport mechanism.

Table 29: Theoretical pK_a of substrate binding site residues E241, H245 and E290 of UraA. Calculations were performed for both available structures of UraA (inward-open, PDB: 3QE7, and inward-occluded, PDB: 5XLS) by Dr. Ahmadreza Mehdipour using a multiconformation continuum electrostatics (MCCE) approach (Song *et al.*, 2009).

	E241		H245		E290	
	Without uracil	With uracil	Without uracil	With uracil	Without uracil	With uracil
3QE7	1.1	2.8	8.2	8.4	4.6	2.8
5XLS	0.8	1.8	9.0	9.8	6.1	4.3

Of the three charged substrate binding site residues, at least the ones on TM8 were demonstrated to not be absolutely crucial for substrate binding (section 4.3.2.1). When mutated to alanine, in *in vivo* transport assays uracil uptake was markedly reduced, while in thermal shift assays, uracil binding to the mutants could still be observed. At the moment, it cannot be excluded that the reduction in transport is caused by a lower affinity of the mutants for the substrate that is not detected in the binding assays as they were performed here. A detailed determination of affinities was not possible within the scope of this thesis due to the high background uracil uptake in the whole cell-based transport assays and the pronounced instability of the purified mutants (section 4.3.2.2). In both available UraA structures, all three residues are involved in coordination of uracil. All three of them together probably contribute to high-affinity uracil binding. When one of them is replaced by alanine, substrate binding in general is still possible, but a lower affinity would lead to reduced transport as observed here.

Replacing either of the residues on TM8 (E256/H260 in 51ThXi) with alanine resulted in a severe reduction of the melting temperature of the apo state (section 4.3.2.2). This might be caused by disturbing the network of electrostatic interactions within the substrate binding site, leading to its destabilization and subsequently of the protein in total. Interestingly, substrate binding was able to recover protein stability. In thermal shift assays, the shift in melting temperature in presence of uracil was similar for wild type 51ThXi as for the E256A or H260A mutant (**Figure 37**). In absence of substrate, the substrate binding site residues most likely interact in a different way to compensate the multiple charges within the binding site, probably involving water molecules. In this context, it is still unclear how substrate binding leads to a more efficient conformational transition of the transporter. Substrate binding might lead to a reorientation of electrostatic interactions within the substrate binding site that might stabilize the transport domain in a state that can move more efficiently against the scaffold domain.

The presence of two glutamates and one histidine in the substrate binding site suggests some form of charge compensation, especially as the glutamate and histidine on TM8 are located immediately next to each other. Two findings argue against a pronounced relevance of this: First, the histidine on TM8 is found to be an aspartate in ~50% of UraA homologs, and also in the uracil transporter 55PyPi that was investigated in this study. Although its functionality was not studied as in the detail as the one of UraA and 51ThXi, no indications for an impaired uracil transport capacity or protein stability were observed (section 4.3.1). Furthermore, also the histidine in UraA and 51ThXi could be replaced by aspartate without a massive impact on transport activity (section 4.3.2.4). Second, in the UapA xanthine transporter an aspartate is found in place of histidine and a glutamine in place of glutamate on TM10. This further indicates that charge compensation of the TM8 glutamate is not necessary for protein stability or functionality.

Strikingly, the conserved negative charge on TM8 can not only be found in SLC23 transporters, but also in the structurally related, but functionally unrelated 7TMIR families of SLC4 and SLC26 transporters (Chang and Geertsma, 2017). The only exceptions are the human SLC26 transporters that harbor a neutral amino acid at that position. For SLC4 and SLC26 proteins, examples of electrogenic and electroneutral symporters and exchangers have been reported, as well as both cases of coupling substrate transport to proton or sodium flux (section 1.4.3). Consequently, the conserved charged residue on TM8 has been suggested to be relevant for coupling (section 4.2.1). Proton-independent transport of SLC23 transporters has so far, within this study, only been shown for one example. However, the data presented here strongly suggests that the existing information needs to be re-evaluated. In light of uncoupled substrate transport of the SLC23 family, the role of the TM8 glutamate in coupling cannot be conserved throughout the whole 7TMIR class of proteins. The effect of removing

the TM8 charge on protein stability might provide another explanation for the extreme conservation of this residue in proteins of different families that exhibit different substrate specificities and different modes of transport. As a common characteristic, substrate binding sites of proteins that have been shown or proposed to work by an elevator mechanism are located in the center of discontinuous helices. The precise relevance of this discontinuous helices for the elevator movement is largely unknown. For some elevator proteins it has been suggested that the dipole moment of the discontinuous helix would be involved in substrate or ion coordination, for example in AE1 that shares the same architecture as the SLC23 family (Ficici *et al.*, 2017; Abbas *et al.*, 2018). For the Na$^+$/H$^+$ exchanger NHE9, D215 and R441 were proposed to neutralize the dipole moments of two oppositely oriented half-helices (Winkelmann *et al.*, 2020). Studies on the *E. coli* Na$^+$/H$^+$ antiporter NhaA indicates that K300, located close to the C-termini of the interrupted helices TM4 and TM11, is crucial for the stability of the protein (**Figure 69**) (Lee *et al.*, 2014; Călinescu *et al.*, 2017). Compensation of the helix dipoles was also suggested to being relevant for other Na$^+$/H$^+$ antiporters as well as the ASBT (SLC10) family (Schushan *et al.*, 2010; Hu *et al.*, 2012; Wang *et al.*, 2014). In UraA, TM3 and TM10 that are protruding only halfway into the membrane might create a surplus of positive charge that is counteracted by E241 in an analogous fashion. Similar as K300 for NhaA, E241 of UraA (or E256 for 51ThXi) was demonstrated to be important for protein stability (section 4.3.2.2). As the SLC4 and SLC26 share the same fold with the SLC23 family, but harbor distinct substrate specificities and different modes of transport, this relevance for protein stability in neutralizing the positive helix dipoles at the substrate binding site might provide an explanation why this residue has been conserved throughout evolution in the 7TMIR proteins.
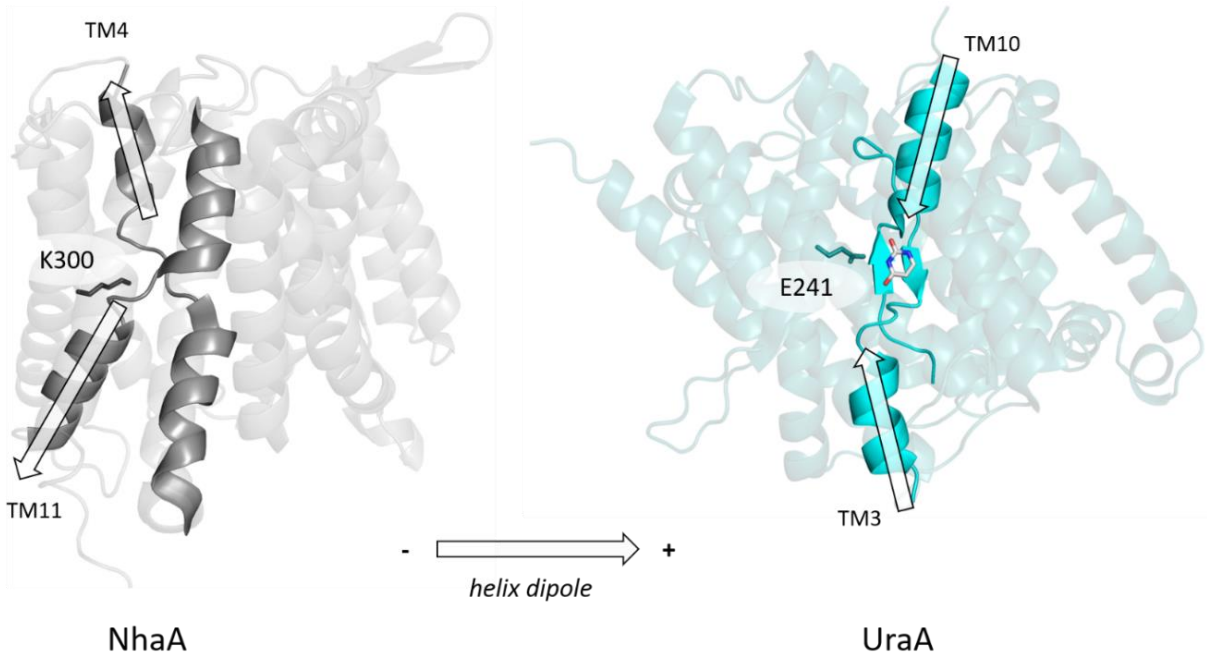
Figure 69: Helix dipole compensation in NhaA and UraA. Elevator proteins contain discontinuous helices which might create a surplus of charge, based on the intrinsic helix dipole moment, at the substrate binding site. The orientation of the proposed helix dipoles is indicated with arrows. For NhaA, K300 was suggested to compensate for the negative charge created by the dipole of helices 4 and 11 (left). In UraA, the dipoles are orientated in an opposite direction, and E241 might serve a similar role by neutralizing the surplus of positive charge (right).

5.3.   Substrate specificity of SLC23 transporters

Concerning the substrate specificity of SLC23 transporters, the best studied ones are the homologs from *E. coli,* albeit only UraA has been structurally resolved. Next to that, UapA from *Aspergillus niduans* has also been functionally well-characterized. Additionally, uracil uptake has been investigated in some protozoan parasites, such as *Trypanosoma brucei brucei* and *Leishmania major*, and the honey bee pathogen *Paenibacillus larvae* (De Koning and Jarvis, 1998; Papageorgiou *et al.*, 2005; Gudin *et al.*, 2006; Stoffer *et al.*, 2018; Stoffer-Bittner *et al.*, 2018; Campagnaro and de Koning, 2020). Protozoan parasites generally lost the ability of *de novo* synthesis of nucleobases and rely on the uptake from the host environment through various transport systems. The reason for the interest of uracil transporters in pathogens is that many of them also efficiently transport 5-fluorouracil, which is a potent cytotoxic agent. As the human SLC23 transporters do not transport nucleobases, but ascorbate, 5-fluorouracil (and other nucleobase analogs) have been suggested, and also used, as antimetabolites against different diseases and for treating viral or bacterial infections (De Koning and Diallinas, 2000). In mammals, nucleobase uptake is mediated mostly by the SLC29 Equilibrative Nucleoside Transporter (ENT) family (Baldwin *et al.*, 2004). 5-fluorouracil (brand name Adrucil) has been used for the last 60 years as effective anticancer drug, despite its considerable adverse effects. Tumor cells were shown to exhibit an increased uptake of nucleobases compared to non-carcinogenic cells in order to promote their fast growth rate (Diasio and Harris, 1989; Longley *et al.*, 2003). In humans, SLC23A1 and SLC23A2 mediate ascorbate uptake in a sodium-dependent manner rather than uptake of purines or pyrimidines (Savini *et al.*, 2008). In rats, though, rSNBT1 (SLC23a4) was found to be transporting a number of nucleobases, including uracil, guanine, urate, xanthine and hypoxanthine (Yamamoto *et al.*, 2010b; Yasujima *et al.*, 2018). The gene for SLC23A4 in humans was also identified, but it is not well or not at all expressed (Yuasa *et al.*, 2020).

Based on the detailed mutagenesis studies on the *E. coli* xanthine transporter XanQ, it was suggested that substrate specificity of SLC23 nucleobase transporters is governed by a few conserved residues, and that changing these residues would also change the specificities of the transporters (Frillingos, 2012; Papakostas and Frillingos, 2012; Karena *et al.*, 2015; Chaliotis *et al.*, 2018; Kourkoulou *et al.*, 2018, 2021). This hypothesis even led to the development of an algorithm that is supposed to predict the substrate specificity of SLC23 proteins solely based on sequence alignments on 1355 NCS2 homologs from 4655 bacterial and archaeal proteomes (the NAT/NCS2-hound, http://bioinf.bio.uth.gr/nat-ncs2/#/) (Chaliotis *et al.*, 2018). Earlier on, several XanQ mutants were identified that showed a changed binding behavior towards different non-natural xanthine analogs (reviewed in Frillingos, 2012). On the other hand, changing the spectrum of substrates that are not only binding, but transported, proved to be much more difficult. Only the N93S and N93A mutants of

XanQ were shown to additionally transport uric acid, but with very low rates compared to xanthine transport and also only when high substrate concentrations were used, indicative of a low affinity (Karena and Frillingos, 2011). Noteworthy, uric acid and xanthine are very similar, differing only in the imidazole moiety (**Figure 70**). Consequently, it can be assumed that only small changes to the protein should suffice to convert a xanthine transporter into a uric acid transporter and *vice versa*. Nevertheless, to date it has not been possible to create a high-capacity uric acid transporter from XanQ. UapA in contrast transports both xanthine and uric acid and shows similar affinities for both substrates by nature (Diallinas *et al.*, 1998). Changing the substrate profile of UapA has so far been more successful in general, enlarging the binding profile to include more natural and unnatural purine compounds by single or multiple side chain mutations (Meintanis *et al.*, 2000; Koukaki *et al.*, 2005; Papageorgiou *et al.*, 2008; Kosti *et al.*, 2010; Diallinas, 2016). Many of these mutants had been investigated before the first crystal structure of UraA was available, and were therefore designed based on a topology model made of 12 transmembrane helices, as prediction algorithms typically fail to identify the two half-helices 3 and 10. A later re-evaluation revealed that many of the mutations influencing UapA substrate specificity map to the substrate binding site (Kosti *et al.*, 2012). The UapA structure largely confirmed the observations made based on the model on UraA (Alguel *et al.*, 2016).
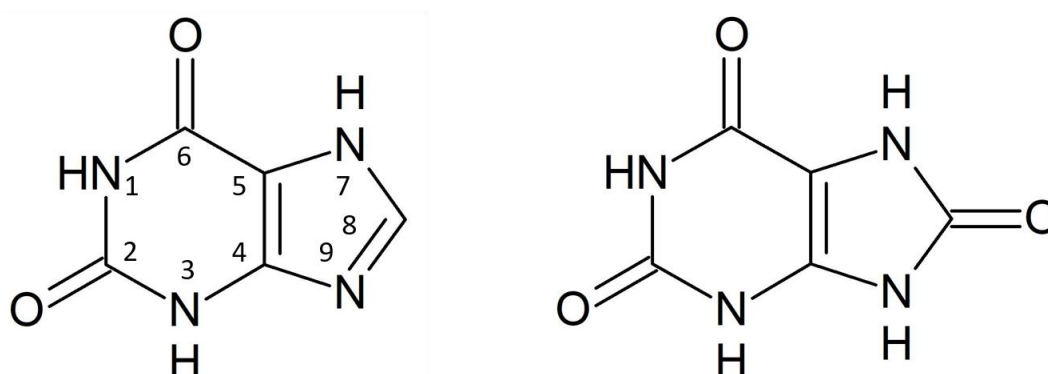
Figure 70: Structures of xanthine and uric acid. Xanthine (left) and uric acid (right) differ in an additional ketone group at position 8 of the imidazole moiety.

Interestingly, some mutations that were found by to alter the substrate specificity are also located away from the binding site, and were therefore suggested to be additional "selectivity filters" (Vlanti *et al.*, 2006). In particular, these are Q113 on TM2, A441 on the helix connecting TM11 and 12, V463 and A469 on TM12, R481 on TM13 and T526, G527 and F528 on TM14 (**Figure 71A**). In this, R481 is extraordinary as it points away from the protein and the proposed substrate translocation pathway. In the dimeric structure of UapA, however, it can be seen that R481 of one protomer is located near the substrate binding site in the other protomer (**Figure 71B**). The effect on substrate binding was explained assuming that R481 of one protomer forms transient hydrogen bonds and π-π-stacking with

the substrate and interacts with Q408 of the substrate binding site of the opposing protomer. This hypothesis was supported by molecular dynamics simulation, but no experimental data was provided. It should be noted that the arginine is not conserved and only found in UapA and UapC, another endogenous xanthine-uric acid transporter from *Aspergillus nidulans* (Kosti *et al.*, 2012). In XanQ, a threonine (T397) is located at that position, in UraA a leucine (L366). The suggested relevance of the arginine for substrate translocation therefore appears to be unique to UapA. The other residues for which mutants with an increased substrate specificity were found are either located at the interface of the substrate binding and scaffold domain, or at a hinge region between both domains (Q113 and A441). How this might affect the transport mechanism by making the movement of the substrate binding domain against the scaffold domain more or less efficient, is easily conceivable; why mutations at these positions would cause an increased substrate specificity is harder to explain, and has not been addressed experimentally in much detail. Interestingly, the observed effects were strongest when several mutants were combined, especially together with Q408 in the substrate binding site. The Q408E mutation on its own allows binding, but not transport of hypoxanthine, which is not a substrate of wild type UapA. In combination with R481G and F528S or T526M, not only hypoxanthine, but also adenine transport was observed (Alguel *et al.*, 2016).
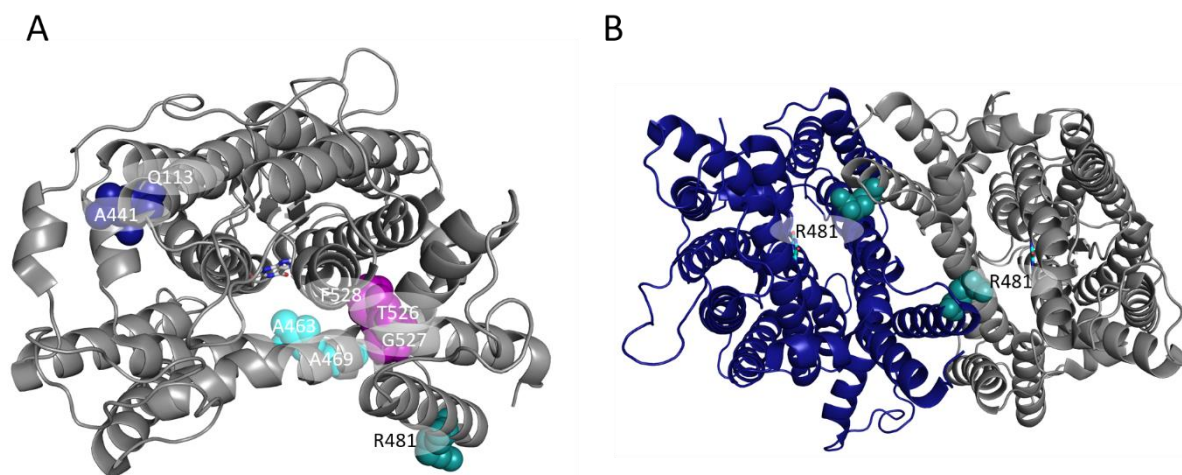
Figure 71: Residues in UapA for which mutants show an increased substrate specificity. (A) When mutants of UapA were identified that showed an increased substrate specificity compared to the wild type protein, several residues were found that are not located at the substrate binding site, in particular Q113 on TM2, A441 on the helix connecting TM11 and 12, V463 and A469 on TM12, R481 on TM13 and T526, G527 and F528 on TM14. All residues are annotated and shown as spheres (Q113 and A441 in blue, V463 and A469 in cyan, R481 in dark cyan and T526, G527 and F528 in purple). UapA is shown in top view and as a monomer for simplification. Xanthine is shown in sticks. (B) In the UapA dimer, R481 of one protomer is located close to the substrate binding domain of the other protomer. The UapA dimer (PDB: 5I6C) is shown in bottom view with one protomer colored in dark blue, the other in grey. The two arginines are shown as spheres in dark teal and the substrate is shown in sticks.

These mutations were identified based on growth assays, selecting for *Aspergillus nidulans* mutants that are able to utilize the respective purine as nitrogen source, which is directly linked to their

capability for transport. So far, this growth-based assay is unique to *Aspergillus nidulans* and was therefore used only to investigate the function of UapA or UapC, and, more specifically, only for purines, as this organism cannot use pyrimidine compounds for nitrogen salvage in a similar fashion. This kind of assay enabled the identification of gain-of-function mutants using a UV-induced, random mutagenesis approach rather than by rational design (Kosti *et al.*, 2010). The development of a similar growth-based assay for the investigation of uracil uptake in *E. coli* failed in our hands, caused by the high background uracil uptake that is also retained in the *E. coli* BW25113(*Δura ΔrutG*) knockout strain lacking both so far identified uracil transporters in *E. coli* (section 4.2.3). Andersen *et al.* identified mutants deficient in uracil uptake based on their ability to use cytosine as a pyrimidine source combined with a resistance to 5-fluorouracil (Andersen *et al.*, 1995). We took that as a hint that it should be possible to develop similar growth-based assays using uracil as pyrimidine source. However, although uracil and 5-fluorouracil are closely related, transporters of uracil are not necessarily also recognizing the toxic analog, i.e., transporters specific for uracil over 5FU are easily missed in this kind of assay.

That mutations of substrate binding site residues in UapA on their own did not result in transporters with an increased or modified substrate specificity fits to our finding that mutating the binding site residues of the uracil transporter 51ThXi to the ones found in UapA, i.e., changing the TM8 histidine to aspartate and the TM10 glutamate to glutamine, did not result in any detectable uptake activity for xanthine, nor uracil (section 4.3.3.1). As the mutant could not be purified due to pronounced instability although expression levels were unaffected (not shown), it could not be determined whether the mutant might be able to bind any of the two compounds. Still the analysis of the specificity of novel UraA homologs (section 4.3.3.3) clearly showed that the binding sites of uracil transporters preferably contain a glutamate on TM10 and xanthine transporters a glutamine on the corresponding position. Furthermore, all newly identified xanthine transporters harbor an aspartate on TM8, while for uracil transporter, histidine or aspartate residues were found at the corresponding position. For the xanthine transporter XanQ, D276 of TM8 could be replaced by glutamate, while all other mutations, including D276H, led to reduced or completely abolished transport activity (Mermelekas *et al.*, 2010). The D360H mutant of UapA shows no functionality and has a higher vacuolar turnover, indicating intrinsic protein instability (Kosti *et al.*, 2012). In our experiments, the UraA or 51ThXi TM8 histidine could be exchanged for aspartate without any pronounced effects on transport, indicating that the nature of this residue is less relevant for specificity and substrate translocation in uracil transporters. On the other hand, the alanine mutants in both cases showed reduced uptake activity, while capacity for binding was preserved, although a reduction in affinity cannot be excluded (**Figure 33**, **Figure 36**). This indicates that charge or polarity of this residue might have some impact on transport. However, aspartate would

rather be expected to be negatively charged, while H245 in UraA is positively charged (**Table 29**). This appears contradictory; replacing the positive charge on TM8 (H245 in UraA, H260 in 51ThXi) with a negatively charged aspartate will also affect the pK_a values of the neighboring glutamates on TM8 and TM10 (E241/E290 in UraA, E256/E303 in 51ThXi). Taken together, it is still unclear how certain residues on specific positions in the substrate binding site are connected to distinct substrate specificities of SLC23 nucleobase transporters. The finding that RutG (and the herein newly identified homolog YeEn03 from *Yersinia enterolitica*) transports both uracil and xanthine while harboring the uracil transporter-like binding site (His/Glu on TM8/10) might provide valuable insights into this. The difficulties in purifying and reconstituting RutG within the scope of this thesis, nevertheless, precluded a more detailed analysis on the binding and transport behavior of this protein so far.

Purines are built up from a pyrimidine ring that is extended by an imidazole moiety. The pyrimidine ring of xanthine, in this respect, perfectly resembles uracil (**Figure 72A**). Xanthine thus can be viewed as uracil with additional (bridged) side groups at C5 and C6. The orientation of the pyrimidine part of xanthine resembling uracil, however, is different in UapA from the orientation of uracil in UraA in the published structures (**Figure 72B**). While uracil appears to be orientated in such a way that E241 is in hydrogen-bonding distance with N3, His245 with O2 and E290 with N1, xanthine is rotated by ~90° so that Q408 interacts with both O6 (O4 in uracil) and N1, while E356 is close to N9. Alguel *et al.* suggest an alternative orientation of xanthine that poses N7 close to E356, but this preserves the relative orientation of xanthine compared to uracil (Alguel *et al.*, 2016). From this structural point of view, there is no obvious reason why a glutamine on TM10 would prevent uracil binding, yet no binding of uracil to the E303Q mutant of 51ThXi could be observed (section 4.3.2.3). Also the general high similarity of the UraA and UapA binding sites and the high degree of conservation within the SLC23 proteins suggest that the transporters should be able to recognize both pyrimidine and purine compounds *per se*. So far, nevertheless, many SLC23 transporters for which substrate specificities were investigated in more detail mostly are restricted to bind either only purines or only pyrimidines, but the reason for this remains largely unclear. In a recent study, Kourkoulou *et al.* describe a mutant of UapA that is able to recognize a broad variety of purines and pyrimidines (Kourkoulou *et al.*, 2021). This UapA variant combines four adjacent mutations on TM10 (T405S/F406Y/A407S/Q408E) with different substitutions on F528, which is located on TM14 on the scaffold domain. Molecular docking simulations revealed that the most probable orientation of uracil in the substrate binding site of this mutant resembles the one in UraA, while it binds in a different orientation to wild type UapA. Interestingly, this mutant binds uracil with a moderate affinity (100 µM - 300 µM) but shows no uracil transport. A model of UapA in the outward-open conformation, based on the structure of AE1, places F528 within ~3.5 Å to the conserved glutamate on TM8 (E356 in UapA, E241 in UraA). In the inward-

open structure of UapA, these residues are ~6.0 Å apart. Kourkoulou *et al.* suggest that an interaction of F528 with E356 *via* delocalized π-electrons 'locks' UapA in the outward-open conformation; substrate binding affects the interaction of these residues in a way that allows the substrate binding domain to move against the scaffold domain. Modification of F528 thus indirectly affects the substrate specificity of UapA. Further experimental data will reveal the extent to which this mechanism can be generalized to other SLC23 transporters.
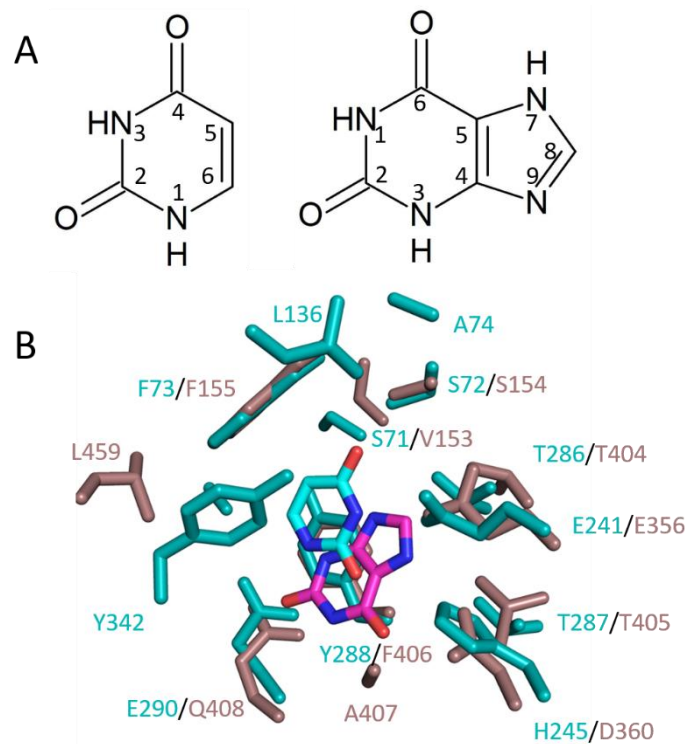
Figure 72: Orientation of uracil and xanthine is different in UraA and UapA. (A) Structures of uracil (left) and xanthine (right). The pyrimidine moiety of xanthine resembles uracil. (B) Superimposition of residues of UraA (PDB: 3QE7, dark teal) and UapA (PDB: 5I6C, dark pink) that are found in 5 Å distance around the respective ligands, uracil in teal and xanthine in pink.

There are a few exceptions of SLC23 transporters that appear to be naturally capable of binding both purines and pyrimidines, for example two homologs from *Paenibacillus larvae*: A UraA-like protein termed PlUP (NCBI reference sequence: WP_023482778.1), that shares 61% sequence identity with UraA, was identified as a transporter for uracil, but did also bind xanthine and uric acid (Stoffer *et al.*, 2018). The putative binding site residues of PlUP (with the exception of F346, that is Y342 in UraA) are absolutely identical to the ones of UraA (**Figure S31**). The other homolog, PlUacP (NCBI reference sequence: WP_023485062.1), transports adenine, guanine and uracil and additionally binds hypoxanthine, xanthine, cytosine and uric acid (Stoffer-Bittner *et al.*, 2018). This protein shares higher sequence similarity with the uric acid transporter UacT from *E. coli* (43%) than with UraA (28%). Both

cases again indicate that the nature of the substrate binding site residues alone cannot be used to predict the specificities of the respective transporters.

To date, it also has not been possible to explain why SLC23 proteins function as nucleobase transporters in almost all kingdoms of life, but transport ascorbate instead in higher eukaryotes. As already mentioned above, nucleobase uptake in humans is performed by the members of the SLC29 family, that also are able to transport nucleosides. Remarkably, SLC29 proteins are only found in (higher) eukaryotes (Young, 2016), suggesting that, evolutionary, SLC23 transporter function might have shifted to the SLC29 family. SLC29 transporters are part of the Major Facilitator Superfamily (MFS) and are structurally unrelated to 7TMIR transporters (Wright and Lee, 2019). Interestingly, together with the ability to transport nucleobases, uracil phosphoribosyltransferase activity appears to be less important in eukaryotes. In humans, a gene coding for a UPRT homolog was identified, but no UPRT function could be detected for the protein (Li *et al.*, 2007). Even more remarkable, SLC23 ascorbate transporters act in a sodium-dependent manner, as has been intensively studied, e.g. for SLC23A1 (SVCT2) and SLC23A2 (SVCT1) from human, which has been summarized in several reviews (Savini *et al.*, 2008; May, 2011; Bürzle *et al.*, 2013). To date, no high-resolution structures for SLC23 ascorbate transporters are available, and the sodium-binding sites have not been identified yet. In general, the transport properties of human SVCT1 and SVCT2, that share a sequence identity of 66%, appear to be very similar. For both, a sodium to ascorbate stoichiometry of 2:1 was determined, while transport by SVCT1 seems to be electrogenic in contrast to electroneutral transport by SVCT2 (Tsukaguchi *et al.*, 1999a; Godoy *et al.*, 2007). Within the SLC23 family, the ascorbate transporters appear to be closer sequence-related to the uracil and xanthine transporters than to the ones specific for guanine and adenine, as exemplified by aligning their sequences to the *E. coli* nucleobase transporters with known functions, indicating no general diversification of SLC23 vitamin C and purine/pyrimidine transporters at an early stage of evolution (**Figure 73A**).

A

GhxP	0.08774	*guanine*
GhxQ	0.09266	*guanine*
AdeP	0.13407	*adenine*
AdeQ	0.12778	*adenine*
UapA	0.35895	*xanthine, uric acid*
XanQ	0.28023	*xanthine*
XanP	0.28075	*xanthine*
SVCT1	0.18313	*ascorbate*
SVCT2	0.17958	*ascorbate*
UacT	0.35439	*uric acid*
UraA	0.30618	*uracil*
RutG	0.30696	*uracil, xanthine*

B
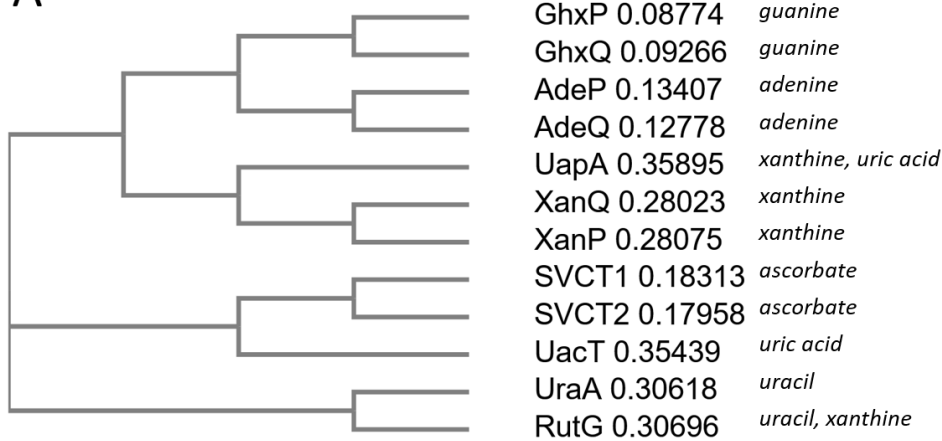

```
                                                                    TM1
UapA  ------------MD-----------NSIHSTDGPDSVIPNSNPK--KTVRQRVRLLARHL       35
UraA  ------------------------------------------------------------       0
SVCT1 -----------------------------------------------------------M       1
SVCT2 MMGIGKNTTSKSMEAGSSTEGKYEDEAKHPAFFTLPVVINGGATSSGEQDNEDTELMAIY      60

UapA  TTREGLIGDYDYGFLFRPELPFMKKDPRAPPFFGLNEKIPVLLAFILGLQHALAMLAGVV      95
UraA  --------------------------MTRRAIGVSERPPLLQTIPLSLQHLFAMFGATV      33
SVCT1 RAQEDLEGRTQHETTRDPSTP-LPTEPKFDMLYKIEDVPPWYLCILLGFQHYLTCFSGTI      60
SVCT2 TTENGIAEKSSLAETLD-STG-SLDPQRSDMIYTIEDVPPWYLCIFLGLQHYLTCFSGTI     118
                            :.:   *      :  *.:**  ::  :....:
                       TM2
UapA  TPPLIISSSLSLPSDLQ--QYLVSTSLIVCGLLSMVQITRFHIYKTPYYIGSGVLSV---     150
UraA  LVPVLFHINPAT-------------VLLFNGIGTLLYLFICKG-KIPAYLGSSFAFISPV      79
SVCT1 AVPFLLAEALCVGHDQHMVSQLIGTIFTCVGITTLIQ-TTVGI-RLPLFQASAFAFLVPA     118
SVCT2 AVPFLLADAMCVGYDQWATSQLIGTIFFCVGITTLLQ-TTFGC-RLPLFQASAFAFLAPA     176
        *.::       .        :      *: :::           : * : .*..  :
         TM3              TM4
UapA  -----------------MGVSFSIISVASGAFNQMYSNGFCQLDEAGN-RLP-CPEAYGA     191
UraA  LLLLPLGYEVALGGFIMCGVLFCLVS----FIVKKAGTGWLDVLF--------------     120
SVCT1 KAILALERWKC-----------PPE------EEIYGN-WSLPLNTSHIWHPRIRDVQGA     159
SVCT2 RAILSLDKWKC-----------NTT------DVSVANGTAELLHTEHIWYPRIREIQGA     218
                                          ..
                                   TM5
UapA  LIGTSACCALVEILLA-FVPPKVIQKIFPPIVTGPTVMLIGISLIGTGFKDWAGGSACMD     250
UraA  ---------------------------PPAAMGAIVAVIGLELAGVAAGM---------     143
SVCT1 IM----VSSVVEVVIGLLGLPGALLNYIGPLTVTPTVSLIGLSVFQ-------------     201
SVCT2 II----MSSLIEVVIGLLGLPGALLKYIGPLTITPTVALIGLSGFQ-------------     260
                            *  .      *  :**:.
                             TM6
UapA  DGMLCPSATAPRPLPWGSPEFIGLGFLVFVS-IILCERFG--------------APIMKS     295
UraA  -AGLLPAEGQTPDSKTIIISITTLAVTVLGS-------------------VLFRGFLAI     182
SVCT1 ------AAGDRAGSHWG-ISACSILLIILFSQYLRNLTFLLPVYRWGKGLTLLRIQIFKM     254
SVCT2 ------AAGERAGKHWG-IAMLTIFLVLLFSQYARNVKFPLPIYKSKKGWTAYKLQLFKM     313
           :  . :: *                                       ::
         TM7
UapA  CSVVIGLLVGCIVA-------------AACGYFSHADIDAAPAASFIWVKTFPLSVYGP-     341
UraA  IPILIGVLVGYALSFAMGIVDTTP----------------IINAHWFALPTLYT---P     221
SVCT1 FPIMLAIMTVWLLCYVLTLTEVLPTDPKAYGFQARTDARGDIMAIAPWIRIPYPCQWGLP     314
SVCT2 FPIILAILVSWLLCFIFTVTDVFPPDSTKYGFYARTDARQGVLLVAPWFKVPYPFQWGLP     373
        :::.::.    :.                                    *.
```

```
                     TM8                                              TM9
        _____                              _____
UapA    -----MVLPIIAVFIICAGECIGDVTATCDVSRLEVRGGTFESRIQGAVLADGINSVVAA    396
UraA    RFEWFAILTILPAALVVIAEHVGHLVVTANIVKK---DLLRDPGLHRSMFANGLSTVISG    278
SVCT1   TVTAAAVLGMFSATLAGIIESIGDYYACARLAGA---PPPPVHAINRGIFTEGICCIIAG    371
SVCT2   TVSAAGVIGMLSAVVASIIESIGDIYYACARLSCA---PPPPIHAINRGIFVEGLSCVLDG    430
             ::  :: :   *: :*:     · · :        :: .::..:*:   :: .

                         TM10                  TM11
        _____            _____
UapA    LATMTPM-TTFAQNNVVIALTRCANRWAGYCCCLILIVAGIFAKFAAAIVAIPNSVMGGM    455
UraA    FFGSTPN-TTYGENIGVMAITRVYSTWVIGGAAIFAILLSCVGKLAAAIQMIPLPVMGGV    337
SVCT1   LLGTGNGSTSSSPNIGVLGITKVGSRRVVQYGAAIMLVLGTIGKFTALFASLPDPILGGM    431
SVCT2   IFGTGNGSTSSSPNIGVLGITKVGSRRVIQCGAALMLALGMIGKFSALFASLPDPVLGAL    490
         :     *:    *: *: *:..*:   · ·       · ..*::* :   :*  ::*.:

            TM12                      TM13
        _____          _____
UapA    KTFLFASVVISGQAIVAKA--P-FTRRNRFILTASMALGYGATLVPTWFGNVFPQTENRD    512
UraA    SLLLYGVIGASGIRVLIESKVDYNKAQNLILTSVILIIGVSG---AK---------VNIG    385
SVCT1   FCTLFGMITAVGLSNLQF--VDMNSSRNLFVLGFSMFFGLTL---PNYL-SPNPGAINTG    485
SVCT2   FCTLFGMITAVGLSNLQF--IDLNSSRNLFVLGFSIFFGLVL---PSYL-RQNP--LVTG    542
          *:. :   *    :         . :* ::     : :*        .         .

                     TM14
        _____
UapA    LEGFENAIELVLETGFAVTAFVAMLLNAIMPAEVEEIGAVTPMPVSAHDNRDGE---AEY    569
UraA    AAELKGM---ALAT--IVGIGLSLIFKLISVLRPEEVVLDAED-ADITDK----------    429
SVCT1   ILEVDQILIVLLTTEMFVGGCLAFILDNTVPGSPEERGLIQWK-AGAHANSDMSSSLKSY    544
SVCT2   ITGIDQVLNVLLTTAMFVGGCVAFILDNTIPGTPEERGIRKWK-KGVGKGNKSLDGMESY    601
         ..        * *    *     :::::.        **         .

UapA    QSKQA-------------------------------------------------    574
UraA    ---------------------------------------------------    429
SVCT1   DFPIGMGIVKRITFLKYIPICPVFKGFSSSSKDQIAIPEDTPENTETASVCTKV    598
SVCT2   NLPFGMNIIKKYRCFSYLPISPTFVGYTWKGLRKSDNSRSSDEDSQATG-----    650
```

Figure 73: Sequence alignments of human SLC23 ascorbate transporters SVCT1 and SVCT2 and other SLC23 transporters. (A) Phylogenetic tree based on a sequence alignment using Clustal Ω of human SVCT1 and SVCT2 and SLC23 transporters from *E. coli* for which functional specificity was determined. Respective substrates are indicated in italics on the right. (B) Clustal Ω sequence alignment of *E. coli* UraA, *A. nidulans* UapA and human SVCT1 and SVCT2. Transmembrane helices as well as the short ß-sheet preceding the discontinuous helices in the substrate binding site according to the UraA structure are indicated. Residues located in the substrate binding sites of UraA and UapA are represented by bold letters, the charged/polar residues in TM8 and TM10 (E/H/E in UraA, E/D/Q in UapA and E/D/P in both SVCT1 and SVCT2) additionally are marked with dashed boxes.

The conserved glutamate of TM8 is also found in the human SVCT1 and SVCT2, in combination with an aspartate in the vicinity, resembling the UapA binding site (**Figure 73B**). Notably, on TM10, where commonly glutamate is found in uracil transporters and glutamine in those specific for xanthine, they harbor a proline residue instead. So far, only one study addresses the question how ascorbate transporters might have evolved from ancestral nucleobase transporters, based on the well-characterized xanthine-uric acid transporter UapA. However, no mutant of UapA, changing single or multiple amino acids to the ones found in humans SVCTs, could be identified that would bind or transport ascorbate with any pronounced affinity (Kourkoulou *et al.*, 2018). Based on competition assays, it was proposed in contrast that wild type UapA is able to bind ascorbate with a very low affinity (Ki ~ 17 mM). Uric acid and ascorbate serve similar physiological functions, namely in acting as radical scavengers and antioxidants. It was therefore suggested that SLC23 transporters evolved from nucleobase to ascorbate transporters due to the increased exposure to oxygen, when life started to populate solid ground (Liang *et al.*, 2001). How this evolutional step happened on a molecular basis,

however, is not yet understood. While nucleobases are planar and poorly water-soluble, ascorbate is non-planar and readily soluble in water (Kourkoulou *et al.*, 2018). Additionally, it is not clear yet why ascorbate transport is sodium-coupled. In general, coupling of substrate transport to a secondary gradient allows accumulation of the substrate against its gradient. Vitamin C and nucleobases exhibit very different functions inside the cell; most importantly, the majority of organisms do not rely on nucleobase uptake from the environment as nucleobases can also be synthesized *de novo*. Additionally, many nucleotides can be converted into one another. Many higher eukaryotes, especially primates, cannot produce vitamin C due to loss of the L-gulono-γ-lactone oxidase (Liang *et al.*, 2001). Expression of SVCT2 was found in most human tissues tested, and *svct2* knockout mice die shortly after birth (Bürzle *et al.*, 2013). While vitamin C can also be metabolized to other compounds, but it also serves physiological functions in itself. For example, ascorbate was shown to inhibit glucose transport in GLUT1, that is commonly upregulated in cancer cells to promote high rates of glycolysis (Reang *et al.*, 2021). Free nucleobases, on the other hand, are mainly degraded or incorporated into nucleotides. From this point of view, there is no physiological need to couple their transport to a secondary gradient. Sodium coupling that allows for accumulation of vitamin C in contrast might therefore be a result of the high physiological importance of this solute.

Discussion

5.4. Cooperativity in dimers of the SLC23 and SLC26 family

To date, all proteins that have been suggested to facilitate alternating access to the two sides of the membrane by an elevator transport mechanism are found to form oligomers, in particular di- or trimers (Garaeva and Slotboom, 2020; Holzhüter and Geertsma, 2020). At the same time, every protomer within the oligomer harbors a complete substrate binding site and translocation pathway. This apparent contradiction suggests a biological relevance of the oligomeric state for the transport mechanism. Herein, in a combination of *in vivo* and *in vitro* transport assays the cooperative behavior between the protomers in SLC23 and SLC26 dimers was investigated (sections 4.4 and 4.5).

The oligomeric state of SLC26 proteins has been investigated in several different studies using a variety of direct and indirect biochemical and biophysical approaches. Detro-Dassen *et al.* observed dimer formation using unspecific crosslinking and BN-PAGE for human SLC26A3, prestin from rat and zebrafish (SLC26A5), and a SLC26 homolog from *Pseudomonas aeruginosa*, also suggesting a conservation of the oligomeric state in different distantly related species (Detro-Dassen *et al.*, 2008). In contrast, using sequential bleaching of fluorophore-coupled protein in isolated membranes of HEK293 cells, which poses a more native-like environment than in the detergent-solubilized state, tetramers were observed for a number of SLC26 family members (prestin from gerbil, rat, and zebrafish, as well as human SLC26A4, SLC26A9, and SLC26A11) (Hallworth *et al.*, 2013). In another study, pendrin (SLC26A4) was described as mostly monomeric based on sucrose gradient sedimentation (Porra *et al.*, 2002). It was suggested that the cause for this different observations lies in only weak interactions between protomers of SLC26 oligomers that might be easily disrupted when the protein is moved out of its native environment (Hallworth *et al.*, 2013). The first structure of a SLC26 family member, the one of SLC26Dg from *Deinococcus geothermalis*, was also solved on the protein in a monomeric state in detergent (Geertsma *et al.*, 2015). Nevertheless, there were indications that SLC26Dg forms a dimer in the lipid environment, and later on, a model of the transmembrane domain of dimeric SLC26Dg reconstituted to proteoliposomes based on EPR constraints was published (Chang *et al.*, 2019). Since then, two other full-length cryo-EM structures of SLC26 proteins became available, of mouse SLC26a9 and its human homolog (Walter *et al.*, 2019; Chi *et al.*, 2020). For both, structures of dimers were obtained on protein solubilized in detergent. For the experiments described herein, it was assumed that SLC26Dg and SLC26Si form dimers when reconstituted to soy PC, as there were no indications for another (higher) oligomeric state found.

For the SLC23 family, the oligomeric state has not been addressed in detail so far. UraA has been crystallized both as monomer and dimer, and the oligomeric state seems to depend on the detergent used (Lu *et al.*, 2011; Yu *et al.*, 2017). UapA also crystallized as a homodimer. Human SVCT1 was found

162

Discussion

to be mostly monomeric on Western blot when purified using Triton X-100, but dimers were detected upon chemical crosslinking of the overexpressed protein in oocyte membranes (Boggavarapu *et al.*, 2013). For UraA, Yu *et al.* suggest a monomer-dimer equilibrium in the detergent-solubilized state based on the observation of two peaks in size-exclusion chromatography (Yu *et al.*, 2017). As they find in *in vivo* uracil uptake assays the activity of a concatemer of UraA, i.e., a constitutive dimer, higher than for the wild type protein, they suggest a similar monomer-dimer equilibrium in the native membrane. However, as discussed in detail in this thesis (section 4.2.3), the *in vivo* based uptake assays have significant drawbacks when it comes to determining precise activities of protein, such as the high background uracil uptake. Furthermore, Yu *et al.* use cells carrying an empty expression vector as negative control. We show that membrane protein overexpression has a significant impact on the uracil uptake background (section 4.2.3), i.e., the activity of cells overexpressing UraA will likely be overestimated when comparing their activity to the one of cells not overexpressing any protein. Uracil uptake in cells is, as outlined in section 4.2.7, critically depending on the activity of UPRT, the downstream enzyme catalyzing the conversion of uracil to UMP. The apparent uracil uptake rate therefore represents the sum of UraA uptake rate and the enzymatic activity of UPRT. At the moment it is unclear how this is reflected in the apparent uracil uptake rate in a whole cell setting.

Here, specific crosslinking in membrane vesicles and unspecific crosslinking in proteoliposomes showed extensive dimerization of UraA (section 4.4.1). However, chemical crosslinking as performed here rather represents an endpoint measurement and cannot report on a dynamic equilibrium. The homolog 51ThXi, on the other hand, appears to be exclusively monomeric in decyl-ß-D-maltoside as assessed by SEC-MALS (**Figure S27**). Also for this protein, dimerization was observed using specific cysteine crosslinking mutants in membrane vesicles and after purification and reconstitution to proteoliposomes (section 4.4.1). A concatemeric version of 51ThXi overexpressed in *E. coli* BW25113(*ΔuraA*) showed similar activity as the wild type protein (section 4.4.3). Therefore, it can be assumed that the protein is a dimer in the lipid environment.

So far, cooperativity in elevator proteins, or in membrane proteins in general, has only been investigated to a limited extent. Broadly speaking, a system of subunits behaves in a cooperative manner, when the action of the system as a whole differs from the sum of the action of the single subunits. Cooperativity by means of interprotomer crosstalk within the oligomers of elevator proteins has recently been reviewed by us (Holzhüter and Geertsma, 2020). In summary, no general conclusion can be drawn for cooperative behavior in elevator proteins; for some, evidence for no functional crosstalk between protomers was provided, while for others positive or negative cooperativity was reported. For SLC26Dg, functional dependency of the two protomers was investigated by mixing wild type protein with an immobilized version in different ratios (Chang *et al.*, 2019). The activity that was

obtained was higher than expected from independent activity of the single subunits, suggesting a positive interaction of the two protomers. Herein, in a similar experiment performed on the homolog of SLC26Dg from *Sulfitobacter indolifex*, termed SLC26Si, a lower activity of the wild type protein was recorded when mixing with a similar immobilized version (section 4.5.3). A distinct negative dominant effect was also observed for heterodimers of 51ThXi, in which one of the protomers was the E256A mutant that was demonstrated to being impaired in transport (section 4.4.3). This experiment was not only performed on reconstituted 51ThXi, but also using the concatemeric version in a cell-based uracil uptake assay. This experiment confirmed the negative dominance of the inactive over the wild type protomer (section 4.4.3). For SLC26Dg, the higher activity of the wild type protomer when paired with the inward-locked version was explained in such a way that the immobilized mutant serves as an extended scaffold that facilitates a more efficient movement of the substrate binding domain against it. Analogous to that, the inward-locked SLC26Si variant might cause a less efficient movement of the active protomer. The activity of the inward-locked SLC26Si mutant was fully reversible in the presence of reducing agent, indicating that the cysteine mutations by itself do not interfere with the transport activity. As no structure of SLC26Si and its dimer interface is available, it appears not feasible to speculate on the molecular basis of the apparent difference in cooperative behavior of SLC26Dg and SLC26Si. This result, however, demonstrates that mechanistic differences also might exist within the same SLC family.

The apparent negative cooperativity in 51ThXi is less straightforward to explain, as the precise effect of the E256A mutant that was used as an "inactive mutant" is not entirely clear. The experiment using the constitutive heterodimer of 51ThXi composed of one wild type protein and one E256A mutant corresponds to the one performed on UraA by Yu *et al.* (Yu *et al.*, 2017). The outcomes yet are in opposition to each other, as in UraA the heterodimer showed a similar activity as the concatemer composed of two wild type protomers. For 51ThXi, on the other side, the heterodimer was as inactive as the one of two inactive mutants. The same result was observed in the proteoliposome-based assay where wild type 51ThXi and E256A were mixed in different ratios. This might again indicate that mechanistic differences exist within the SLC23 family. For UapA, co-expression experiments of the wild type transport with several transport compromised mutants indicated a similar dominant negative effect of the mutant on the transport activity of the wild type (Alguel *et al.*, 2016). In this case, it is as well not entirely clear what effect the mutations have on the transport mechanism. Notably, one of the mutants was Q408E, which retains the ability for substrate binding with an affinity similar to wild type UapA, but appears not to transport. This relates to some extent to the E256A mutant of 51ThXi, which also is able to bind uracil, but exhibited no (*in vitro*) or only low (*in vivo*) transport activity. It is tempting to speculate that the E256A mutation would cause a similar immobilization of the protein as

the disulfide bridge in the inward-locked versions of SLC26Dg and SLC26Si. No direct experimental evidence for this, though, can be provided at the moment. Although non-mammalian SLC23 transporters appear not to be proton-coupled, it cannot be excluded that the negative charge provided by E256 is relevant for high-affinity substrate binding and efficient reorientation of the transporter. As discussed above, E256 might be relevant in neutralizing the positive dipoles caused by the discontinuous helices 3 and 10; the E256A mutant can be expressed and purified, but shows a remarkable down-shift in melting temperature indicating a destabilization of the protein. E256 might be relevant in stabilizing the discontinuous helices in a position that allows the elevator mechanism to take place. The elevator movement is envisioned as the motion of two rigid bodies against each other. Removal of E256 might cause a higher intrinsic flexibility of the substrate binding domain incompatible with the elevator motion.

For the cell-based uptake assay using the concatemer of 51ThXi, next to E256A as functionally impaired mutant, also the E256A/H260A/E303A triple mutant was used that is lacking most of the substrate binding site residues that are thought to be involved in uracil coordination. In contrast to E256A, this mutant is expected not to be able to bind the substrate. It was not possible to purify this mutant due to pronounced aggregation, which is in line with the idea that the electrostatic interactions within the binding site stabilize the substrate binding site and, subsequently, the protein in total. In the cell-based uptake assay, however, expression and folding of the triple mutant as a single protein as well as in the context of the concatemer was, judged by the fluorescence of GFP as folding indicator, not markedly compromised. This might be an indication that lipids are involved in stabilization of the 51ThXi dimer. The concatemer of 51ThXi, that exhibited wild type-like activity *in vivo*, showed absolutely no activity when reconstituted into proteoliposomes, although protein folding and substrate binding was not affected (section 4.4.3). This might also indicate that lipid composition or membrane curvature are mechanistically relevant. To promote transport, the dimer interface might need a certain degree of flexibility, that is (too) restrained in the reconstituted concatemer. This also fits to the finding that the stapling mutants, in which the scaffold domains were linked by disulfide bridges, were also impaired in transport (section 4.4.2). For the elevator protein vcINDY, a reduction in transport activity was observed when inter-protomer crosslinks were introduced at certain positions at the dimer interface, while for other positions, transport activity was not impaired (Mulligan *et al.*, 2016). This also indicates that local distortions of the dimer interface can affect an efficient movement of the transport domain against the scaffold domain.

The membrane environment may also suggested provide the basis for interprotomer crosstalk leading to cooperative behavior within elevator proteins (Holzhüter and Geertsma, 2020). The elevator movement, that is characterized by a relatively large movement of the substrate binding domain,

would place a significant part of the hydrophobic protein surface out of the lipidic environment, which would probably result in a large energetic penalty. On a rational basis, it is much more likely that the elevator movement is accompanied by a membrane deformation that keeps the protein within the hydrophobic environment. Recent MD simulations on Glt_{Ph} and vcINDY, two well-characterized elevator proteins of different families, showed indeed that the transition from the outward- to the inward-open conformation leads to large deformations of the membrane around the substrate binding domains (Zhou *et al.*, 2019). Experimental evidence for this was provided by structures in different conformations of Glt_{Tk}, a homolog of Glt_{Ph}, reconstituted to nanodiscs (Arkhipova *et al.*, 2020). Detailed smFRET and HS-AFM studies on Glt_{Ph}, however, consistently showed that the protomers move independently from each other, i.e., no cooperativity exists within the trimer of Glt_{Ph}.(Erkens *et al.*, 2013; Ruan *et al.*, 2017). The MD simulations suggested that the membrane perturbations caused by the conformational change of Glt_{Ph} are restricted to the substrate binding domain, meaning, the membrane deflections ceased beyond the transport domain. Conceivably, in a case where the movement of the substrate binding domain would include a membrane deformation at the level of the scaffold domain, this effect might be transferred to the scaffold domain of the adjacent protomer, providing a molecular basis for interprotomer cross-talk.

The negative effect on the activity of the wild type protomer paired with an inactive mutant, observed for both 51ThXi and SLC26Si, indicate that for both proteins, both protomers need to be active to promote efficient transport. In SLC26Dg, on the other hand, active transport was also possible when one protomer was immobilized (Chang *et al.*, 2019). At the moment it is unknown whether a monomeric version of SLC26Dg would also facilitate transport. Monomerized UraA was shown to be largely inactive (Yu *et al.*, 2017). Monomeric NhaA on the other hand retained its function (Rimon *et al.*, 2007). So far, the scaffold domain was suggested to remain mostly static during the transport cycle, acting mostly as a membrane anchor in order to allow efficient movement of the substrate binding domain against it. It remains to be elucidated whether oligomerization of elevator proteins serves a more structural role – in such a way that the interaction of several scaffold domains leads to a more stable membrane anchor – or a more functional role by providing a platform for cooperativity between protomers, or both.

5.5. The role of the STAS domain for transport in the SLC26 family

The SLC23 and SLC26 families share the same 7TMIR architecture of the transmembrane part that is harboring the substrate binding site and the translocation pathway. The SLC26 proteins contain an additional C-terminal, cytosolic domain, the so-called STAS domain. Removal of the STAS domain causes a drastic reduction in transport activity, while the affinity for the substrate remains unchanged, as seen for both SLC26Dg and SLC26Si (Geertsma *et al.*, 2015, and Chang, Geertsma, unpublished). The STAS domain is not confined to the SLC26 family, but in all kingdoms of life, soluble STAS domain-only proteins can be found as well as other STAS-fusion proteins (Sharma *et al.*, 2011). Based on the different functions of these STAS domains, different functions have been suggested for the STAS domain of SLC26 transporters as well, but experimental data is scarce.

The full-length structures of the SLC26A9 homologs from human and mouse revealed the orientation of the STAS domain with respect to the transmembrane domain in these two dimeric proteins: the two protomers intersect, placing the STAS domain of one protomer close to the transmembrane domain of the other protomer, and *vice versa*. Herein, it was demonstrated that the overall architecture of SLC26Dg is similar (section 4.5.1). For both SLC26Dg and SLC26Si, the interaction of the STAS domain with the transmembrane domain was suggested to be specific, as chimeras of both proteins (SLC26Dg carrying the STAS domain of SLC26Si, and the other way around) showed little activity (Chang, Geertsma, unpublished). A similar effect was described before for prestin (SLC26A5) in which the STAS domain was exchanged for the one of other closely related homologs (SLC26A4 and SLC26A6) (Zheng *et al.*, 2005). For the three *Arabidopsis thaliana* sulfate transporters, one chimeric construct was active, the other two not (Shibagaki and Grossman, 2004).
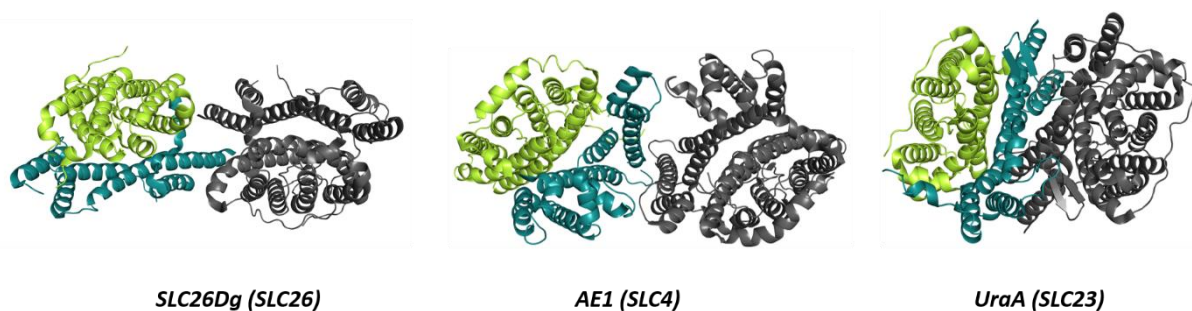

Figure 74. Dimer interfaces of 7TMIR proteins. The buried area at the interface of SLC26Dg (left) is smaller compared to the ones of SLC4 (middle) and SLC23 (right). For each family, a representative member is shown. Proteins are shown in top view, with one protomer colored in grey, the other with the substrate binding domain in light green and the scaffold domain in teal. Only the transmembrane part of SLC26Dg and AE1 are shown.

It appears that the STAS domain is not a general prerequisite for dimerization, as the STAS-less proteins of SLC26Dg and SLC26Si also form dimers (Chang *et al.*, 2019). As already mentioned, the dimer

interface within the transmembrane part of SLC26 is small compared to the dimer interfaces of the SLC4 and SLC23 families (buried areas: ~350 Å for SLC26, ~1000 Å for SLC4 and ~2000 Å for SLC23, respectively (**Figure 74**)). The overall dimer architecture of the SLC26 proteins suggest that the STAS domains might serve in extending the dimer interface outside of the membrane. Supporting this hypothesis, in the structures of both SLC26A9 homologs, an extensive interface between the two STAS domains can be identified, apparently also including electrostatic interactions (**Figure 75**). However, the STAS domains of several SLC26 transporters on its own were reported to being monomeric in solution (Pasqualetto *et al.*, 2010; Compton *et al.*, 2011; Bai *et al.*, 2016). This suggests that the affinity of the two STAS domains for each other might be low. Their fusion to the transmembrane domain places them in proximity to each other, promoting their dimerization. This is as well reflected in the finding that the composition of the linker that connects transmembrane and STAS domain is relevant for transport activity (section 4.5.2): a correct positioning of the STAS domain appears to be crucial in order to allow an efficient movement of the substrate binding domain against the scaffold domain. In the mutants in which the residues of the linker are either glycine or serine, thereby creating a more flexible connection of the STAS domain to TM14, transport activity is lost (**Figure 63**). As judged by the PELDOR measurements, however, the distances between the transmembrane and STAS domain change only to a minor extent when the linker is mutated. This indicates that the relative orientation of the STAS domain to the transmembrane domain is mostly preserved. As the linker sequence is not well-conserved among SLC26 transporters, it is difficult to predict whether similar mutations in other transporters would cause a similar effect on activity. Additionally, without structural information, the precise borders of the linker can be difficult to determine. Overall, as SLC26 transporters in general have various functions, the STAS domain might also be of different importance for different proteins. The small interface within the transmembrane domain, formed mostly by interactions of TM14, on the other hand appears to be conserved from bacteria (as in SLC26Dg) to higher eukaryotes. Currently, we assume that the transmembrane domains of the SLC4, SLC23 and SLC26 families evolved from one common ancestor. The additional N-terminal domain in the SLC4 and the additional C-terminal domain in the SLC26 family probably originated from gene fusion events. These kind of gene fusions are believed to be a major driving force of evolution by creating proteins with novel functions by combining already existing, fully functional ones (Michael, 2017). That STAS domains can also be found in other fusion proteins (Sharma *et al.*, 2011) and also exist on their own (recently reviewed in Moy and Seshu, 2021), supports this hypothesis. Subsequently, the SLC26 proteins evolved in such a way that the dimer interface between the STAS domains gained importance, while the one within the transmembrane domains, that still is responsible for dimerization in the SLC4 and SLC23 families, got reduced. The loss of function of the domain was counterbalanced by the gain of function due to the C-terminal STAS

domain that increased the functional diversity of the SLC26 family by acting as regulator and interaction platform for other proteins.
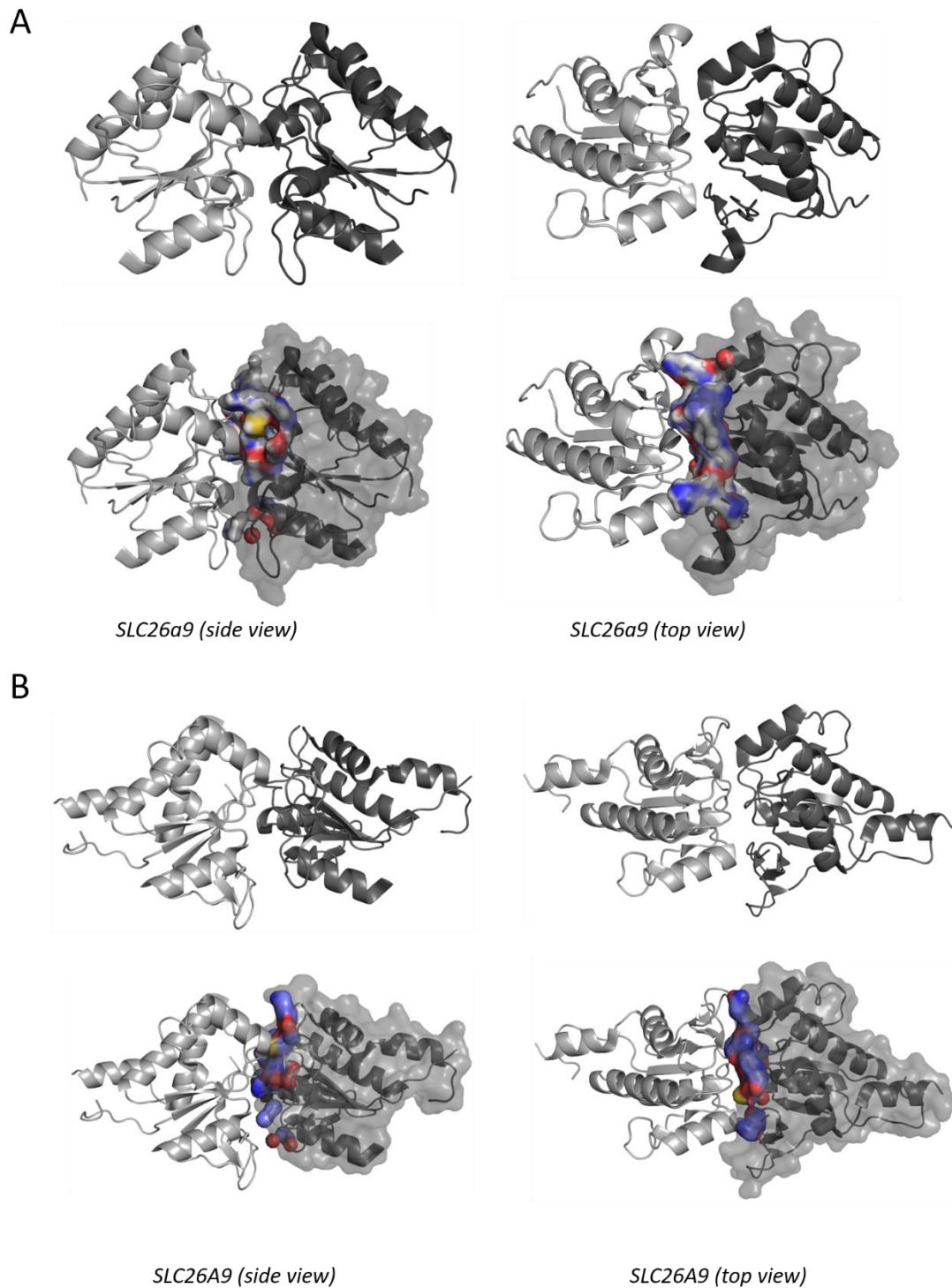
Figure 75: STAS domain interface in mouse SLC26a9 and human SLC26A9. (A) Structure of the STAS domains only as found in the full-length structure of the SLC26a9 dimer (PDB: 6RTC). In the lower panel, the surface of one STAS domain is represented with residues found within 5 Å of the other STAS domain colored in blue. (B) Same as in (A), but for human SLC26A9 (PDB: 7CH1).

169

Discussion

To date, multiple structures of isolated STAS domains of SLC26 proteins are available, e.g., for *E. coli* DauA (PDB: 3NY7), rat prestin (PDB: 3LLO) or chicken prestin (PDB: 5EZB). All of them crystallized as monomers. Although they share low to moderate sequence identity with SLC26a9 (20.3% for DauA, 39.4% for rat prestin and 41.3% for chicken prestin STAS domains), their structures can be superimposed with low RMSD (3.8 Å over 96 residues for rat prestin, 1.5 Å over 136 residues for chicken prestin STAS domains (**Figure 76A**)). Remarkably, although no sequence similarity can be detected between SLC26a9 and SLC26Dg using the BlastP algorithm, the structures of the STAS domains align well with an RMSD of 3.0 Å (88 residues) (**Figure 76B**).[8]
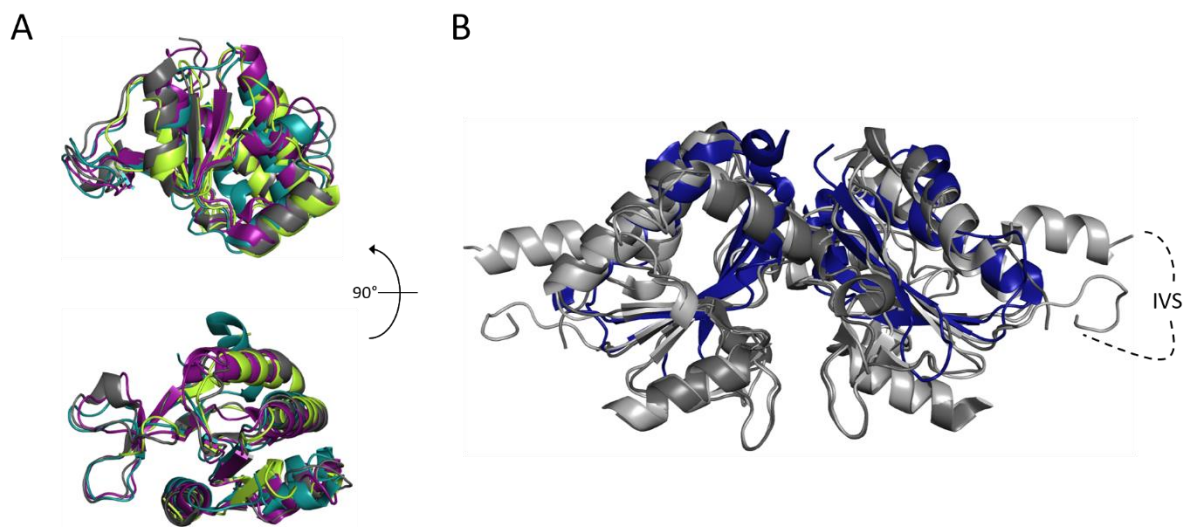
Figure 76: **Superimposition of STAS domains from SLC26 transporters.** (A) Structural alignment of STAS domains from SLC26a9 (dark grey, PDB: 6RTC), *E. coli* DauA (light green, PDB: 3NY7), rat prestin (dark teal, PDB: 3LLO) and chicken prestin (purple, PDB: 5EZB). The proteins were aligned using the cealign tool of PyMOL and are shown in side view on top and in top view below. (B) Superimposition of the STAS domains of mouse SLC26a9 (dark grey), human SLC26A9 (light grey, PDB: 7CH1) and SLC26Dg (dark blue, PDB: 5da0). IVS, intervening sequence.

The STAS domain of SLC26Dg is composed of a ß1-ß2-α1-ß3-α2-ß4-α3 secondary structure. In general, STAS domains from eukaryotic organisms fold in a similar way, but contain a so-called "intervening sequence" (IVS) inserted in the loop between α1 and ß3 (Sharma *et al.*, 2011). To obtain the structure of SLC26a9, the IVS was removed in order to stabilize the protein (Walter *et al.*, 2019). For the structure determination of human SLC26A9, the IVS was included, but not structurally resolved (Chi *et al.*, 2020). The localization of the intervening sequence is indicated for one protomer in **Figure 76B**. As it is on the opposite side of the STAS-STAS interface, it is not expected to influence their dimerization. The larger STAS domains of the SLC26a9 homologs are composed of a six-stranded ß-sheet, interspersed with 4 α-helices. However, as can be seen from the superimposition with the SLC26Dg STAS domain, most of the additional secondary structure elements do not contribute to the dimer interface. Within the STAS

---

[8] The overall structures of SLC26a9 and SLC26Dg align with an RMSD of 4.3 Å (368 residues).

domain dimers, only the additional forth helices are located in proximity and might provide additional stabilization of the eukaryotic SLC26 proteins compared to the prokaryotic ones.

Several disease-related mutations located on the STAS domains of human SLC26 are known. Many of them were shown to influence the correct folding of the STAS domain (Dorwart *et al.*, 2008; Pasqualetto *et al.*, 2010; Alper and Sharma, 2013; Bai *et al.*, 2016). The apparent extensive contribution of the STAS domain to the overall dimer interface of SLC26 transporters suggests that incorrect STAS domain folding might result in improper dimer formation. For many SLC26 transporters, mutations on the STAS domain apparently can be a cause for ER retention (Karniski, 2004; Shibagaki and Grossman, 2006; Dorwart *et al.*, 2008). This suggests that dimerization might be relevant for correct trafficking of SLC26 transporters to the plasma membrane. For the Sultr sulfate transporters of *A. thaliana*, certain STAS domain truncations for both Sultr1;2 and Sultr1;1 resulted in proteins that were expressed to wild type protein levels, but mostly found in the ER fraction (Shibagaki and Grossman, 2004). In the same study, chimeras of STAS and transmembrane domains of the different Sultr transporters were investigated. Fusing foreign STAS domains to the transmembrane part appeared to be able to rescue plasma membrane trafficking, while function was not restored. However, the oligomeric state of the mutants was not investigated therein. For several other transporters, oligomerization in the ER was suggested to be a signal for incorporation into cargo vesicles and ER exit, independent from specific signal sequences (Springer *et al.*, 2014). Also for the SLC23 transporter UapA, oligomerization was reported to be crucial for correct localization to the plasma membrane (Martzoukou *et al.*, 2015b). Detailed re-evaluation and systematic analyses of STAS domain mutations in future experiments, based on the now available full-length dimeric SLC26 structures, will reveal whether incorrect dimer formation caused by STAS domain misfolding can explain disease-related phenotypes described for the human SLC26 transporters.

5.6.   Concluding remarks and outlook

This study provides important insights into distinct aspects of the transport mechanism of the SLC23 and SLC26 families. Most importantly, it demonstrates a uniport transport mode for a SLC23 nucleobase transporter, using a novel and thus far unique proteoliposome-based transport assay. We conclusively show that the apparent proton-coupling observed in whole cell transport assays for uracil transporters, but probably also for SLC23 transporters of other specificities, is caused by an indirect effect of pH on uracil phosphoribosyltransferase, an enzyme catalyzing the downstream conversion of uracil to UMP. A dependence of uptake on downstream modifying enzymes has also been described for other nucleobase transporters, and the activity of many of these enzymes has been reported to be pH-dependent. Future experiments and a re-examination of the existing data will therefore most likely reveal also proton-independent transport of other SLC23 nucleobase transporters. With this, the SLC23 transporter family holds both uncoupled as sodium-coupled transporters, a functional divergence which is similar to the one found for the structurally related SLC4 and SLC26 families.

In particular, the crucial relevance of UPRT activity for uracil uptake in *E. coli* has so far been neglected in interpreting the available transport data for UraA, despite early suggestions that UPRT activity is the limiting factor for the uracil uptake system (Andersen *et al.*, 1995). In line with that, we show that overexpression of both UraA and UPRT massively increases uracil uptake in *E. coli* (section 4.2.7). It follows that instead of actual uracil translocation, the conversion of uracil to UMP is the rate-limiting step of overall uracil uptake. Together with the high background uracil uptake in *E. coli*, this renders the *in vivo*-based uracil transport assays largely useless for most studies on activity of SLC23 uracil transporter. With the newly established *in vitro* transport assay for the SLC23 family, although at the moment it is only available for the UraA homolog 51ThXi, uracil transport now can be studied free of background and independent of downstream enzymatic activity.

Early on, it has been suggested that substrate specificity of SLC23 nucleobase transporters can be predicted on the basis of a few substrate binding site residues (Frillingos, 2012; Papakostas and Frillingos, 2012; Karena *et al.*, 2015; Chaliotis *et al.*, 2018; Kourkoulou *et al.*, 2018). However, this observation was based on the functional analysis of only a few homologs, most of them originating from *E. coli*, which show high specificity for their respective substrates. More recently, SLC23 nucleobase transporters with a higher promiscuity have been described, which phenotype cannot be attributed to the presence of certain binding site residues. Additionally, detailed studies on the two best functionally characterized SLC23 homologs (UapA of *Aspergillus nidulans* and XanQ of *E. coli*) indicate that also residues away from the substrate binding site contribute to substrate specificity. The molecular basis, and its relevance for the substrate translocation mechanism, are yet poorly

understood. To be able to *de novo* predict specificities of SLC23 transporter that are of medical relevance accordingly requires a more thorough investigation of the molecular basis of substrate recognition than is thus far available.

Another aspect of the substrate translocation mechanism of SLC23 transporter that is not well understood to date is how substrate binding promotes the conformational change. As was shown in this study, substrate transport of the uracil transporter 51ThXi happens independent of secondary gradients with re-orientation of the empty carrier as rate-limiting step (sections 4.2.5 and 4.2.6). A possible explanation lies in a stabilization of the transport domain due to substrate binding that enables a more efficient movement against the scaffold domain. Within this study, several substrate binding site residues were identified that interestingly led to a destabilization of the protein and also severely impaired transport (sections 4.3.2.1 and 4.3.2.2). Unfortunately, it was not possible to determine the effect of the mutations on affinity for the substrate. In the future, it might be interesting to determine whether a lower affinity can be correlated with a lower transport rate.

Regarding the importance of dimerization for transport in the SLC23 and SLC26 families, two main insights were obtained and extended. First, different modes of cooperativity (positive and negative) exist within both families. The impact of the lipidic environment as a basis for functional cross-talk still needs to be investigated in detail. For this, single-molecule techniques, e.g. high-speed AFM or single-molecule FRET, as has been successfully used to investigate conformational transitions in the SLC1 family, will be of asset. Also solving multiple structures of reconstituted dimers of one family member in multiple conformations might provide more insight into this. Second, evidence has been presented that a main task of the STAS domain in (prokaryotic) SLC26 homologs lies in extending the dimer interface outside of the membrane, which is crucial for proper function of the protein. It still needs to be discovered whether it serves the same role in eukaryotic SLC26 homologs, for which also other functions have been suggested, such as regulation and interaction platform for other proteins.

Supplementary data

A

B
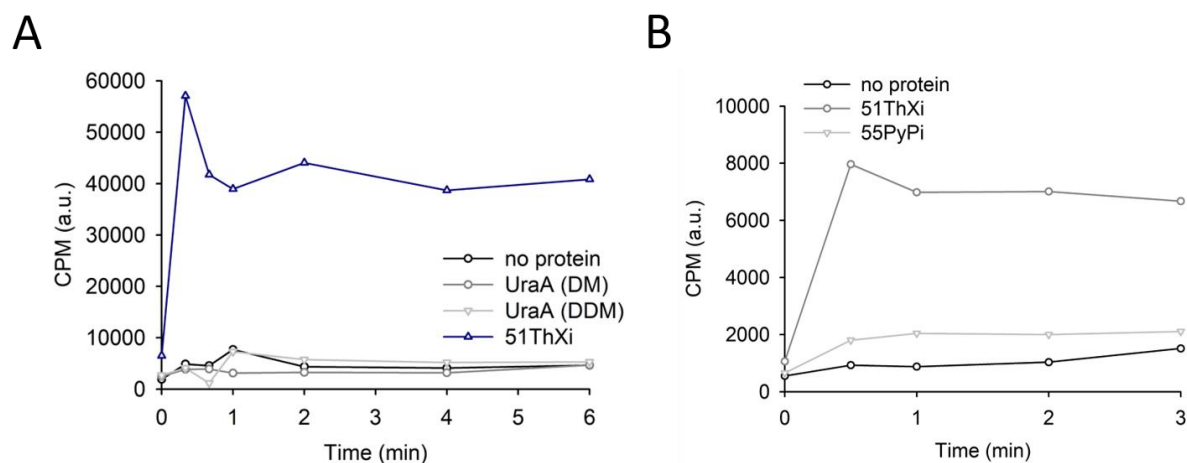
Figure S1: Initial uptake experiments on UraA, 51ThXi and 55PyPi in proteoliposomes. (A) Uptake of proteoliposomes of UraA purified using DM or DDM (as indicated) and 51ThXi compared to empty liposomes ("no protein"). Proteoliposomes were incubated with 50 nM uracil at 30°C. Inside buffer was 50 mM KPi pH 7.5, 2 mM $MgSO_4$, outside buffer 50 mM KPi pH 6.0, 2 mM $MgSO_4$. (B) Uptake of proteoliposomes of 55PyPi and 51ThXi compared to empty liposomes ("no protein"). Proteoliposomes were incubated with 25 nM uracil at 30°C. Inside buffer was 50 mM KPi pH 7.5, 2 mM $MgSO_4$, outside buffer 50 mM KPi pH 6.0, 2 mM $MgSO_4$.
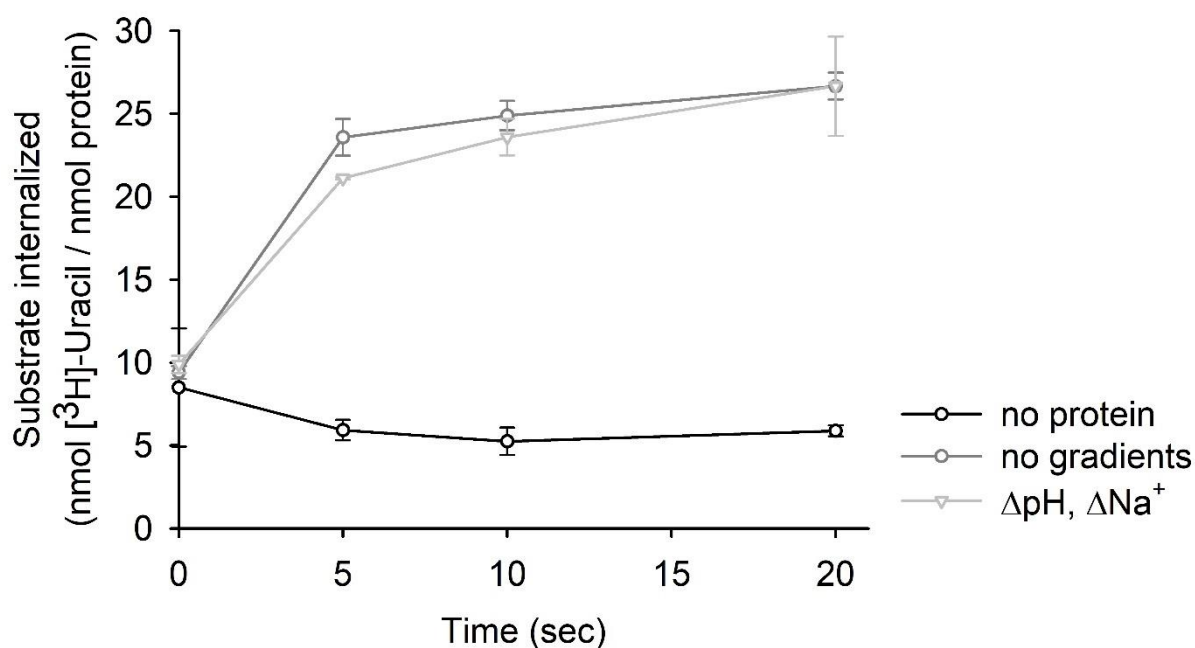
Figure S2: **Uracil uptake in proteoliposomes of 51ThXi at 20°C**. Uptake of 1 µM uracil was measured at 20°C for liposomes containing no protein (black), for proteoliposomes of 51ThXi without additional gradients (dark grey) (KPi pH 7.5 inside/outside) and applying both pH- and sodium-gradients (light grey) (KPi pH 7.5 inside/NaPi pH 6.0 outside). Uptakes were determined in triplicates, errors shown represent standard deviations.
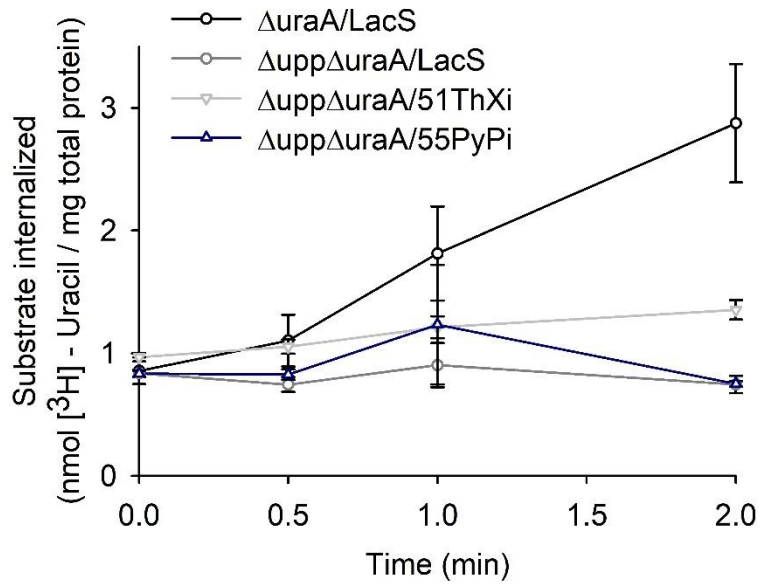
Figure S3: **Uracil uptake of 51ThXi and 55PyPi in *E. coli* BW25113(ΔuppΔuraA).** Uptake by both 51ThXi and 55PyPi in *E. coli* BW25113(*ΔuppΔuraA*) is reduced to levels below uracil uptake in *E. coli* BW25113*(ΔuraA)* overexpressing LacS. Under identical conditions, uracil uptake by 51ThXi and 55PyPi in *E. coli* BW25113(*ΔuraA*) gives rise to significant uptake over the background in cells overexpressing LacS.

Table 30: Co-localization of putative genes coding for homologs of *E. coli* UPRT with genes of UraA homologs. Homologous genes of *E. coli upp* (coding for UPRT) were identified in a number of bacterial and archaeal genomes using the blastp suite (Altschul *et al.*, 1997) and their locations were compared to the ones of the genes of previously identified UraA homologs used in this study (section 4.1). Close proximity of the genes can be an indication of a common operon, as found for UPRT and UraA of *E. coli*. For some organisms, no homologous *upp* genes were identified. For all proteins, the NCBI protein IDs are given. Identities to *E. coli* UPRT are based on Clustal Ω (Sievers and Higgins, 2018).

Organism	Homolog	Gene location	Protein ID	UPRT gene location	Protein ID	Identity to *E. coli* UPRT	mapping?
Thermococcus sibiricus DSM 12597	01ThSi	1128919..1130202	506329793	1130209..1130907	WP_015849513.1	33.98	yes
Methanococcus maripaludis DSM 14266	02MeMa	876529..877770	500195449	877807..878502	WP_012193013.1	32.69	yes
Desulfosporosinus orientis DSM 765	03DeOr	1408099..1409448	503949848	5764223..5764852	AET70904.1	48.79	no
Desulfosporosinus orientis DSM 765	04DeOr	4932124..4933347	503952955				
Desulfosporosinus orientis DSM 765	05DeOr	2453994..2455343	357213331				
Halomonas elongata DSM 2581	06HaEl	2328245..2329486	503097935	2337740..2338369	WP_013332747.1	65.07	yes
Halomonas elongata DSM 2581	07HaEl	85985..87331	503096019				
Halomonas elongata DSM 2581	08HaEl	3837826..3839241	503099227				
Ilyobacter polytropus DSM 2926	09IlPo	1462036..1463259	503153126	1534186..1534812	ADO83197.1	55.77	no
Ilyobacter polytropus DSM 2926	10IlPo	137567..138913	*503151929*				
Ilyobacter polytropus DSM 2926	11IlPo	1189971..1191302	503152888				
Proteus mirabilis DSM 4479	12PrMi	3147454..3148845	739027515				
Proteus mirabilis DSM 4479	13PrMi	1672052..1673353	739013753	1673462..1674088	WP_107034590.1	88.46	yes
Shimwellia blattae DSM 4481	14ShBl		488373146				
Shimwellia blattae DSM 4481	15ShBl	1091496..1092776	*488371914*	1090795..1091421	AFJ46172.1	95.19	yes
Thermotoga thermarum DSM 5069	16ThTh	681421..682641	503697910	766931..767560		50.00	no
Desulfotomaculum kuznetsovii DSM 6115	17DeKu	742225..743610	333821636				
Thermanaerovibrio acidaminovorans DMS 6589	18ThAc	441697..443091	502632588	983915..985027	YP_003317456.1	51.69	no
Thermanaerovibrio acidaminovorans DMS 6589	19ThAc	176437..177792	502632329				
Thermanaerovibrio acidaminovorans DMS 6589	20Thac	453672..455048	502632596				
Halanaerobium saccharolyticum DSM 6643	21HaSa	247727..249025	460789235	246923..247477	CAUI01000005.1	51.24	yes
Desulfomonile tiedjei DSM 6799	23DeTi		390623458				
Roseobacter litoralis DMS 6996	24RoLi	334421..335839	503726220	488311..488943	WP_013960429.1	52.22	no
Treponema caldarium corrig DSM 7334	25TrCa	672549..673802	338833593	1692631..1693692	AEJ19639.1	35.42	

Organism	Homolog	Gene location	Protein ID	UPRT gene location	Protein ID	Identity to *E. coli* UPRT	mapping?
Marinobacter hydrocarbonoclasticus DMS 8798	26MaHy	1372580..1373836	500109449	1373839..1374477	NC_017067.1	69.71	yes
Syntrophobacter fumaroxidans DMS 10017	27SyFu	1492944..1494254	500017353	1312218..1312778	WP_011697929.1		
Ruegeria lacuscaerulensis DMS 11314	28RuLa	226833..228260	492824019	2062305..2062937	EEX10121.1	54.19	no
Oceanicola granulosus DMS 15982	29OcGr	1316797..1318230	494465259		WP_007256093.1	52.22	no
Roseobacter denitrificans DMS 7001	30RoDe	4022011..4023429	*109457388*	3835792..3836424	ABG33430.1	52.22	no
Yersinia enterocolitica DMS 4780	31YeEn	1268234..1269523	500139814	1267515..1268141	YP_001005456.1	93.75	yes
Yersinia enterocolitica DMS 4780	32YeEn	49933..51318	*491318058*				
Yersinia enterocolitica DMS 4780	33YeEn	2124441..2125778	491312021				
Dinoroseobacter shibae DMS 16493	34DiSh	3366483..3367907	501131016	839244..839876	ABV92576.1	55.17	no
Shewanella woodyi DSM 12036	35ShWo	2655252..2656499	501281772	3359593..3360219	WP_012325355.1	83.17	
Natrialba asiatica DSM 12278	36NaAs	226210..227610	493048647	225470..226015	WP_006167680.1		yes
Natrialba asiatica DSM 12278	37NaAs	155667..157250	493048526	333551..334228	ELY98152.1	33.33	
Brachyspira murdochii DSM 12563	38BrMu		502880074				
Octadecabacter arcticus DSM 13978	39OcAr	572141..573556	505306917	3517779..3518417	WP_015496313.1	53.40	
Desulfotomaculum carboxydivorans DSM 14880	40DeCa	888433..889704	503575783	889864..890502	AEF93688.1	40.49	yes
Marinithermus hydrothermalis DSM 14884	41MaHy	1772536..1773768	503469898	1595580..1596206	AEB12315.1	49.04	no
Marinithermus hydrothermalis DSM 14884				1415338..1415874	CP002630.1		
Dialister invisus DSM 15470	42DiIn	1161845..1163125	260403711	943550..944173	WP_007069945	51.66	no
Dialister invisus DSM 15470	43DiIn	1217295..1218659	260403769				
Marinomonas ushuaiensis DSM 15871	44MaUs	15314..16555 contig12	738207174	168682..169338 contig6	ETX10942.1	50.97	no
Clostridium asparagiforme DSM 15981	45ClAs	804386..805759 scfld 5	225045503	379173..380453 Sc5			
Clostridium asparagiforme DSM 15981	46ClAs	1194855..1196207 scfld6	494987778	135363..135716 Sc6			
Clostridium asparagiforme DSM 15981	47ClAs		494980145				
Clostridium asparagiforme DSM 15981	48ClAs		*494988505*				
Clostridium asparagiforme DSM 15981	49ClAs		*225044682*				
Thermovirga lienii DSM 17291	50ThLi	837028..838443	355397094	1115449..1116558	AER66791.1	50.74	no
				1159855..1160397			

| Organism	Homolog	Gene location	Protein ID	UPRT gene location	Protein ID	Identity to *E. coli* UPRT	mapping?
Thalassospira xiamenensis DSM 17429	51ThXi	3882797..3884092	494152249	2452665..2453315	AJD52404.1	49.76	no
Selenomonas flueggei DSM 19720	52SeFl	474911..476227 Sc1	238884962	52109..52765 Sc3	WP_006690928	51.49	
Mitsuokella multacida corrig DSM 20544	53MiMu	680355..681722 Sc0	492423604	41250..41888 Sc3	EEX67749	50.72	
	54MiMu	333057..334343 Sc1	*492422708*				
Pyramidobacter piscolens DSM 21147	55PyPi	15701..17065 contig00027	750258270	8633..9730		50.00	
Anaerobaculum hydrogeniformans DSM 22491	56AnHy		*496493878*	1710778..1711875	WP_009202543.1	50.24	
	57AnHy		496492038				
Marinobacter adhaerens DSM 23420	58MaAd	2429063..2430319	*504390518*	2428422..2429060		69.23	yes
Lactococcus lactis subspecies NZ9000 DSM 20481	59LaLa	1319733..1321037	300070941	2145563..2146198		56.10	
	60LaLa	862536..863828	500160180	861915..862436			yes
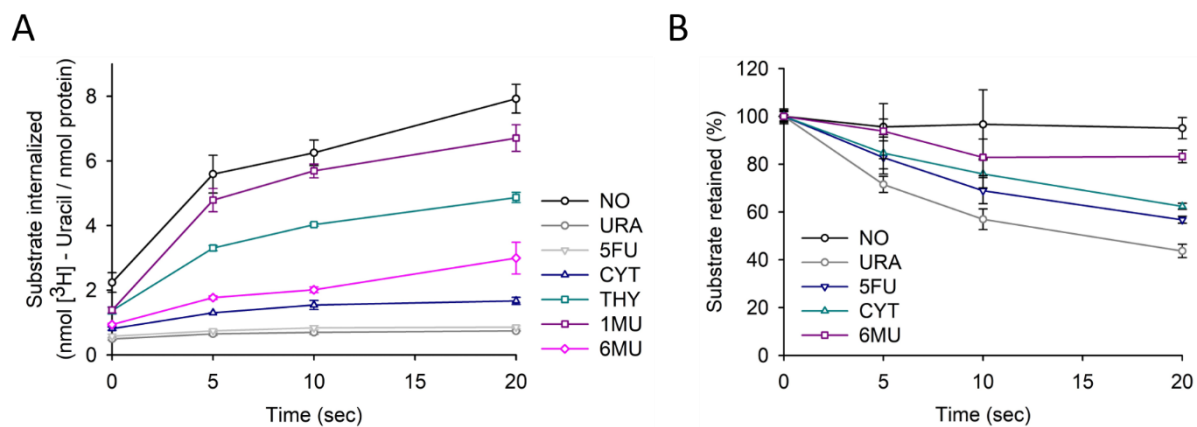
Figure S4: Substrate specificity of reconstituted 51ThXi. (A) Uptake of 1 µM uracil in absence ("NO") or presence of 1 mM competitor in proteoliposomes of 51ThXi. (B) Exchange assay to determine the transport specificity of reconstituted 51ThXi. Proteoliposomes were preloaded with 10 µM partially radiolabeled uracil and either diluted in buffer containing no countersubstrate ("NO") or 10 µM of indicated substrate analogs. All assays were performed at 10°C in triplicates, errors represent standard deviations.
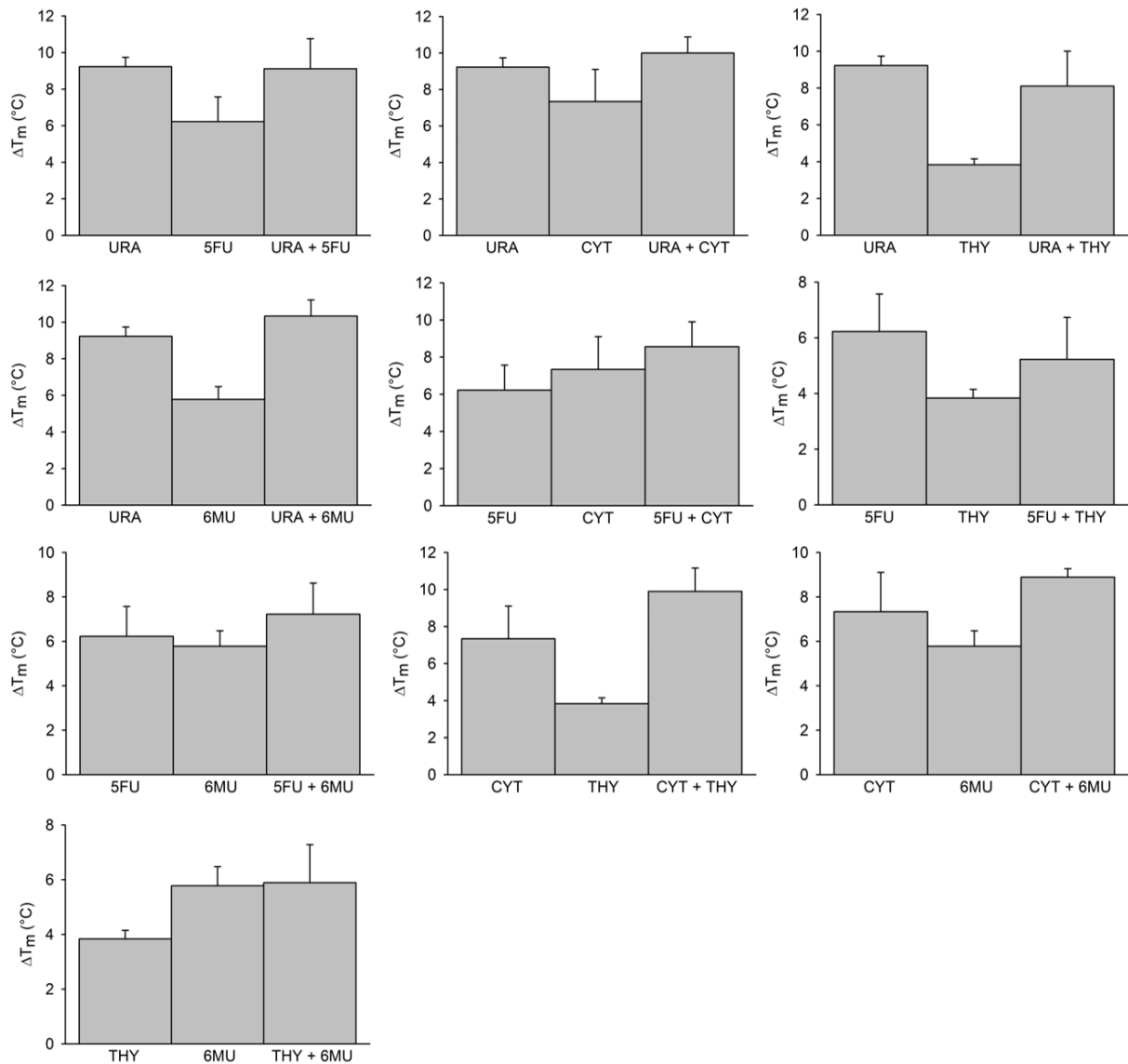
Figure S5: Thermal shift experiments on UraA using substrates in combination. 0.5 mg/mL (18 μM) SEC-pure UraA were preincubated either with 0.5 mM of one substrate, or with 0.5 mM each of two substrates as indicated and melting temperatures were determined in triplicates using DSF. ΔT_m was determined by comparison to the protein's T_m in absence of any substrate. Errors are derived from error propagation of standard deviations. URA, uracil; 5FU, 5-fluorouracil; CYT, cytosine; THY, thymine; 6MU, 6-methyluracil.
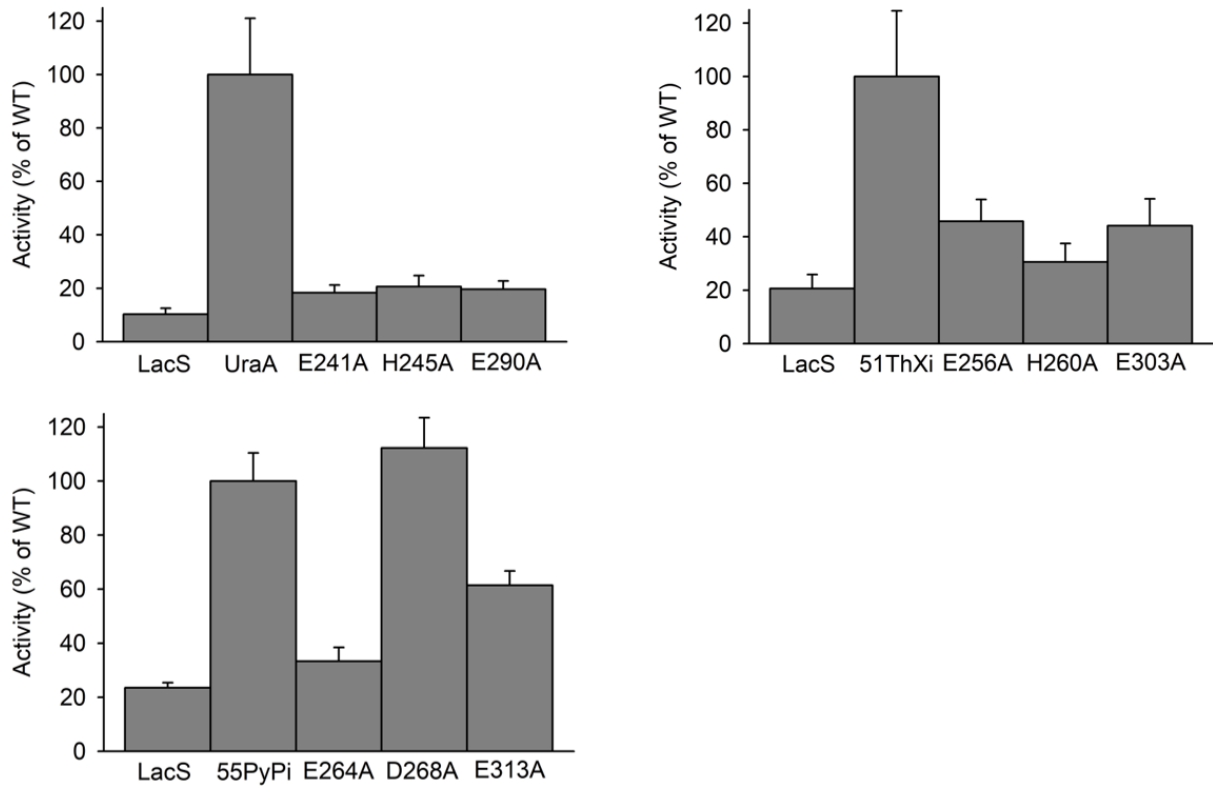
Figure S6: Transport activity of substrate binding site mutants of UraA, 51ThXi and 55PyPi. Initial transport rates of uracil uptake of substrate binding site mutants of UraA, 51ThXi and 55PyPi (**Figure 33**) as indicated were obtained by linear fitting using SigmaPlot 10. Activity of mutants was normalized to the one of wild type protein.
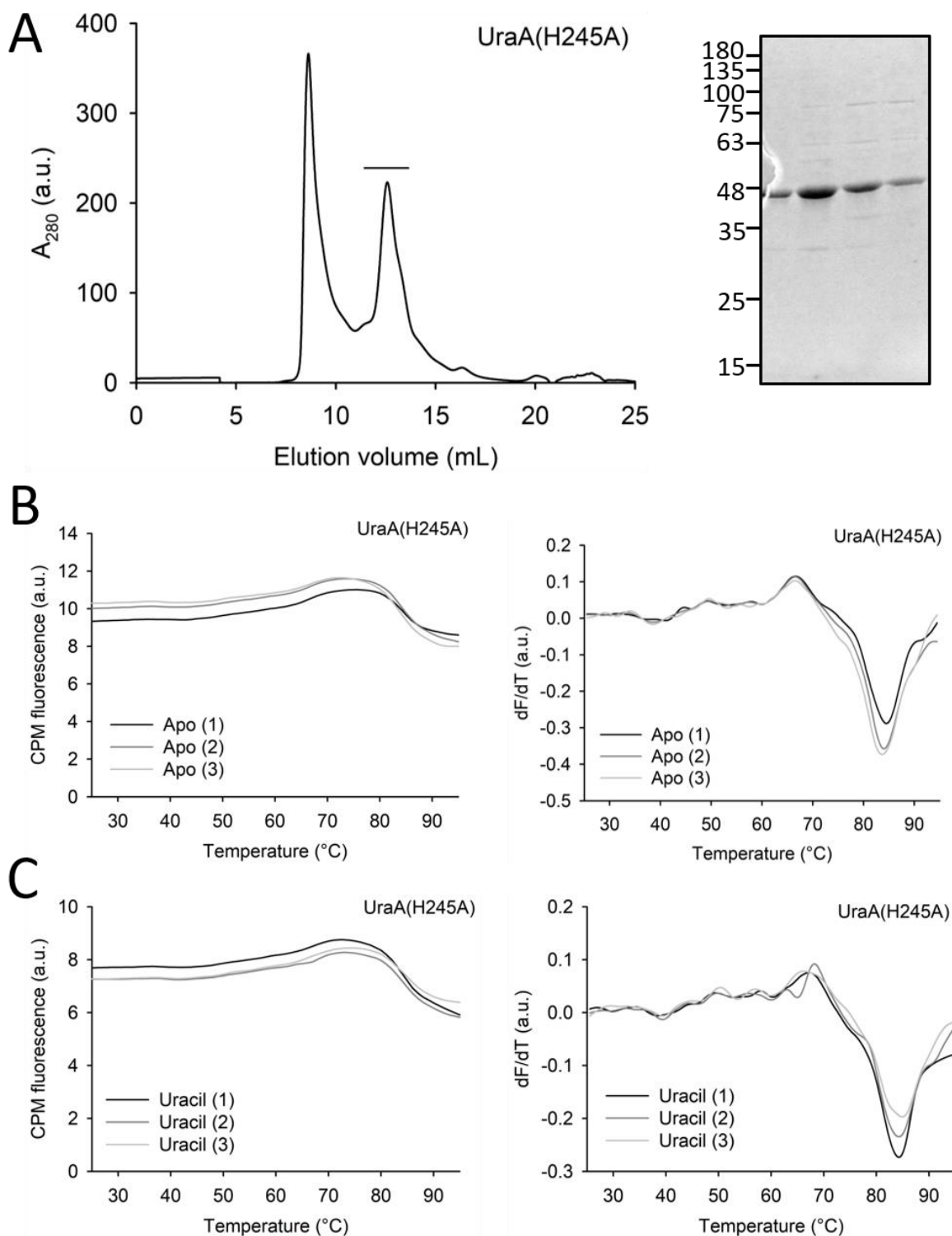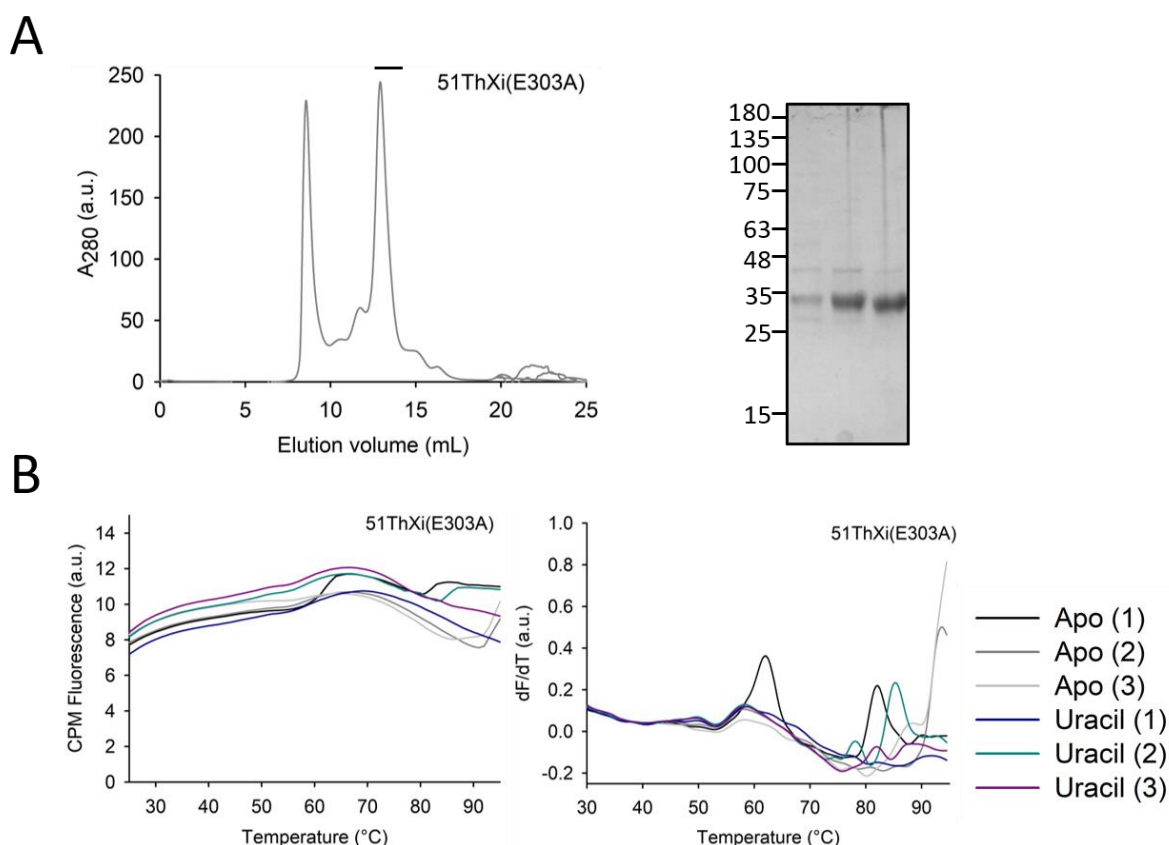
Figure S7: Purification and thermal stability of UraA(H245A). (A) Size-exclusion chromatogram of UraA(H245A) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) DSF on UraA(H245A). 18 μM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Shown are raw fluorescence (left) and dF/dT (right). (C) DSF on UraA(H245A) pre-incubated with 1 mM uracil. 18 μM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Shown are raw fluorescence (left) and dF/dT (right).
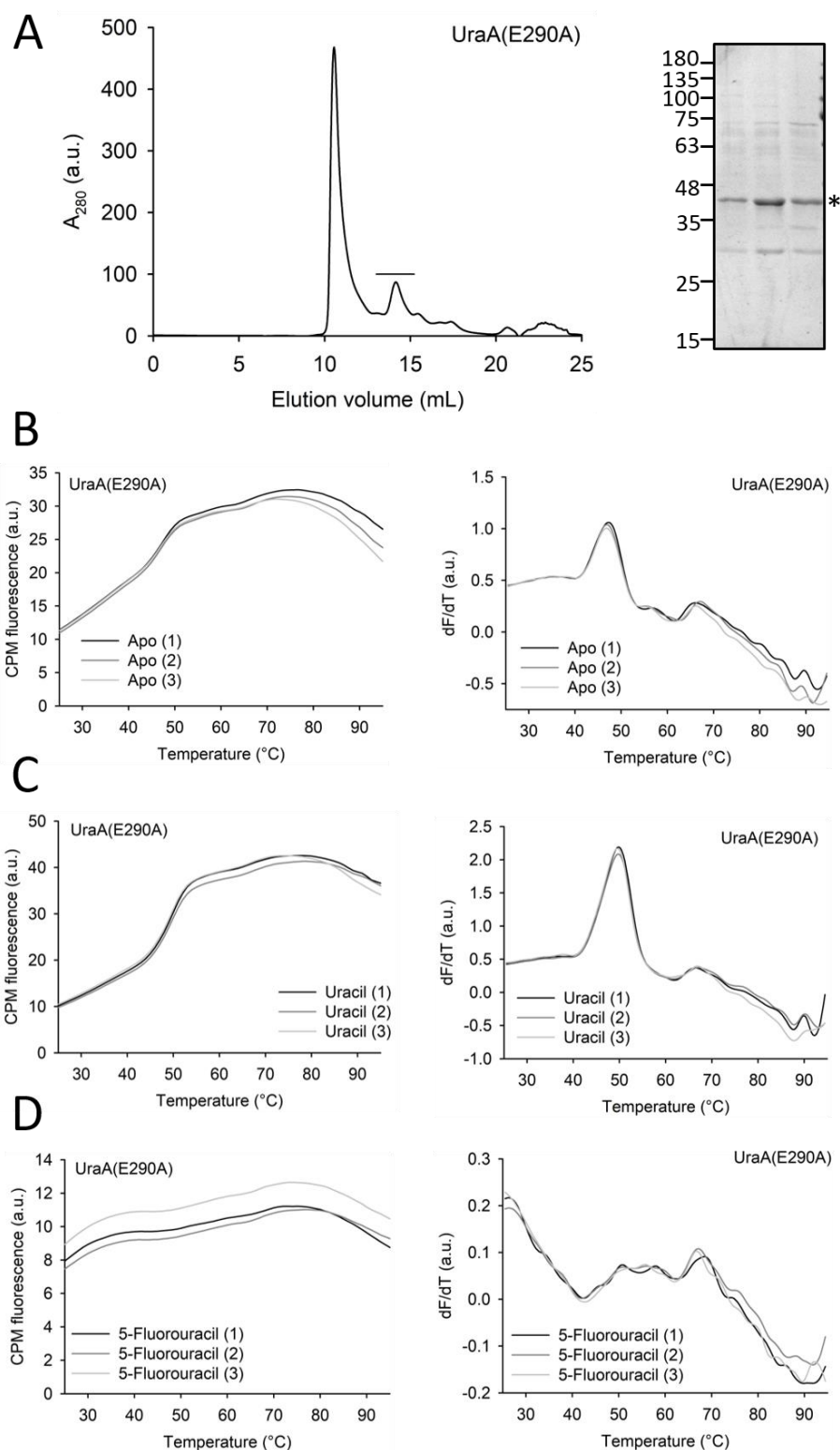
Figure S8: **Purification and thermal stability of UraA(E290A)**. (A) Size-exclusion chromatogram of UraA(E290A) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) DSF on UraA(E290A). (C) DSF on UraA(E290A) pre-incubated with 1 mM uracil. (D) DSF on UraA(E290A) pre-incubated with 1 mM 5-fluorouracil. For all DSF measurements, 18 µM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Shown are raw fluorescence (left) and dF/dT (right).

Figure S9: Purification and thermal stability of 51ThXi(E303A). (A) Size-exclusion chromatogram of 51ThXi(E303A) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) Thermal stability of 51ThXi(E303A) in absence of substrate (dark and grey lines) and pre-incubated with 1 mM uracil (colored lines). 18 µM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Shown are raw fluorescence (left) and dF/dT (right).
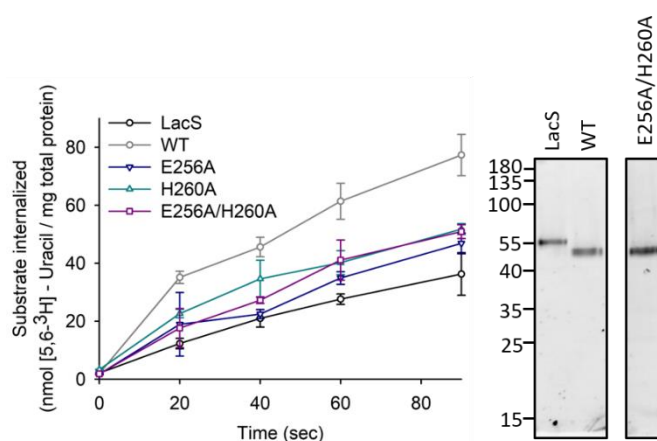

Figure S10: Uracil uptake into *E. coli* BW25113(*ΔuraA*) overexpressing 51ThXi(E256A/H260A). Uracil uptake of *E. coli* BW25113(*ΔuraA*) overexpressing GFP-fusions of LacS as negative control (black), wild type 51ThXi (grey) or alanine mutants of substrate binding site residues of 51ThXi (E256A, dark blue, H260A, teal, or E256A/H260A, dark pink). Uptakes of single mutants is identical to the own shown in **Figure 33B**. Uptakes were measured in triplicates, errors shown represent standard deviations. Expression levels of proteins were monitored by *in gel* GFP fluorescence of whole cell lysates of uptake samples.
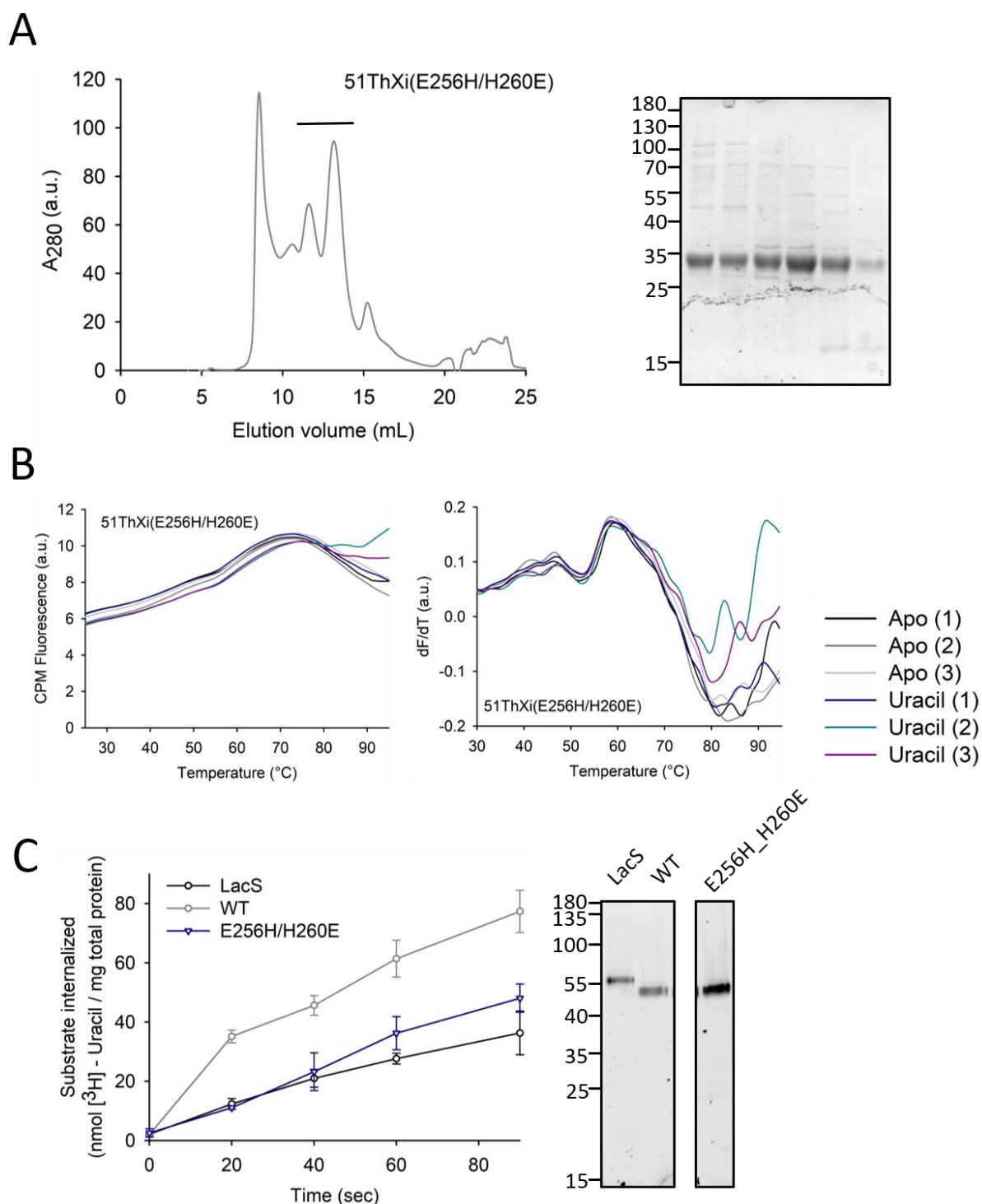
Figure S11: Binding and transport of 51ThXi(E256H/H260E). (A) Size-exclusion chromatogram of 51ThXi(E256H/H260E) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) DSF measurement on 51ThXi(E256H/H260E). 18 μM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Uracil was used at a final concentration of 1 mM. Raw CPM fluorescence is shown on the left, dF/dT on the right. (C) Uptake of uracil into *E. coli* BW25113(*ΔuraA*) overexpressing GFP fusions of LacS as negative control (black), wild type 51ThXi (grey) or 51ThXi(E256H/H260E) (dark blue). Uptakes were measured in triplicates, errors shown are from standard deviations. Expression levels of proteins were monitored by 12% SDS-PAGE and *in gel* GFP fluorescence of whole cell lysates of samples used for uptake experiments.
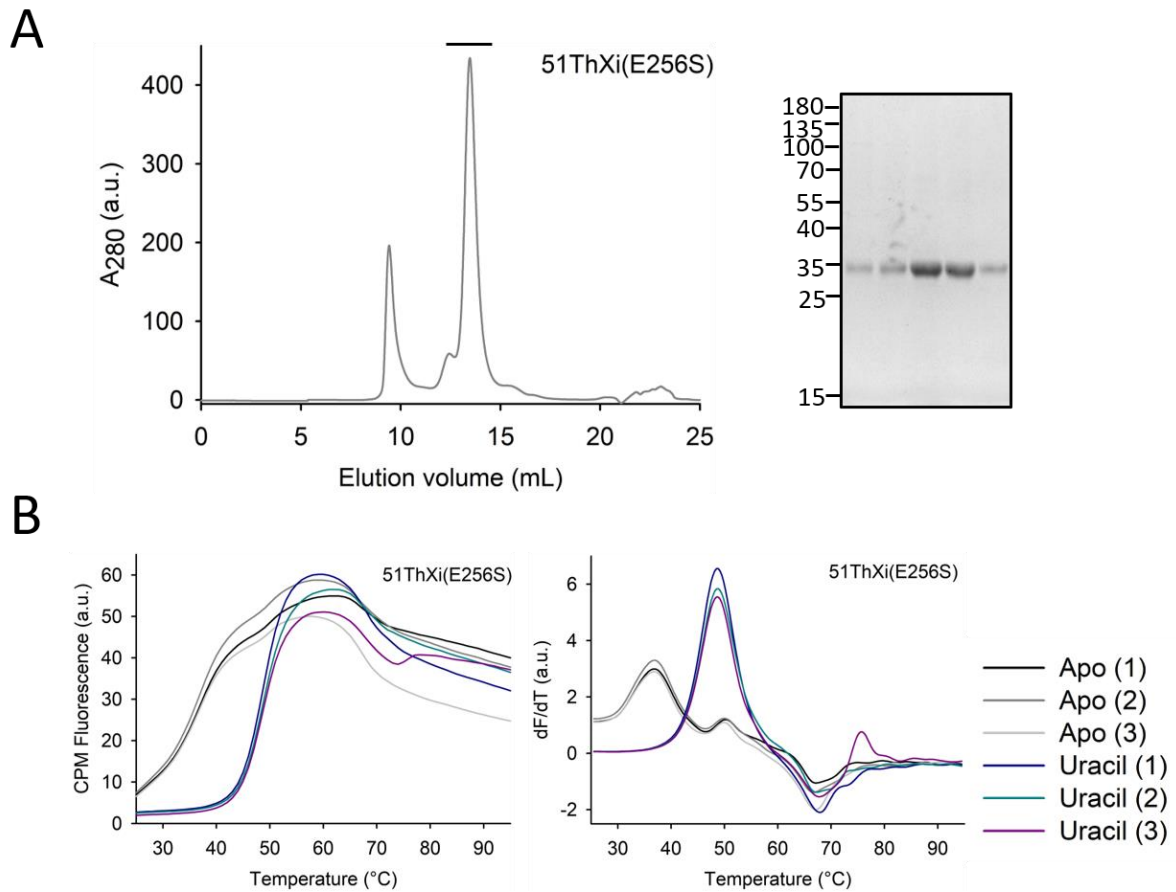
Figure S12: Binding and transport of 51ThXi(E256S). Size-exclusion chromatogram of 51ThXi(E256S) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) DSF measurement on 51ThXi(E256S). 18 µM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Uracil was used at a final concentration of 1 mM. Raw CPM fluorescence is shown on the left, dF/dT on the right.
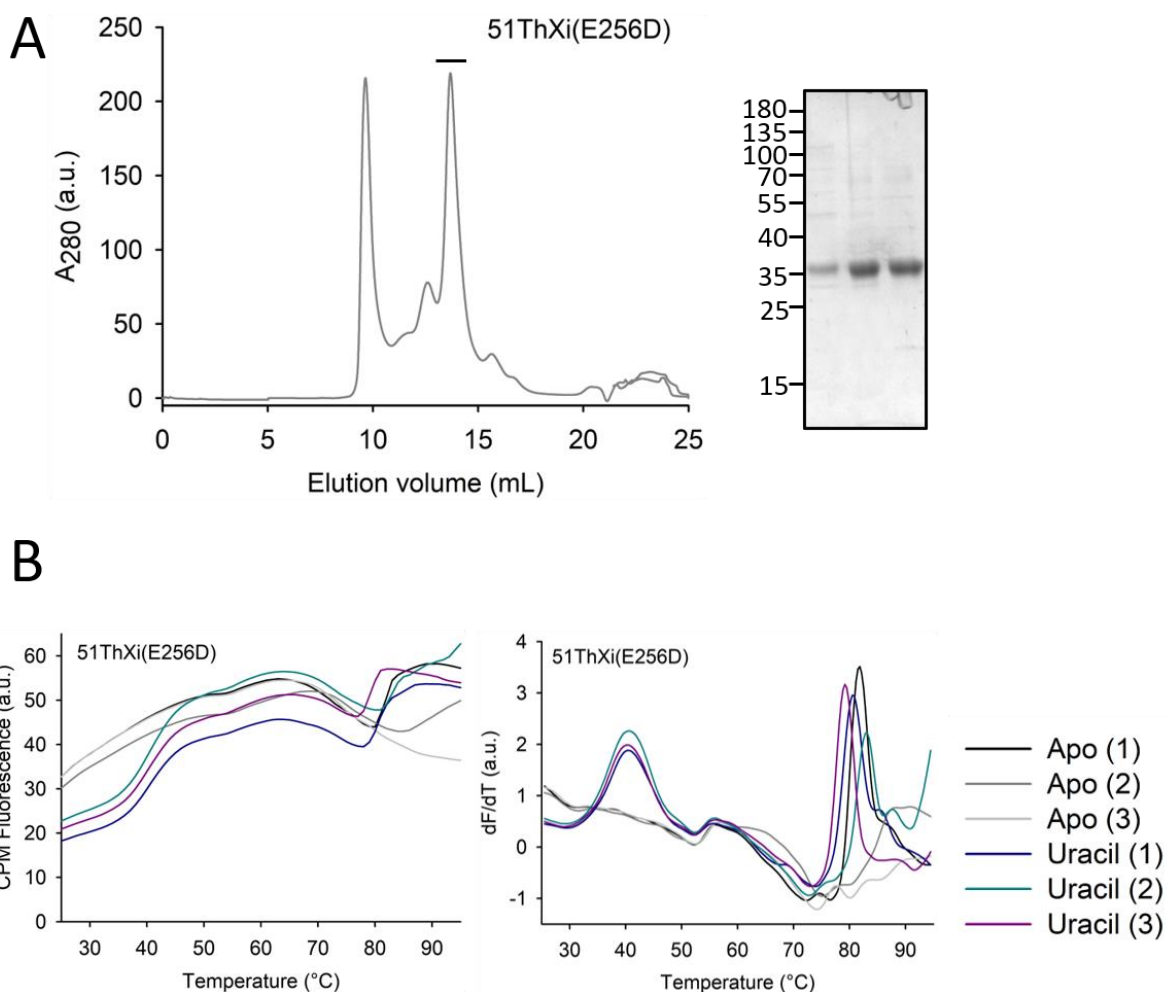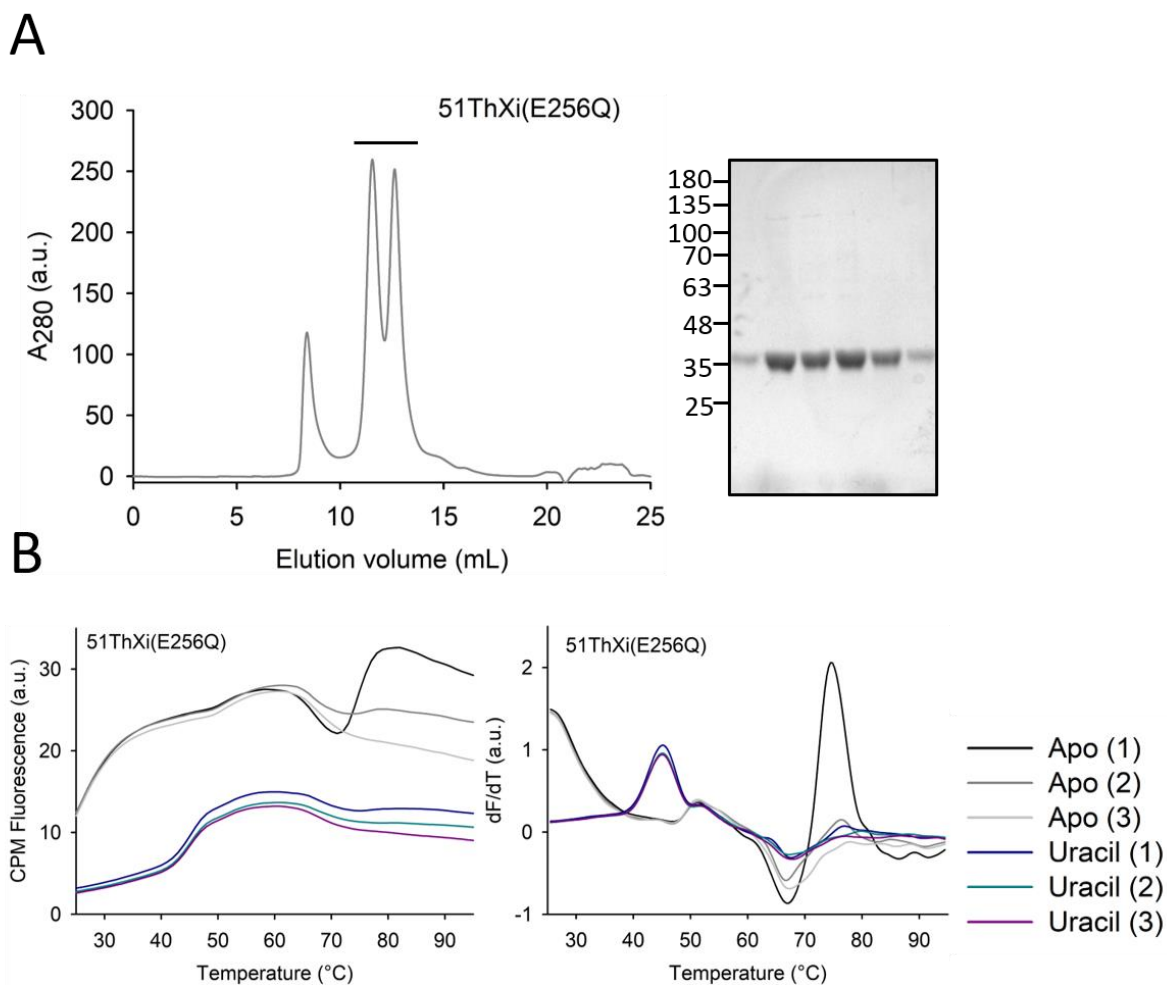
Figure S13: **Binding and transport of 51ThXi(E256D)**. Size-exclusion chromatogram of 51ThXi(E256D) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) DSF measurement on 51ThXi(E256D). 18 µM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Uracil was used at a final concentration of 1 mM. Raw CPM fluorescence is shown on the left, dF/dT on the right.

Figure S14: Binding and transport of 51ThXi(E256Q). Size-exclusion chromatogram of 51ThXi(E256Q) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) DSF measurement on 51ThXi(E256Q). 18 µM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Uracil was used at a final concentration of 1 mM. Raw CPM fluorescence is shown on the left, dF/dT on the right.
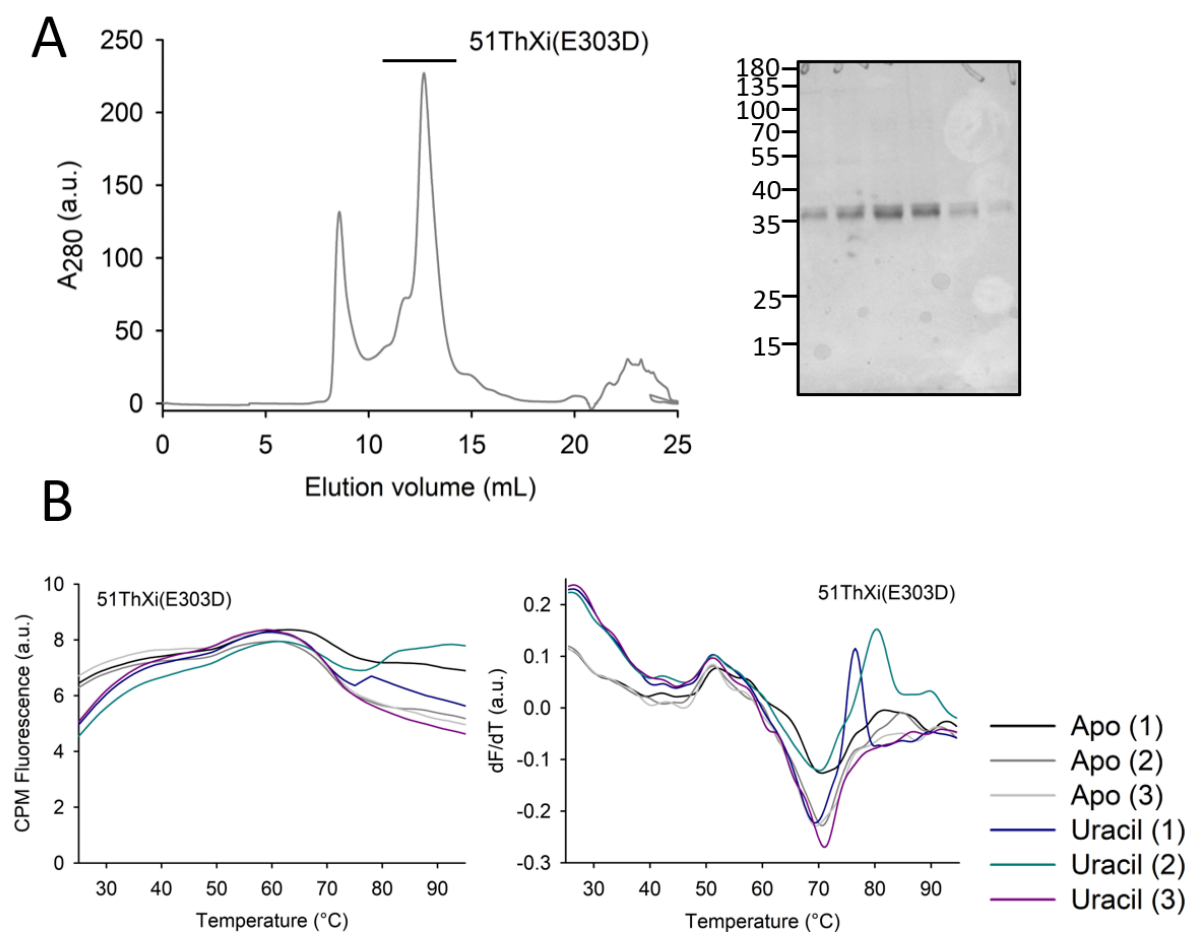
Figure S15: Binding and transport of 51ThXi(E303D). Size-exclusion chromatogram of 51ThXi(E303D) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) DSF measurement on 51ThXi(E303D). 18 μM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Uracil was used at a final concentration of 1 mM. Raw CPM fluorescence is shown on the left, dF/dT on the right.
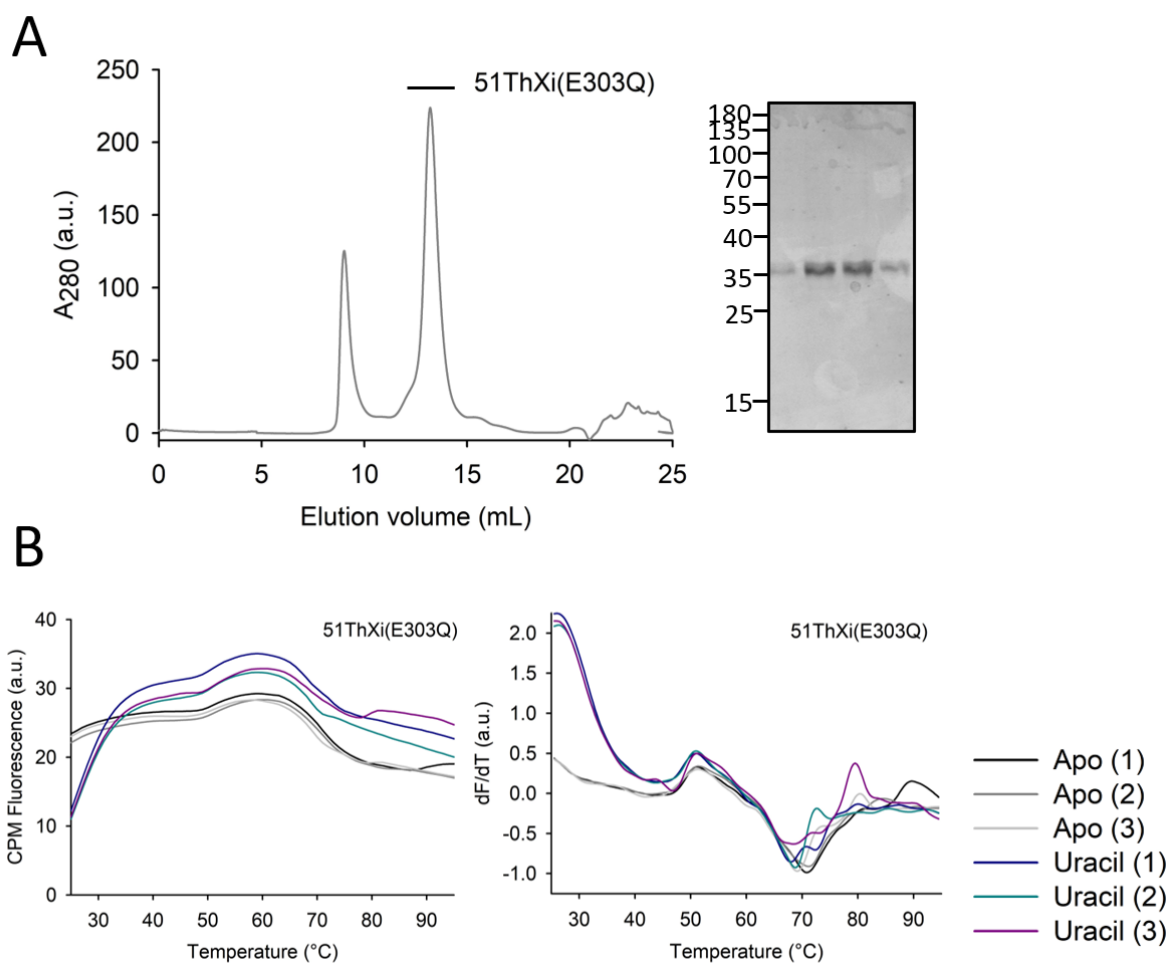
Figure S16: Binding and transport of 51ThXi(E303Q). Size-exclusion chromatogram of 51ThXi(E303Q) on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining (inset). (B) DSF measurement on 51ThXi(E303Q). 18 µM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 10 sec incubation per step and CPM dye fluorescence was recorded. Uracil was used at a final concentration of 1 mM. Raw CPM fluorescence is shown on the left, dF/dT on the right.

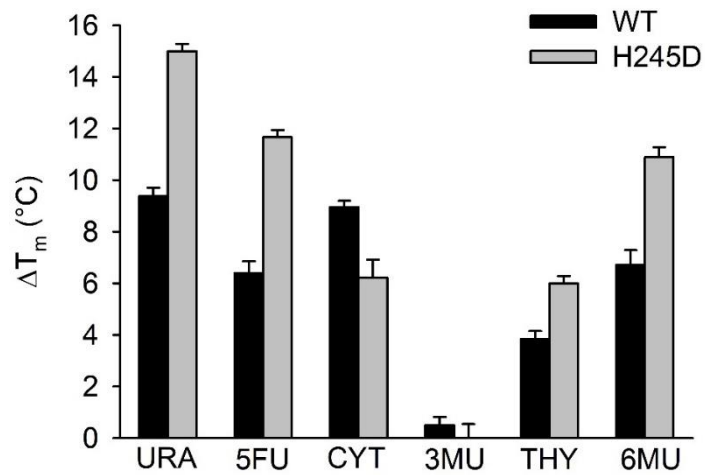**Figure S17: Binding of substrate analogs to UraA(H245D).** Binding of 1 mM of different uracil analogs to 18 µM (0.5 mg/mL) SEC-pure wild type UraA (black) or UraA(H245D) (grey) was tested. To obtain ΔT_m, T_m in presence of respective compound was compared to the T_m of the protein obtained without additional substrate. All T_ms were determined in triplicates, errors shown are derived from error propagation of standard deviations. URA, uracil; 5FU, 5-fluorouracil; CYT, cytosine; 3MU, 3-methyluracil; THY, thymine; 6MU, 6-methyluracil.

```
UraA    ------------------------------------------------------------          0
RutG    ---------------------------------------------MAMFGFPHWQLKS          13
UapA    MDNSIHSTDGPDSVIPNSNPKKTVRQRVRLLARHLTTREGLIGDYDYGFLFRPELPFMKK          60
XanQ    -------------------------------------------------------MSDIN           5
XanP    --------------------------------------------MSVSTLESENAQPVAQ          16
```

```
                    ┌──────────α1──────────┐   ┌──────α2──────┐
UraA    --MTRRAIGVSERPPLLQTIPLSLQHLFAMFGATVLVPVLFHINP-----------ATVL          47
RutG    TSTESGVVAPDERLPFAQTAVMGVQHAVAMFGATVLMPILMGLDP-----------NLSI          62
UapA    DPRAPPFFGLNEKIPVLLAFILGLQHALAMLAGVVTPPLIISSSLSLPSDLQQYLVSTSL         120
XanQ    HAGSDLIFELEDRPPFHQALVGAITHLLAIFVPMVTPALIVGAALQLSAETTAYLVSMAM          65
XanP    TQNSELIYRLEDRPPLPQTLFAACQHLLAMFVAVITPALLICQALGLPAQDTQHIISMSL          76
            .:: *.  :  .   * .*::   :    ::         :
                    ┌──────────α1──────────┐   ┌──────α2──────┐
```

```
        ┌─────────α2─────────┐       ▶▶▶   ┌──────────α3──────────┐
UraA    LFNGIGTLLYLF---------ICKGKIPAYLGSSFAFISPVLLLL---------------          83
RutG    LMSGIGTLLFFF---------ITGGRVPSYLGSSAAFVGVVIAATGF---NGQG------         104
UapA    IVCGLLSMVQITRFHIYKTPYYIGSGVLSVMGVSFSIISVASGAFNQMYSNGFCQLDEAG         180
XanQ    IASGIGTWLQVNRYGI------VGSGLLSIQSVNFSFVTVMIALGSSMKSDGFH------         113
XanP    FASGVASIIQIKAWGP------VGSGLLSIQGTSFNFVAPLIMGGTALKTGGAD------         124
        :   *: : :     . :  :           ::
        ┌────α2────┐              ▶▶▶   ┌──────α3──────┐
```

```
        ┌──α3──┐┌────────────α4────────────┐┌──α4-5──┐┌──────α5──────┐
UraA    -----PLGYEVALGGFIMCGVLFCLVSFIVKKAGTGWLDVLFPPAAMGAIVAVIGLELAG         138
RutG    ----INPNISIALGGIIACGLVYTVIGLVVMKIGTRWIERLMPPVVTGAVVMAIGLNLAP         160
UapA    NRLPCPEAYGALIGTSACCALVEILLAFVP----PKVIQKIFPPIVTGPTVMLIGISLIG         236
XanQ    ----EELIMSSLLGVSFVGAFLVVGSSFIL----P-YLRRVITPTVSGIVVLMIGLSLIK         164
XanP    ----VPTMMAALFGTLMLASCTEMVISRVL----H-LARRIITPLVSGVVVMIIGLSLIQ         175
                    :*    .        .:    ::  *  . *   *   **:.*
                    ┌──────α4──────┐    ┌──α4-5──┐┌──────α5──────┐
```

```
        ┌─α5─┐                    ┌──────────α6──────────┐  ┌──α7──┐
UraA    VAAGMAGLLPAE---G------Q--TPDSK----TIIISITTLAVTVLGSVLFRGFLAII         183
RutG    IAVKSVSASAF--------------DSW----M---AVMTVLCIGLVAVFTRGMIQRL          197
UapA    TGFKDWAGGSACMDDGMLCPSATAPRPLPWGSPEFIGLGFLVFVSIILCERFGAPIMKSC         296
XanQ    VGIIDFGGGFAAKSSG-----------TFGNYEHLGVGLLVLIVVIGFNCCRSPLLRMG         212
XanP    VGLTSIGGGYAAMSDN-----------TFGAPKNLLLAGVVLALIILLNRQRNPYLRVA         223
               .                            .  .               :
        ┌─α5─┐                    ┌──────α6──────┐
```

```
        ┌──────α7──────┐                              ┌──────α8──────┐
UraA    PILIGVLVGYALSF----A---MGIVDTTPIINAHWFALPTLYTPRFEWFAILTILPAAL         236
RutG    LILVGLIVACLLYGVMTNVLGLGKAVDFTLVSHAAWFGLPHFSTPAFNGQAMMLIAPVAV         257
UapA    SVVIGLLVGCIVAA----ACGYFSHADIDAAP-AASFIWVKTFPLSVYGPMVLPIIAVFI         351
XanQ    GIAIGLCVGYIASL----CLGMVDFSSM-RNLPLITIPHPFKYGFSFSFHQFLVVGTIYL         267
XanP    SLVIAMAAGYALAW----FMGMLPESNEPMTQELIMVPTPLYYGLGIEWSLLLPLMLVFM         279
           : :.:: ..            .          .       .:.:. :
        ┌──────α7──────┐  ┌──α7-8──┐                  ┌──────α8──────┐
```
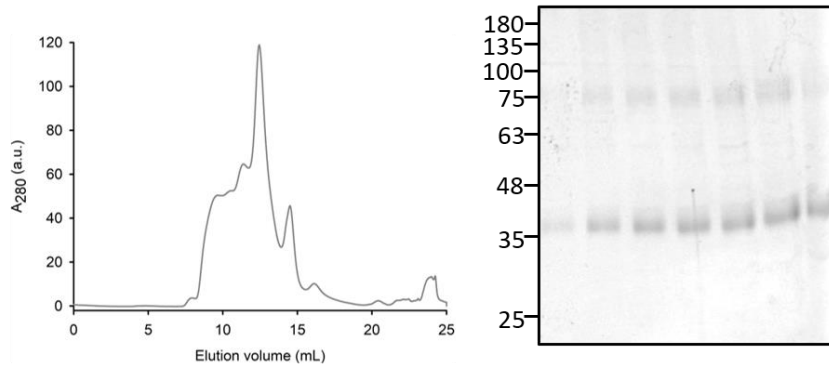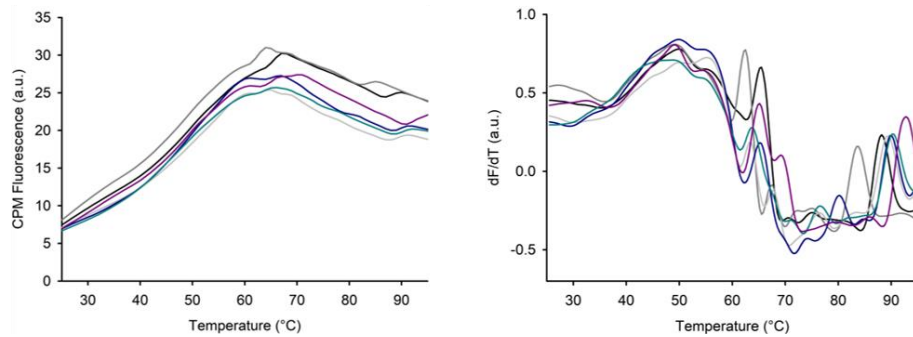
```
             α8                    α9                    α10
UraA  VVIAEHVGHLVVTANIVKKDLLRDPGLHR---SMFANGLSTVISGFFGSTPNTTYGENIG     293
RutG  ILVAENLGHLKAVAGMTGRN--MDPYMGR---AFVGDGLATMLSGSVGGSGVTTYAENIG     312
UapA  ICACECIGDVTATCDVSRLEVRGGTFESRIQGAVLADGINSVVAALATMTPMTTFAQNNV     411
XanQ  LSVLEAVGDITATAMVSRRPIQGEEYQSRLKGGVLADGLVSVIASAVGSLPLTTFAQNNG     327
XanP  ITSLETIGDITATSDVSEQPVSGPLYMKRLKGGVLANGLNSFVSAVFNTFPNSCFGQNNG     339
          :    *  :*.:  ... :          *     ....:*: :.::.      : .:*
             α8                    α9                    α10
```

```
            α10         α11          α11-12               α12
UraA  VMAITRVYSTWVIGGAAIFAILLSCVGKLAAAIQMIPLPVMGGVSLLLYGVIGASGIRVL     353
RutG  VMAVTKVYSTLVFVAAAVIAMLLGFSPKFGALIHTIPAAVIGGASIVVFGLIAVAGARIW     372
UapA  VIALTRCANRWAGYCCCLILIVAGIFAKFAAAIVAIPNSVMGGMKTFLFASVVISGQAIV     471
XanQ  VIQMTGVASRYVGRTIAVMLVILGLFPMIGGFFTTIPSAVLGGAMTLMFSMIAIAGIRII     387
XanP  VIQLTGVASRYVGFVVALMLIVLGLFPAVSGFVQHIPEPVLGGATLVMFGTIAASGVRIV     399
       *: :*         .   .:: ::   ...  **  *:**  .::. :  :*    .
            α10         α11          α11-12               α12
```

```
      α12        α13                                       α14
UraA  IESKVDYNKAQNLILT-SVILIIGVSGAKV------------------NIGAAELKGMA     393
RutG  VQNRVDLSQNGNLIMV-AVTLVLGAGDFAL------------------TLGGFTLGGIG     412
UapA  AKAPF--TRRNRFILTASMALGYGATLVPTWFGNVFPQTENRDLEGFENAIELVLETGFA     529
XanQ  ITNGL--KRRETLIVATSLGLGLGVSYDPEIFKIL-----------PASIYVLVENPIC     433
XanP  SREPL--NRRAILIIALSLAVGLGVSQQPLILQFA-----------PEWLKNLLSSGIA     445
          .  .:   :*:. :: :  *.                      :          :
          α13                                       α14
```

```
            α14
UraA  LATIVGIGLSLIF--KLIS--VLRPEEV-VLDAEDADI--TDK--     429
RutG  TATFGAILLNALLSRKLVD--VPPPE-V-VHQEP----------     442
UapA  VTAFVAMLLNAIMPAEVEEIGAVTPMPVSAHDNRDGEAEYQSKQA     574
XanQ  AGGLTAILLNIILPGGYRQENVLPGITSAEEM--D----------     466
XanP  AGGITAIVLNLIFPPEKQ------------------------     463
        : .: *. ::
          α14
```

Figure S18: Sequence alignments of UraA, RutG, UapA, XanQ and XanP. The sequence alignment of the uracil transporter UraA, the uracil/xanthine transporter RutG and the xanthine transporters XanQ and XanP (all *from E. coli*) as well as the *Aspergillus nidulans* xanthine transporter UapA were generated using Clustal Ω. Secondary elements as deduced from the structures are indicated above for UraA (in teal, PDB: 3QE7) and below for UapA (in dark pink, PDB: 5I6C) with ß-sheets indicated as arrows. Helices of the scaffold domain are colored in a lighter shade compared to helices from the substrate binding domain. Substrate binding site residues are given in bold letters.

A

B

Figure S19: Purification and DSF measurement of the *E. coli* uracil-xanthine transporter RutG. Size-exclusion chromatogram of RutG on Superdex 200 30/100 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM. The protein was overexpressed as GFP fusion in *E. coli* MC1061 and purified by IMAC and SEC. Size and purity of the protein were monitored on 12% SDS-PAGE and Coomassie staining. (B) DSF measurement on RutG. 18 µM (0.5 mg/mL) protein were incubated from 25°C to 95°C in steps of 1°C with 20 sec incubation per step and CPM dye fluorescence was recorded. Uracil was used at a final concentration of 1 mM. The melting temperature can be found at the maximum of the fluorescence derivative (dF/dT). Melting curves were recorded in triplicates.

Supplementary data

Supp. S1: Sequences of new SLC23 homologs tested for uptake of uracil and xanthine.

>DeOr01

MANTQNGVAAVDEMLPAGKLFLYGLQHVLAMYAGAVAVPLIIAAAAGLTQTQTAFLINADLFTCGIATLLQTLGIW
KIGIKIPVIQGVTFAAVTPMVIMAQEGGMTLIFGSVIVAGLFTFLAAPFFSKLIRFFPPVVTGSIITIIGISLLPVGINWAG
GGVGNKAFGSLTFIFIAAVVLVTILLLNKLFTGFISHIAVLLGLIVGLIVAIPFGLVDFSGVSAAPWLGIDIPFHFGVPKFD
IGSIIAMILVMLVVMVESTGDFLAIGEIVGKKIGEEELTAGLRADGLATALGGILNAFPYTAFAQNVGLVGLTGVRSRF
VVAVSGVILVVMGLFPKLATIIASLPNAVLGGAGIAMFGIVAGSGIKTLAKVDFDRNVNNIFIVAISIGIGLIPTLVPTFF
QNFPSWSQTIMHSGITLGSITAIILNAFFNGSKGAGNLDELKANAGMRE

>DeOr02

MKEVQINERLPLTRAIPLGVQHLFAMTGSTILVPFLVGLSPATALFCSGIGTIVFLLLTRSKVPAYLGSSFAFIASLTAFV
KDQNNLSSAMAGVLSVGLAYILIYFVLRLFGTKWINKLIPPVVAGSVVAIIGLSLTPVALQMAGANWIVAIFTLAVAII
VSVYGKGFSKVIPVLLAIVAGYILAYFMGLVDTAKIIESFKTPFVLPFSSFGTIHIDVTAILTFAPLALITLIEDLGHMMILG
NITHTDIIEDPGFDKVVLGNGLATGIASIFGGVPLTTYAENIGVLAITKVYSSLNLWIAAIAAIILSTFNPLGVLIMSIPTP
VMGGVVILLFGMIGAAGLRTLIEAKVDFSKNKNLIIAAVIFAVGIGLANHGIMFATLAGIFLNLVLKDEPEELINTSKQV
AK

>DeOr03

MENSKRIIHSDEKLSISQSLFYGFQSVLACNFFLGPIVLIGAFQLDVKDAAALIAFTFLACGLATIVQSGFFLKFQVIQG
MSFATLGAIISIGMKTDFATLFGALMIGSGVLIVIGVTKIFSKIVSRLIPGLVAGTVIICIGLALMPITWNSLIAIPGNPGI
NFLEAGVTFAAMLIFMRLGSFQSSLGRILSIGSVIYAIVLGTIVASLFGHVDLSPVGTAPWFSIPKLFPYGPPKFDLNAS
LVMTFILLIVLVESIGTWFTYTALSGETMDNQRIDRGVIGEGLGCLIGTLFGGMPVTSYGSNSGVLAVTKVFSRYAAL
GAGVIAIAMALCPKLMYLIAVIPSSVIWGIYAVICVAVVMSGLASIRVYPFSERNNLVVGVSILTTIGASLLPPQLVNSL
PSLISYVLGSAIIVGALTAIILNLIIPEKEEDGISLGQIEANPAAVATK

>IlPo01

MTTLGTKTKYILGLQHVLAMFGATVLVPFLTGLNPSIALLTAGAGTLLFHYCTKKIVPVFLGSSFAFIGAISLVLKEEGIG
VVKAGVIGAGLVYIAMAYIIKTYGVERVSSFFPAIVTGPTIMVIGLRLSPVALGMIGYSNGHFDIKGLIIALIVVTSMIAI
SVLEKSFFRLVPILISVTLGYLASIPAGLVNFEPIKNAGWIGFSPEALHDLTAIPEFSVTGLLAIAPIAMVVFIEHIGDITTN
GAVVGKDFFKEPGIHRTMLGDGVATLFAGFLGGPANTTYGENTGVLAVTKVYDPSVLRIAAGYAIVLSFIGKFGAILQ
TIPTPVMGGVSLILFGMIASVGVRTLIEADLDFGHSRNLIIASLIFVLGIAIDNVVVWKTVSLSGLAIAAFTGVLLNKILP
KNI

>IlPo02

MKNTSPYHLDGVPPLKVAFPLGLQHILAMFVSNLTPIIIVSGVLGLPQEQKTFLIQCTMLVAGLNTMIQAYSLGPIGA
KLPVVVGTSFAFVPVAISIGTKYGFEAVLGAALVGGIFEALLGSVIGKIRKFFPPIVTGVVVLSIGLSLLPVGIKYFAGGV
GAADFGSPVNMTIGMIVLLTVIFFKQFTKGITSTASVVIGTIVGFFVAALFGKVDLGAVSQANFFIVPKPFTYGFAFHI
DAILAMVLMFVVSAVETVGDMSGVTMGGAGRETTDRELSGGIMADGFGSALASAFSILPTTSFSQNTGLVAMTGI
MSRHVVAVGASLLVMGAFIPKIGALFTIIPPSVIGGSLVMIFAMISISGINLITKDKLEGRNSVIIAVSLGLGFGLGSVPE
ALAHFPKTIQLIFGGSGIVVSGAIALILNIVLPQDEVEEKETEAKLKNA

Supplementary data

>PrMi01

MVTSENRNLSQSSQISPQKTSELLFALEEKPPLLQTLFAACQHLLAMFVAVITPAILICQALGLPAHDTQRIISMSLFAS
GIASLIQIRAWGPVGSGLLSIQGTSFNFVAPLIMGGLALKNGGADIPTMMAALFGTLMVASLTEVLLSRFLHLARRIIT
PLVSGIVVMIIGLSLIQVGLTSIGGGYAAIESNTFGSPKNLLLAGSVLVVIIILLNRQRNPYLRVASLVIAMAVGYILAWW
LDMLPTPPEQQETPIITVPEPFYYGLSFDWHLLIPLMLVFMITSLETIGDITATSDVSEQPVSGPLYMKRIKGGVLANG
LNSMVSAFFNTFPNSCFGQNNGVIQLTGVASRYVGYVVAAMLIILGLFPSVAEFVQQIPEPVLGGATLVMFGTIAAS
GVRIVSKEALNRRAIMILAISLAVGLGVSQQPQILQFAPDWLKTLLSSGIAAGGLTAIILNVIFPPEK

>PrMi02

MTRRAIGVEERPPLLQTIPLSFQHLFAMFGATVLVPILFKVNPATILLFNGIGTLLYLFICKGRIPAYLGSSFAFISPVLLLL
PLGYELALGGFIVCGVLFCLVAGIVKIAGRGWINVMFPPAAMGAIVAVIGLELAGTAANMAGLLPSADNPVDSQTLI
ISMVTLGVTILGSVMFRGFLAIIPILIGVLAGYALSFFMGVVDVTPIKEANWFALPTFYSPRFEWSAIFIILPAALVVIAE
HVGHLVVTANIVQRDLMKNPGLHRSMFANGFSTIISGFFGSTPNTTYGENIGVMAITKVYSTWVIGGAAIIAILLSCV
GKLAAAIAAVPVPVMGGVSLLLYGVIGASGIRVLLDSKVDYNKPQNLILTAIILIVGVSGASIQIGAAELKGMALATIVG
IVLSLVFKLVSIIRPETELVDDTLESINKHTNTK

>ShBl01

MSVKAVESENRKPETQDRPSELIYRLEDKPPLPQTLFAACQHLLAMFVAVITPALLICQALGLPAADTQHIISMSLFAS
GVASIIQIKAWGPVGSGLLSIQGTSFNFVSPLIMGGMALKNGGADIPTMMAALFGTLMLASCTEMIISRGLHLAQRI
ITPLVSGVVVMIIGLSLIQVGLTSIGGGYTAINNHTFGAPSNLLLAGAVLVVIIALNRQRNPYLRVASLVIAMAVGYLLA
WFMGMLPENTTPSNSDLFMIPTPLYYGLGIDWSLLIPLMLVFMVTSLETIGDITATSDVSEQPVTGPLYLRRLKGGVL
ANGMNSFVSAVFNTFPNSCFGQNNGVIQLTGVASRYVGFVVALMLIILGLFPAVSGFVQHIPEPVLGGATIVMFGTI
AASGVRIVSREPLNRRAIMIIALSLAVGLGISQQPLILQFAPDWLKTLLASGIAAGGITAIVLNLIFPEDQPQKR

>ShBl02

MTRRAIGVSERPPLLQTIPLSLQHLFAMFGATVLVPIMFHINPATVLLFNGIGTLLYLFICKGKIPAYLGSSFAFISPVLLL
LPLGYEVALGGFIMCGVLFCLVALIVKKAGTGWLDVMFPPAAMGAIVAVIGLELAGVAANMAGLLPADGTTPDTK
AITVSLVTLAVTIFGSVLFRGFLAIIPILIGVLVGYALSFAMGMVDITPIRDAHWFALPTFYTPRFEWFAIFTILPAALVVI
AEHVGHLVVTANIVKRDLLRDPGLHRSMFANGFSTIVSGFFGSTPNTTYGENIGVMAITRVYSTWVIGGAAIFAILLS
CVGKLAAAIQIIPVPVMGGVSLLLYGVIGASGIRVLIESKVDYSKAQNLILTSVILIIGVSGAKVHIGAAELKGMALATIV
GVGLSLIFKVISLVRPEEEIQESDEKRA

>YeEn01

MSRRTIGVSERPPLLQTIPLSFQHLFAMFGATVLVPILFKINPATVLLFNGIGTLLYLFICKGKIPAYLGSSFAFISPVLLLL
PLGYEVALGGFIMCGVLFCLVALIVKKAGTGWLNVLFPPAAMGAIVAVIGLELAGVAAGMAGLLPAEGVTVDSTTIII
SMVTLGVTILGSVLFRGFFAIIPILIGVLVGYALSFVMGVVDLTPIREAHWFALPTFYTPRFEWFAILTILPAALVVIAEHI
GHLVVTANIVKKDLMRDPGLHRSMFANGISTVISGFFGSTPNTTYGENIGVMAITRVYSTWVIGGAAILAIMLSCIG
KLAAAIQAVPVPVMGGVSLLLYGVIAASGIRVLIESKVDYNKAQNLILTSVILIIGVSGAKVNIGATELKGMALATVVGI
GLSLLFKVISLIRTEEEIIEATEDKPALK

>YeEn02

MSTQSAELNTPQQATTRPSELIYRLEDRPPLAQTLFAACQHLLAMFVAVITPGLLICQALGLPAEDTQRIISMSLFASG
LASLLQIKTWGPVGSGLLSIQGTSFNFVSPLIMGGLALKNGGADIPTMMAALFGTLMVASCTEIILSRFLHLARRIITP
LVSGIVVMIIGLSLIQVGLTSIGGGYGAMSDHTFGSPKNLMLAGAVLVVIIILLNRQRNPYLRVASLVIAMAVGYLLAW
TLGMLPESRPVVDTALITIPTPLYYGLSFDWNLLVPLMLIFMVTSLETIGDITATSDVSEQPVHGPLYMKRLKGGVLA
NGLNSMLSAVFNTFPNSCFGQNNGVIQLTGVASRYVGFVVALMLIVLGLFPAVAGFVQHIPEPVLGGATLVMFGTI
AASGVRIVSRETLNRRAIMIMALSLAVGMGVAQQPLILQFAPDWIKTLLSSGIAAGGITAIVLNLIFPQEK

Supplementary data

>YeEn03

MANNWFPTWRKRSGNLDGAIIAPDERLPVGATLIMGLQHAVAMFGATVLMPLLMGFDANLAILMSGIGTLLFFLI
VGGRVPSYLGSSAAFVGLVIAVTGYTGSGPNPNIALALGGIIACGAIYTLIGFIVMSIGTRWIERLMPPVVTGAVVMAI
GLNLAPIAVNSVSGSSFNSWMAVVTILCIGVVAVFTKGMIQRLLILVGLILSYVIYWVVTTILGWGTPVDFGPISQAA
WFGWPQFTTPVFDAHAMLLIAPVAVILVAENLGHIKAVAGMTGQNLDPYMGRAFVGDGLATMLSGSVGGTGVT
TYAENIGVMAVTKIYSTLVFVAAALVAILLGFSPKFGALIHTIPGPVLGGASIVVFGLIAVAGARIWVQNKVDLSDNG
NLIMVAVTLVLGAGNFALTLGNFTLGGIGTATFGAILLNALLQRRKILPKLGSDGKPLPQDG

>DiIn01

MEGFMEKRIIEVEERLPLLPTIPLSLQHLFAMFGSTVLVPFLLHVDPATCLFMNGVGTLLYLFICKWKLPAYLGSSFAFI
SPVLAVTATEGMTYADAQSGFICFGLSFMILALLVKKIGTSWIDVLFPPAAMGSIVAIIGLELAPLAMSMSGFSGEAQ
GMTNAMAVMISMFTLIVTILATVLGRGFISIIPILIGVIAGYILSIIMGVVDFSNVAAKEWFAVPTFYSPRFNFSAVMM
IMPALFVVFAEHLGHLFVTSDIVGRDLIKDPGLHRSLFADGLSNMISGFAGSTPNTTYGENMGVMAMTGVYSTWV
IGGAAVFAIIFSFIGKVAAVIHAIPTPVMGGVCILLFGFIASSGIRMLIEKKVDYTKSRNLILTAVTMISGLSGATVTLGP
VQLKGMGLATIVAMILSLSFLIFDRLHLSNDKK

>DiIn02

MPKQASLRSAEEAVMNTDPIYEADGKISVRKAVPFGLQHVLAMFVANIAPILIVTGVVKMPASEAGAVVQAAMIIA
GIGSLLQMYPVFRLGSGLPVIMGISFTFVSVFCVIGLKYGYGAILGAVLIGGVLEGILGLGAAWWRKLVPPIVSATVVT
AIGFSLLPIGANSFGGGFGHPEFGDVRFLIVGTITLVSCLIFNIRAKSFYKQLSVLFGLFVGYVTAYFYGMVDLGRLTEV
SLVSLPVFMPYSLEFHYDAIFSVFLIFLVSATETLGDTSALSAMGFNREAKDREISGSIAVDGFVSAASSLFGCLPITSFS
QNVGLIAMTHVVNRKAIASGAVIMVLAGLIPALGVILASLPEAVLGGCTLMMFGSIVVSGVQMISRCGYSQRNMSI
AALSLSIGLGFTQTPQIFRIFPELLRSVFAENCVAVVFIVAVLLNLVFPKEEEEKAAAGAAE

>ClAs01

MGMKRSEGDEMSGETKRADILYNIDDRPPVGKSVIFAFQHILAMFAGNVTVPLLVINIVGLNSEEGTFLIQCALLVA
GVATLLQVRGIKAVGSRLPIVMGTSNAFLSTVVAITSQYGIGACLGASFIGGLFEAVLGNFIGRLKKIFNPLVSGIVVM
TIGITLIPTGMKQAAGSKTAAGLGAPVNLLLSGLVILVIVLCSRSRNKTLKSASILVGIVVGYVVAAVAGLVDFSAIGQA
AAFSVPLPFRYRWEFHWSAIVAMLFMYVATTVETVGDMTALTVVAQNRQPTPEESRGGILADGLGSSLAAVFNAF
PNTSYTQNIGVVNLTGVFSRSIVNIGAVILVGMSLFPKLSAVILCVPEPVLGGATLITFMMVFISGVSLITSVDLGSRN
MLIMAVSLGIGVGFSLVPDVTKVFGESVSVCLNNGIVPASLIAIALDWFLPKDGNGTAGEEEIREAQ

>ClAs02

MDNRRIIQVEEKVPFNLLVPLSLQHMFAMFGASVLVPSIFHISPAIVLFMNGIGTLLFIFVTKGKAPAYLGSSFAFLAP
AGVVIANFGYEYALGGFVAVGFCGCILAYIVYKFGTDWIDVVLPPAAMGPVVALIGLELSGSAASNAGLLDPVINPK
NVIVFLVTLGTAVFGSILFRKFLSVIPILIAVVAGYVAAVACGIVDFSQVAAAPIFALPNFTTPKFNLNAIMMILPVLLVI
TSEHIGHQVVTSKIVGRDLLKDPGLHRSLFGDNFSTMISGLIGSVPTTTYGENIGVMAVTRVYSVRVIGGAAVLSIACS
FIGKLSTLIQTIPGPVIGGISFLLYGMIGTSGLRILVDAQVDYSKNRNQALTSVIFVAGLSGIALKIGSIQLTGMVLACVV
GMVMSLVIYLLDRAHLTNDQD

>MiMu01

MTTTAAASVHPVDQKLPWGRTIIYAFQHVLAMYAGAVAVPLVLAGAVHLSTEELIYLINADLFTCGIATLVQSLSIRN
FIGSKLPIVQGCTFTAVTPMIIIANTHGGGAAGLQVVYGAAIFAGIACFVISNFFSKMLRFFPPVVTGSVITVIGLSLMP
VAVNWIAGANPAASDYCEPTHILMAALVLAIILVIYRVCRGFLRQISVLLGLVIGTLIADALGMADFHAVGEAAALGIT
TPFAFGYPVFEPVSCLAMTLVMLVTMAETTGDIVAVTEIVGKPMTESMLTKALRADGFSTALGGVFNTFPYTAFAQ
NIGLIGMTGVRSRYIVAVCGVMLMILGLFPKMAAVVASVPTMVLGGAGLAMFGMVAASGFQALSKAELHKKGNI
MVVAISVAVALIPIGVPSFYSKFPTWVQIICNSGITIGSITAIVLNLLLRDAPEPDEEKLPFSRSKDA

Supplementary data

>MiMu02

MQKRPIGVEEKLPLLETIPLSLQHLFAMFGSTVLVPILFKINPATVLLFNGIGTILYLILCKGKIPAYLGSSFAFLSPVFLVL
AQYSYEAALGGFICVGIVFCLVGLLVRAAGTGWIDVIFPPASMGAIVAVIGLELMPTAAGMAGLTGDKTDSTAIFVS
LATLAITIFASVAFKGFLSIIPILIGVVSGYIIAALCGIVDWSLVESAPWFALPTFYTPKFDLGAILIILPASLVVIVEHIGHLI
VTGNIVGHDLTKDPGLDRSIFGNGLSTIISGFFGSTPNTTYGENIGVLAITKVFSTWVIGGAAVFAILLSCLGKLAALIQS
IPTPVMGGVSMLLFGVIAASGIRILVEAKVDYNNPMNLLLTSIVMGIGVSTASLTIGTVSFKGMSLATVVAIILSLAFRL
IAMVRGDKTPLVGQKAEDNTHQ

>LaLa01

MPHKNEQSQEQSNSKSAVLGLQHLLAMYSGSILVPIMIAGALNYSATQLTYLISADIFMCGLATLLQLQMRKYFGIG
LPVVLGVAFQSVAPLIIIGQRHGSGAMFGSLMVAGVFVILISGIFSKIRKLFPPIVTGSVITTIGLSLIPVAIGNMGNNVE
KPTIQSLILAVSTILIILLINIFTTGFIRSIAILIGLIAGTIIAASMGLVDFSVVSQAPWAHLPQPFYFSAPKFYLADSLMMIII
AIVSLVESTGVYLALADITGENLDEKRLRNGYRAEGFAVFLGGIFNTFPYTGFSQNVGLVQLSGIKTRKPIYFTAGFLIIL
GLIPKFAAVAQLIPTPVLGGAMLIMFGMVATQGIRMLAKVEFEGNQNLLIAAVAIAMGVGFNSTNLFTALPSFIQPF
VSNGIVMSTVSAIILNLVFNHGKKDELAADEAK

>LaLa02

MHDNDIILKVDEKPAASQWFGLSFQHLFAMFGSTVLVPILVGINPAIALLSSGLGTLAHMSVTKFKVPAYMGSSFAY
IGAMTLLMKNGGMPAIAQGAMTGGLVYLIVALIVKFAGKGWIDKVLPPIVVGPIVMVIGLSLAPTAINDAMYTDVA
NLKGYSLAYIIIALITVLSIVVYSIYGKGFLSVVPILLGIITGYVAAMIIGKITGMNIVSFTGISQAKWLTLPPMEIPFASYK
WAFYPSAILTMAPIAFVTMTEHFGHIMVLNSLTKKDYFKEPGLEKTLTGDGLAQIIAGFIGAPPVTSYGENIGVMAIT
KIHSIYVIAGAAVLAIVVSFVGKITALLQSIPAPVIGGASIALFGVIAASGLKILVENKVDFDIKRNLLISSVVLVIGIGGMII
NITQNLQISSVAIATILGIVLNLVLPKDASEI
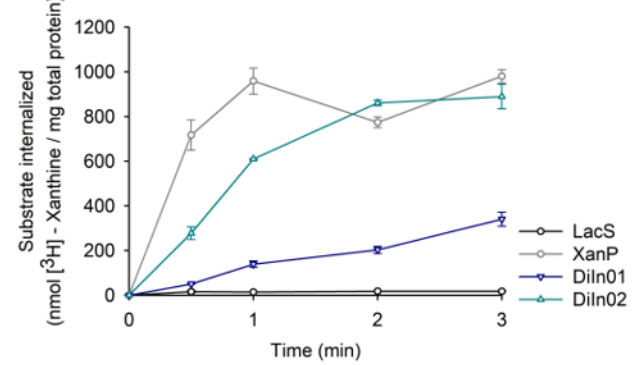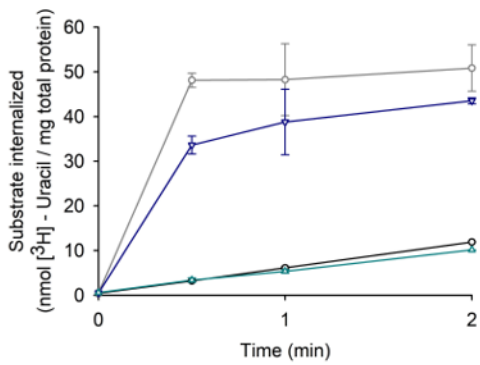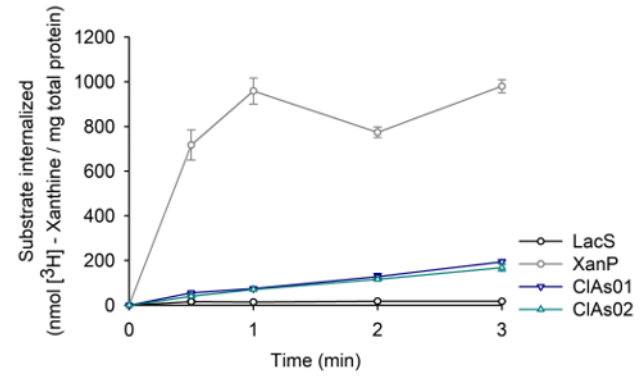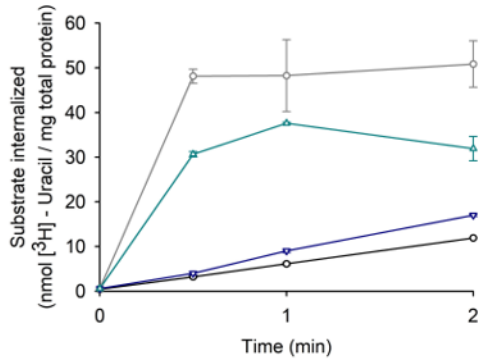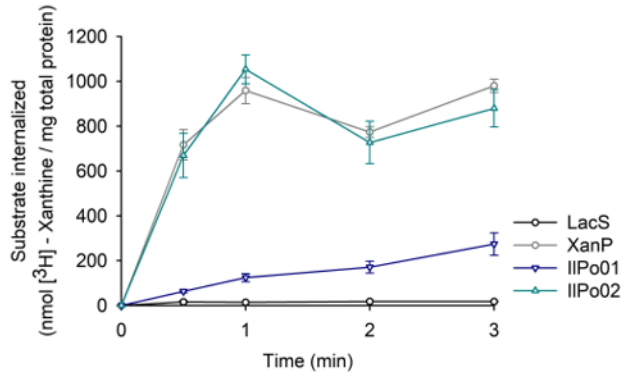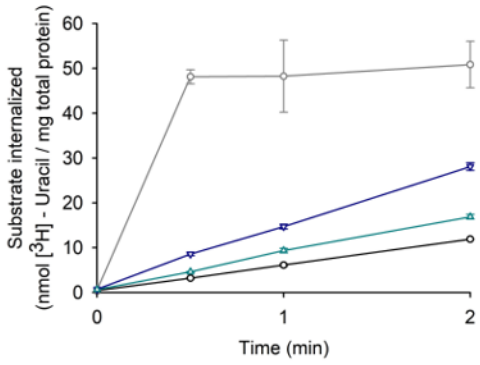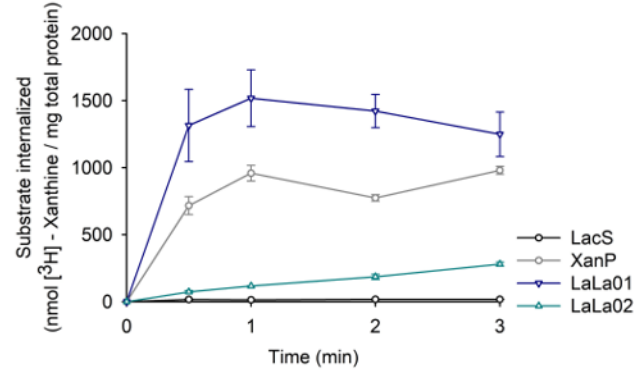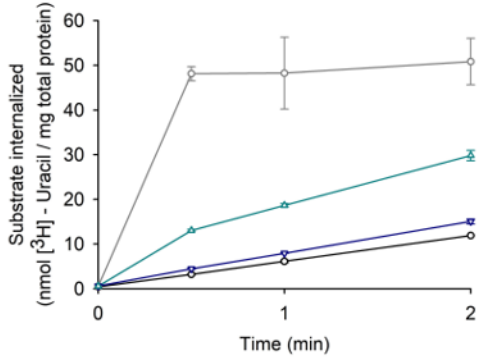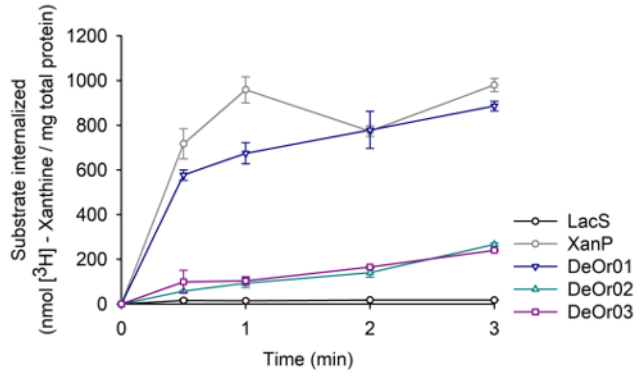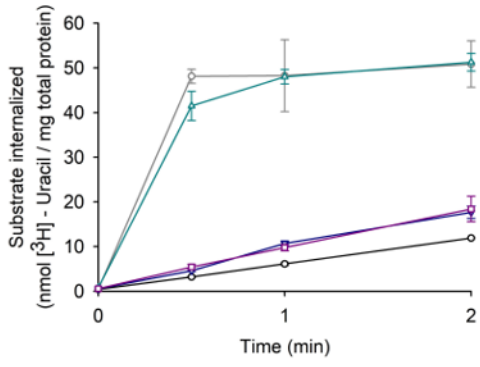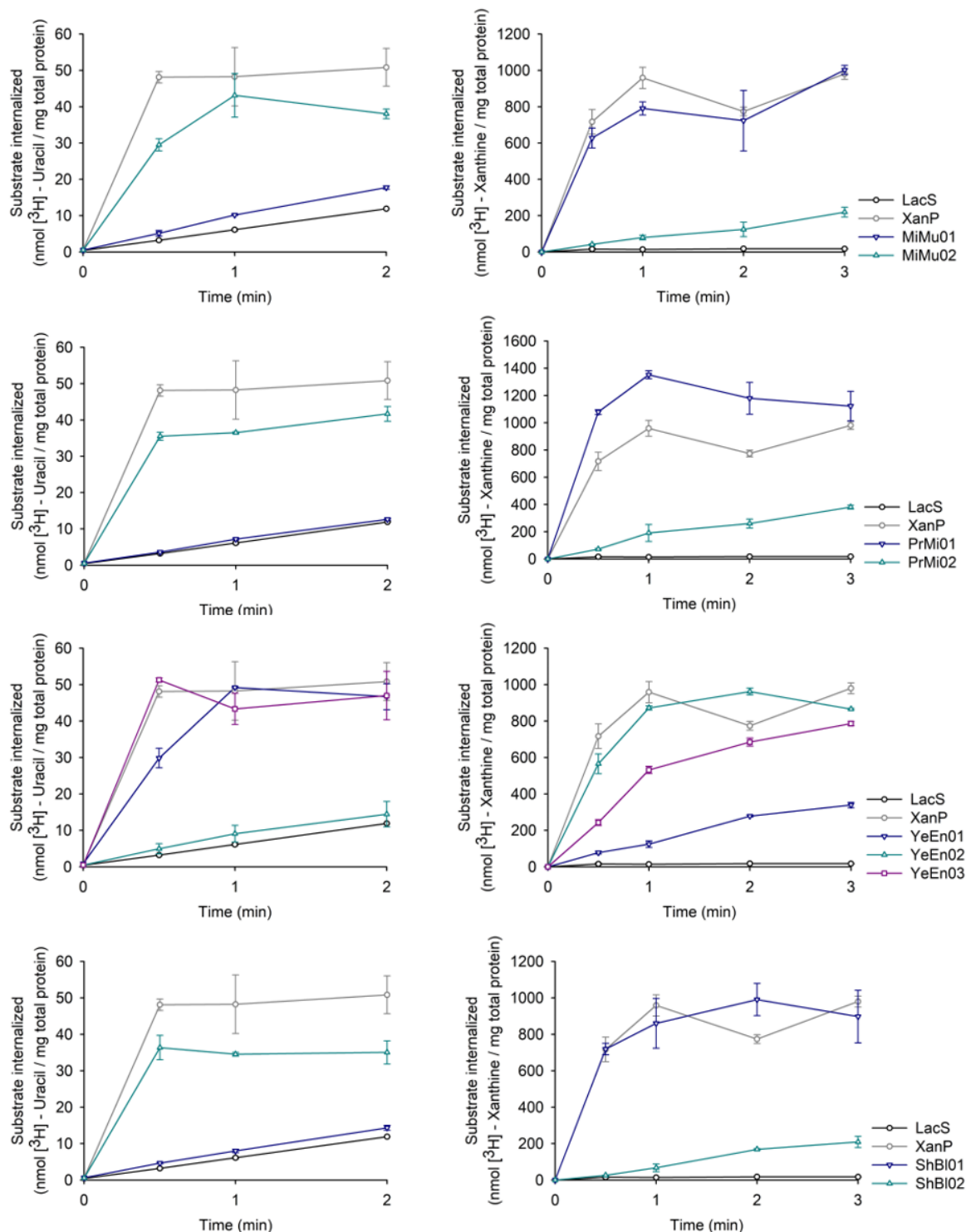
Figure S20: Uptake of uracil and xanthine into *E. coli* BW25113 *ΔuraA* (left) or *ΔxanP* (right) overexpressing homologs of UraA. All proteins were expressed as GFP fusions. In all cases, LacS was used as a negative control (black lines) and UraA or XanP as positive controls for uracil or xanthine uptake, respectively (grey lines). Annotated labels of homologs are abbreviations of the organisms they originate from and an arbitrary number (DeOr: *Desulfosporosinus orientis*, LaLa: *Lactococcus lactis*, IlPo: *Ilyobacter polytropus*, ClAs: *Clostridium asparagiforme*, DiIn: *Dialister invisus*, MiMu: *Mitsuokella multacida*, PrMi: *Proteus mirabilis*, YeEn: *Yersinia enterolitica*, ShBl: *Shimwellia blattae*). In all cases, cells were incubated at 30°C in 50 mM NaPi pH 6.0, 2 mM MgSO$_4$, 0.1 % (w/v) glucose, supplemented with 0.2 µM [^{3}H] - uracil or 5 µM partially radiolabeled xanthine. Uptakes were performed in triplicates, errors shown are standard deviations.
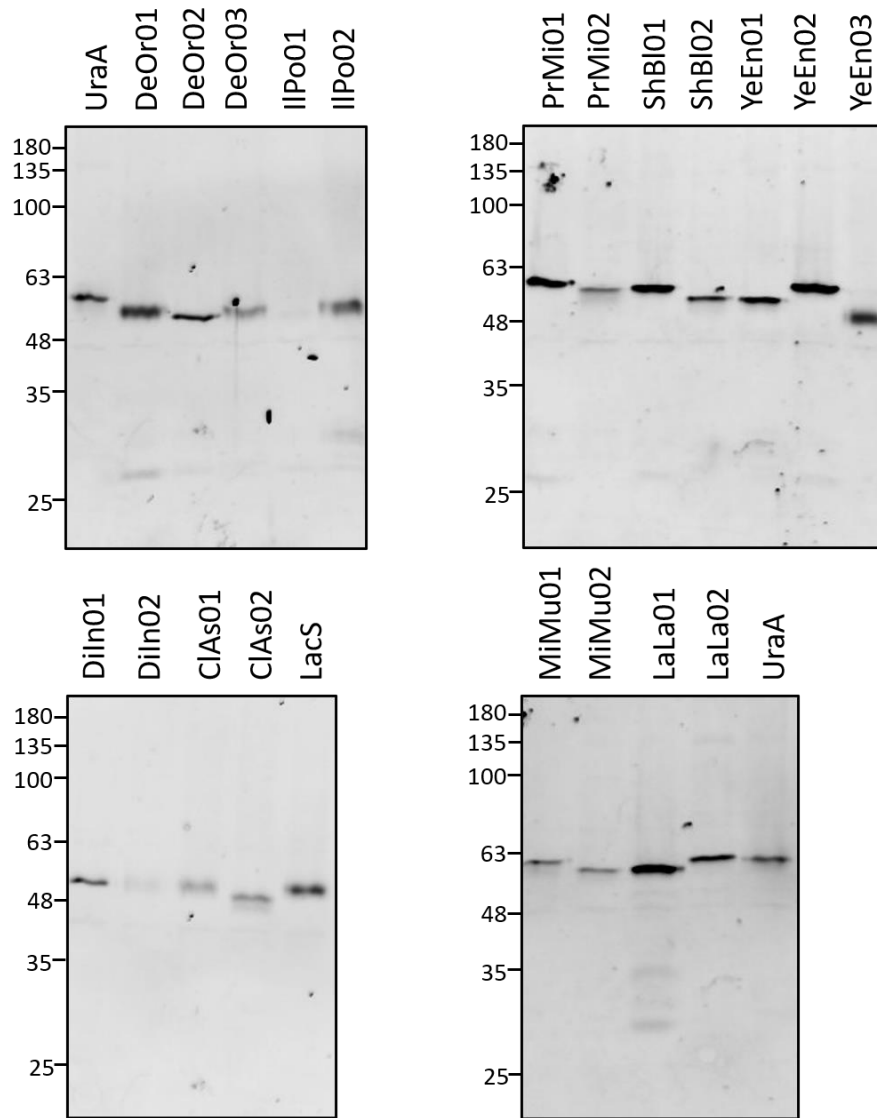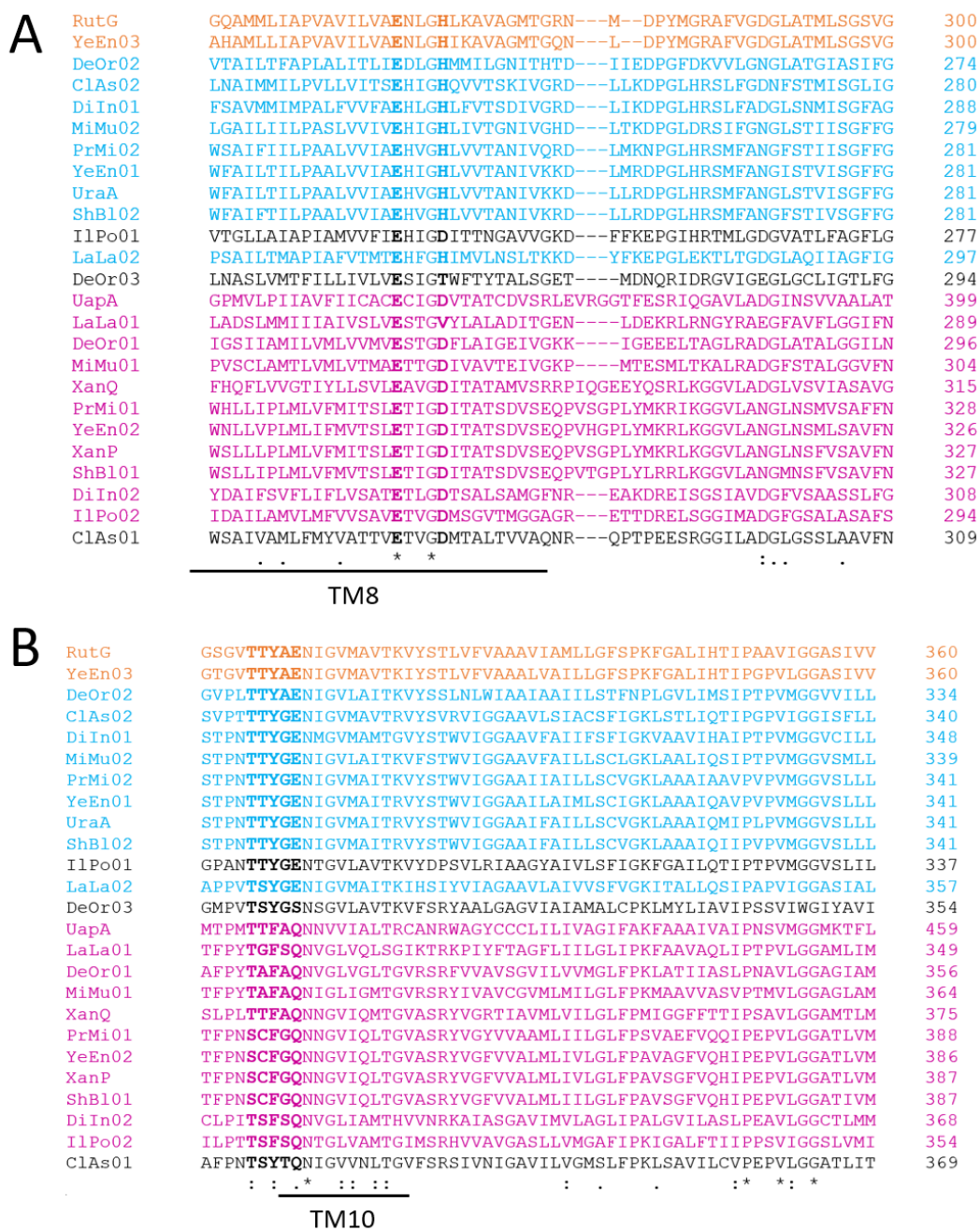
Figure S21: Expression of UraA homologs used for determination of substrate specificities. Samples from uptake of uracil into *E. coli* BW25113 *ΔuraA* expressing GFP fusions of different UraA homologs were analyzed for their expression levels. Whole cell lysates were subjected to 12% SDS-PAGE and proteins were visualized by *in gel* GFP fluorescence.

Figure S22: Sequence alignment of SLC23 homologs. (A) Sequence alignment of functionally analyzed SLC23 transporters, depicted are the residues found around TM8 of UraA. (B) Sequence alignment of functionally analyzed SLC23 transporters, depicted are the residues found around TM10 of UraA. The sequence alignment was generated using Clustal Ω, substrate binding site residues are shown in bold. Sequences are colored according to substrate specificity, with uracil transporters in teal, xanthine transporters in pink and homologs transporting both uracil and xanthine in orange.

Table 31: Sequence identities of SLC23 homologs. Sequence identities of the novel SLC23 uracil and/or xanthine transporters to UraA, RutG, XanQ, XanP or UapA were determined using the Blastp algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and are given in percentages with query coverage in brackets. The specificity that was determined for the homologs is stated by U for uracil and X for xanthine, respectively.

|	UraA	RutG	XanQ	XanP	UapA	Specificity
DeOr02	41.78 (99)	36.64 (97)	26.8 (81)	26.92 (97)	23.86 (89)	U
LaLa02	44.44 (96)	36.28 (97)	23.82 (90)	23.19 (98)	22.42 (87)	U
ClAs02	55.61 (97)	38.50 (92)	24.06 (80)	25.70 (83)	25.34 (65)	U
DiIn01	61.87 (97)	40.87 (95)	26.49 (80)	25.68 (95)	24.03 (85)	U
MiMu02	67.37 (98)	37.62 (91)	27.37 (60)	25.00 (92)	25.56 (89)	U
PrMi02	80.89 (99)	37.12 (93)	23.95 (77)	26.52 (92)	24.91 (92)	U
YeEn01	89.36 (98)	37.81 (90)	25.14 (78)	25.80 (86)	24.06 (67)	U
ShBl02	90.31 (99)	36.92 (96)	25.75 (96)	25.62 (96)	24.68 (68)	U
YeEn03	37.38 (91)	78.60 (96)	26.28 (84)	24.68 (96)	21.27 (84)	U/X
LaLa01	29.47 (92)	29.69 (97)	28.80 (96)	32.79 (93)	27.47 (84)	X
MiMu01	27.51 (83)	27.71 (97)	29.71 (96)	32.75 (97)	24.69 (93)	X
DeOr01	27.70 (82)	29.06 (86)	30.86 (95)	29.49 (96)	29.02 (92)	X
IlPo02	26.22 (84)	27.69 (94)	34.46 (96)	30.91 (95)	30.52 (94)	X
DiIn02	25.79 (81)	26.12 (83)	32.60 (96)	30.49 (94)	26.68 (94)	X
PrMi01	26.88 (83)	23.93 (88)	46.05 (95)	81.98 (98)	33.06 (94)	X
YeEn02	26.77 (83)	25.65 (88)	46.59 (95)	84.63 (100)	33.06 (95)	X
ShBl01	26.70 (92)	23.89 (84)	43.40 (95)	87.66 (99)	30.86 (96)	X
UraA	100 (100)	36.94 (93)	25.00 (78)	25.68 (93)	24.05 (67)	U
RutG	36.94 (93)	100 (100)	27.03 (84)	23.13 (84)	22.53 (94)	U/X
XanQ	25.92 (85)	27.03 (84)	100 (100)	45.05 (94)	29.31 (94)	X
XanP	25.68 (93)	24.44 (85)	45.15 (95)	100 (100)	24.44 (85)	X
UapA	23.49 (72)	22.77 (85)	29.31 (83)	31.43 (83)	100 (100)	X
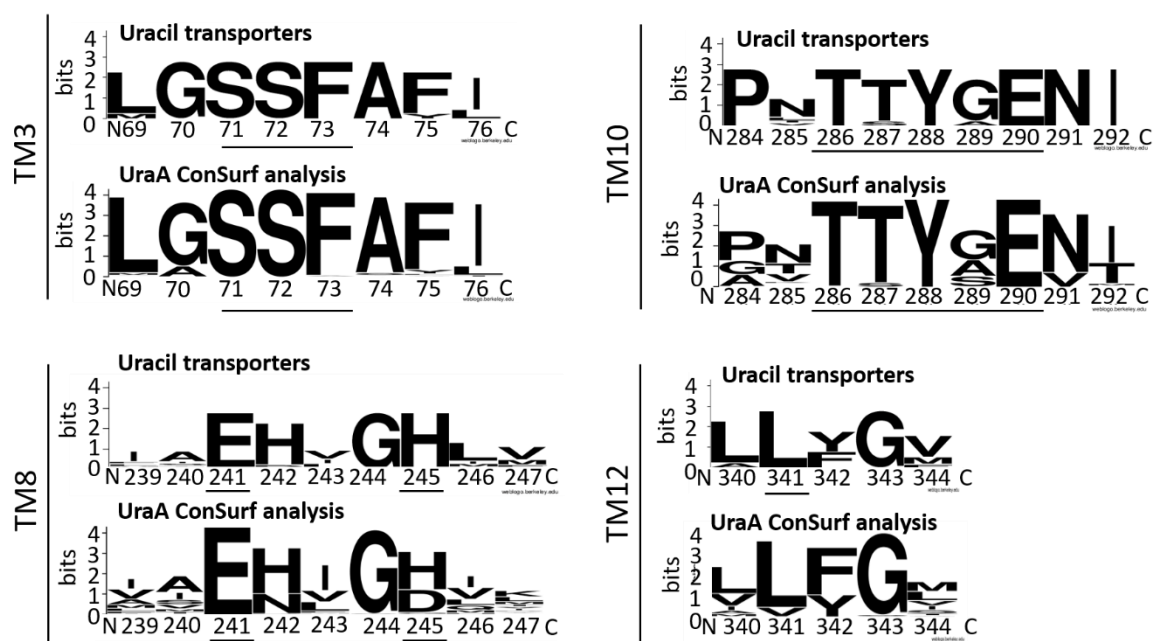
Figure S23: **Sequence conservation of substrate binding site residues in uracil transporters and comparison to functionally characterized uracil transporters.** Sequence alignments of uracil transporters identified in this study were generated using ClustalΩ. Alignments of relevant positions (within 5 Å of uracil in the UraA structure, underlined) are depicted as sequence logos generated using WebLogo (https://weblogo.berkeley.edu/logo.cgi) (Schneider and Stephens, 1990). In comparison, a ConSurf analysis on 500 UraA homologs (sequence identities 36.34% - 90.95%) was performed and relevant residues again are depicted as sequence logos.
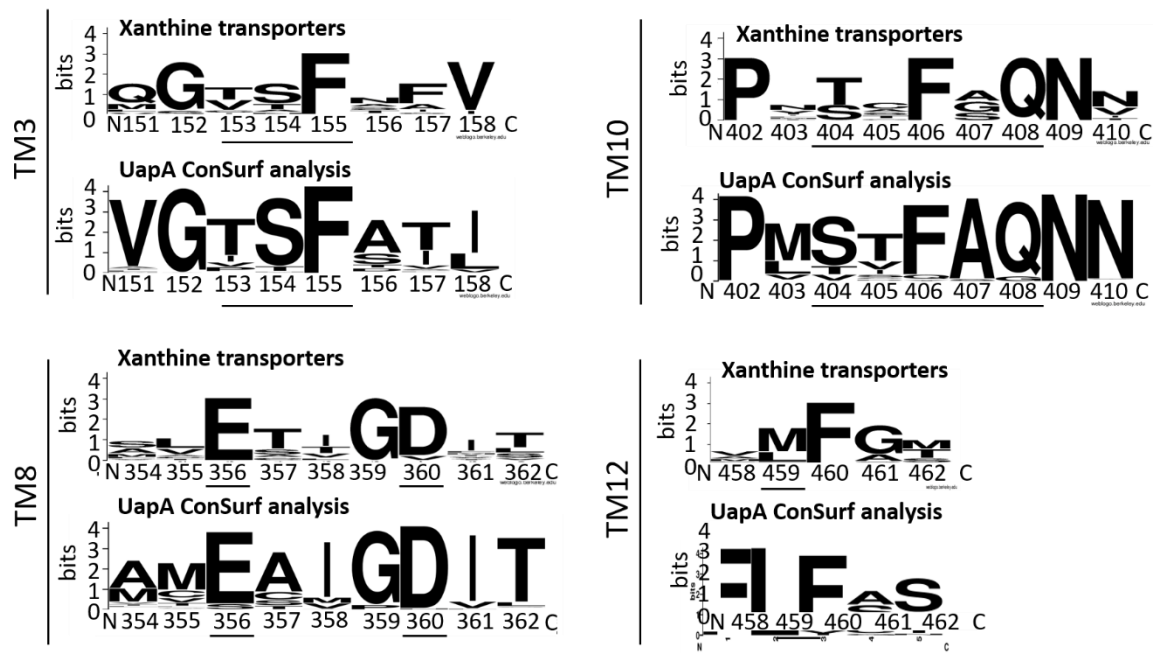

Figure S24: Sequence conservation of substrate binding site residues in xanthine transporters and comparison to functionally characterized xanthine transporters. Sequence alignments of xanthine transporters identified in this study were generated using ClustalΩ. Alignments of relevant positions (within 5 Å of xanthine in the UapA structure, underlined) are depicted as sequence logos generated using WebLogo (https://weblogo.berkeley.edu/logo.cgi) (Schneider and Stephens, 1990). In comparison, a ConSurf analysis on 500 UapA homologs (sequence identities 46.12% - 89.02%) was performed and relevant residues again are depicted as sequence logos.
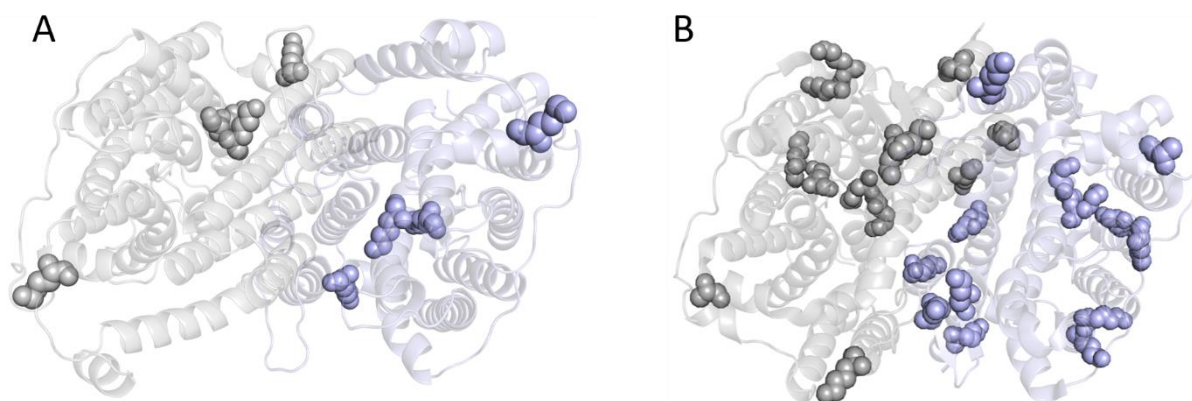
Figure S25: Location of lysines in 51ThXi and UraA. (A) Location of four of five lysine residues in 51ThXi. A Phyre2 model was generated for 51ThXi based on UraA (3QE7) which was used to build a dimer model based on the UraA dimer (5XLS). 51ThXi contains an additional lysine in a C-terminal amino acid stretch that is missing in UraA and therefore not included in the model. One protomer is colored in grey, the other in light blue, lysines are shown in spheres. (B) Location of lysines in the dimer of UraA (5XLS). One protomer is colored in grey, the other in light blue, lysines are shown in spheres.
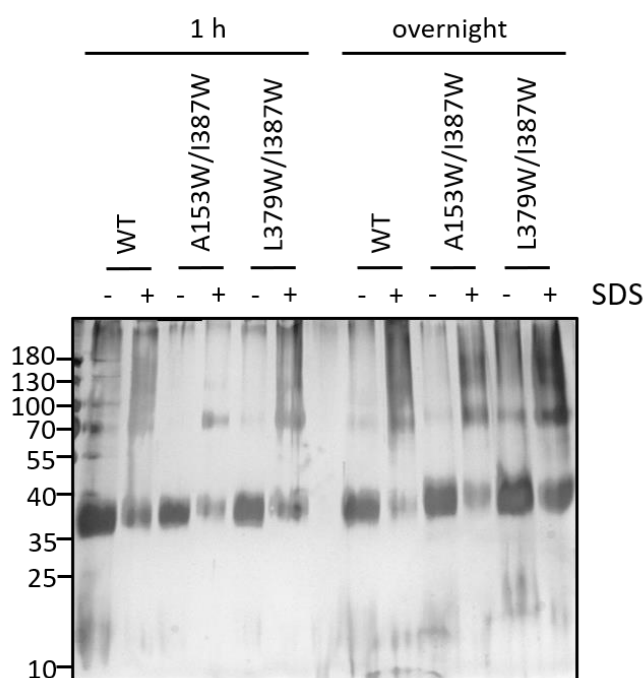
Figure S26: Unspecific crosslinking of 51ThXi tryptophan mutants in proteoliposomes. Unspecific crosslinking using glutaraldehyde on proteoliposomes of wild type 51ThXi, 51ThXi(A153W/I387W) or 51ThXi(L379W/I387W). A pre-incubation step with SDS is indicated by a plus (+) in contrast to not-preincubated samples, indicated by a minus (-). Proteoliposomes were incubated with 0.5% (w/v) glutaraldehyde either for 1 hour or overnight. Samples were subjected to 10% SDS-PAGE and proteins were visualized by silver staining.

51ThXi(cysl)
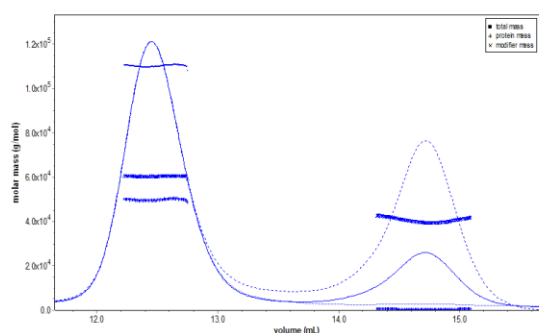
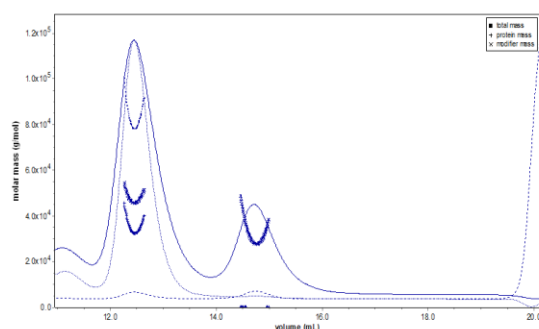M_W: 45294 Da
ε_0:   40910 M^{-1} cm^{-1}

	Volume (mL)	M_W (kDa)	Assignment
peak 1	12.226-12.752	49.85	51ThXi monomer
peak 2	14.314-15.095	0.39	DM micelle

51ThXi(A153W/I387W)

M_W: 45514 Da
ε_0:   51910 M^{-1} cm^{-1}

	Volume (mL)	M_W (kDa)	Assignment
peak 1	12.259-12.646	35.65	51ThXi monomer
peak 2	14.460-14.996	0.12	DM micelle

51ThXi(L379W/I387W)
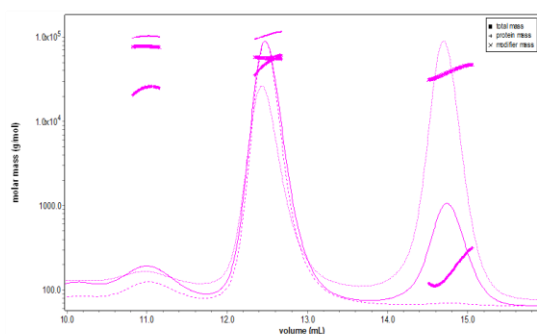
M_W: 45472 Da
ε_0:   51910 M^{-1} cm^{-1}

|	Volume (mL)	M_W (kDa)	Assignment
peak 1	10.816-11.152	245.90	
peak 2	12.343-12.679	48.22	51ThXi monomer
peak 3	14.511-15.061	0.16	DM micelle

Figure S27: SEC-MALS on 51ThXi and 51ThXi tryptophan mutants demonstrates its monomeric state in detergent. SEC-MALS on purified 51ThXi(cysl), 51ThXi(A153W/I387W) or 51ThXi(L379W/I387W) was performed on Superdex 200 10/300 GL Increase column, running buffer 50 mM KPi pH 7.5, 150 mM NaCl, 0.2% (w/v) DM, flowrate 0.5 mL/min. Light scattering was recorded using a miniDAWN TREOS instrument, the refractive index with an Optilab rEX. Measurements and data analysis as well as preparation of figures was done by Dr. Rupert Abele.
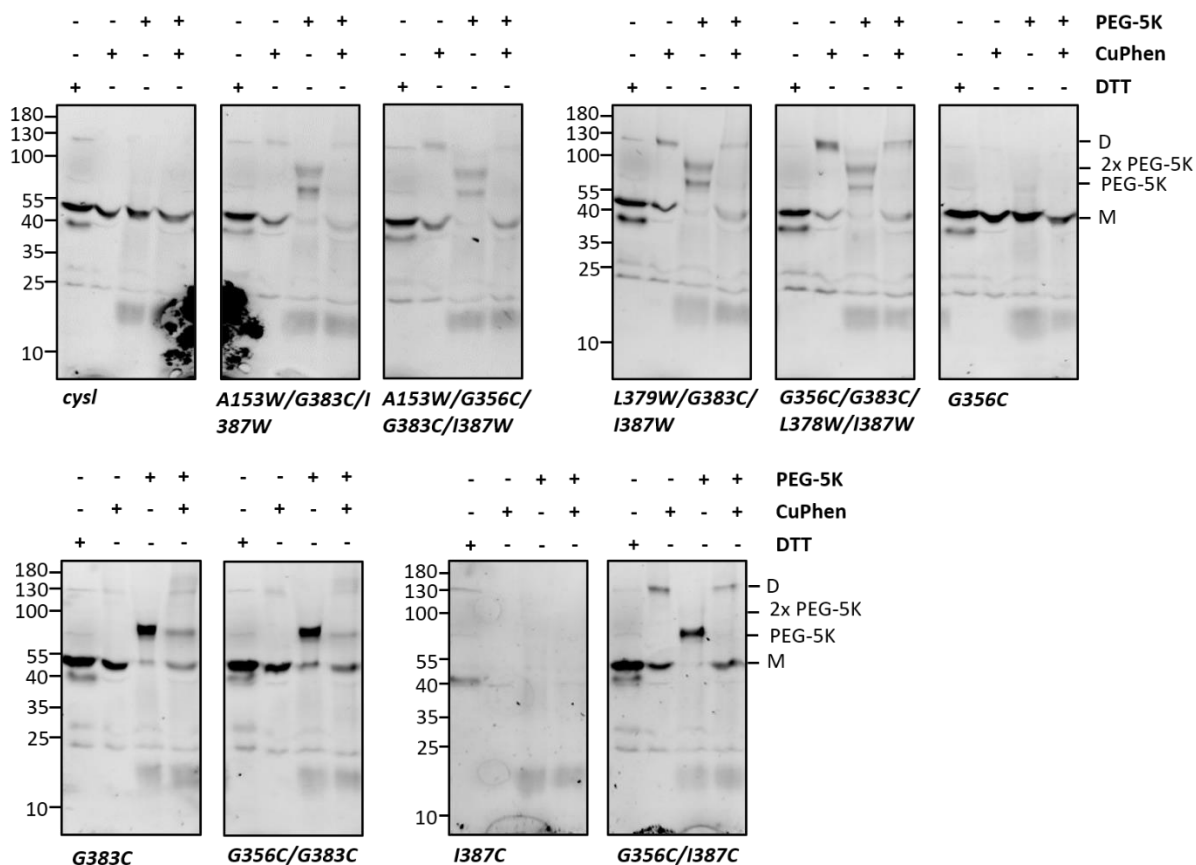
Figure S28: Crosslinking to assess the oligomeric state of the A153W/I387W and L379W/I387W mutants of 51ThXi. Cysteine-specific crosslinking in whole cell vesicles of GFP fusions of 51ThXi(A153/I387W) and 51ThXi(L379W/I387W) in combination with the interface crosslinking mutations G383C and G356C/G383C (**Figure 50**). Whole cells were disrupted in buffer containing 10 mM DTT that was subsequently removed using Biospin 6 columns, followed by incubation with 1 mM CuPhen and/or 1 mM PEG-5K + 1 % (w/v) SDS. Proteins were separated on 10% SDS-PAGE and visualized by *in gel* GFP fluorescence. Monomers and dimers are indicated by M and D, respectively, as well as one-fold and two-fold PEG-5K-labeled protein.
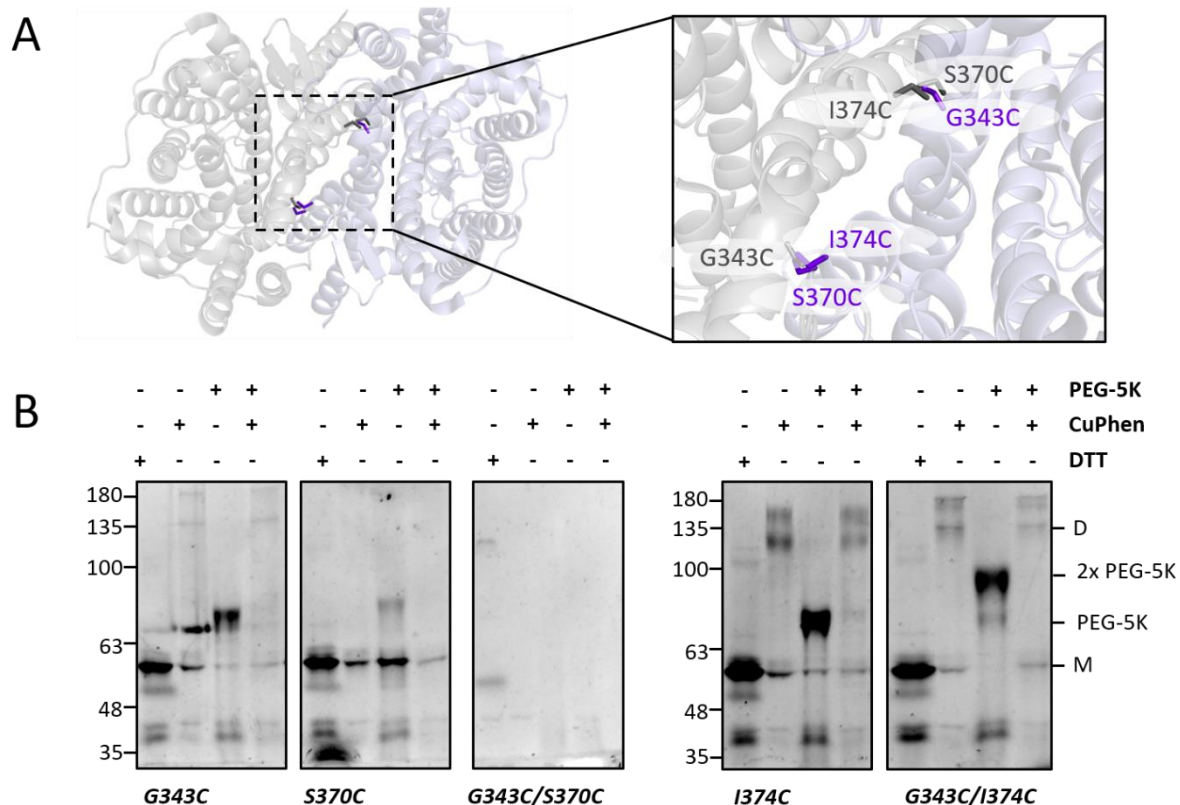
Figure S29: Crosslinking on cysteine mutants of UraA located at the dimer interface. (A) Location of cysteine mutations at the interface of the UraA dimer (PDB: 5XLS). Cβ-Cβ distances as measured by the built-in PyMOL tool: G343C-S370C, 2.8 Å; G343C-I374C, 4.2 Å; G343C-G343C, 18.9 Å; S370C-S370C, 21.2 Å; I374C-I374C, 17.2 Å. (B) Cysteine-specific crosslinking in whole cell vesicles of cysteine mutants of UraA expressed as GFP-fusion proteins. Cells were disrupted in buffer containing 10 mM DTT that was subsequently removed using Biospin 6 columns, followed by incubation with 1 mM CuPhen and/or 1 mM PEG-5K + 1 % (w/v) SDS. Proteins were separated on 8% SDS-PAGE and visualized by *in gel* GFP fluorescence. Monomers and dimers are indicated by M and D, respectively, as well as one-fold and two-fold PEG-5K-labeled protein.
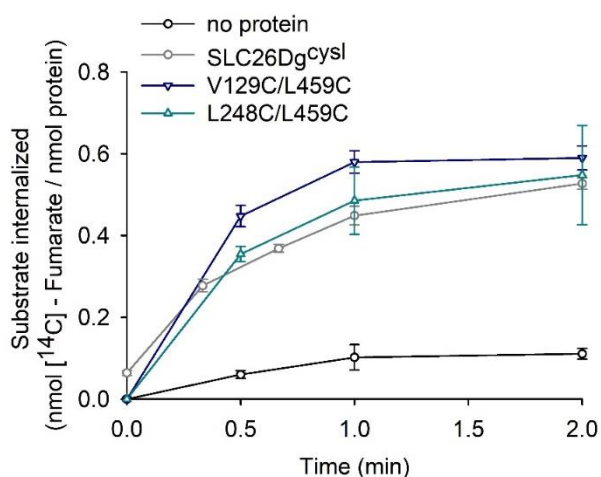
Figure S30: Uptake of SLC26Dg mutants used for PELDOR measurements. Uptake of 50 μM fumarate into proteoliposomes of cysteine-free SLC26Dg (SLC26Dg^{cysl}, grey) or the mutants used for PELDOR measurements (V129C mixed with L459C, dark blue; L248C mixed with L459C, teal) was measured at 30°C in presence of a pH gradient and an inside positive membrane potential. As negative control, liposomes containing no protein were used (black). Uptakes were measured in triplicates, errors shown represent standard deviations.

TM1

```
RutG     TSTESGVVAPDERLPFAQTAVMGVQHAVAMFGATVLMPILMGLDP-----------NLSI      62
UraA     --MTRRAIGVSERPPLLQTIPLSLQHLFAMFGATVLVPVLFHINP-----------ATVL      47
PlUp     --MKDRVIQVDERLPFLQSIPLSLQHLFAMFGSTVLVPMLLQINP-----------AICL      47
UapA     DPRAPPFFGLNEKIPVLLAFILGLQHALAMLAGVVTPPLIISSSLSLPSDLQQYLVSTSL     120
XanQ     AGSDLIF-ELEDRPPFHQALVGAITHLLAIFVPMVTPALIVGAALQLSAETTAYLVSMAM      65
UacT     SGQDRPTDEVDRILSPGKLIILGLQHVLVMYAGAVAVPLMIGDRLGLSKEAIAMLISSDL      71
PlUacP   ------------MRKSKVLTLGFQHVLAMYAGAVIVPLIVGSSLKLNAEQLAYLVSIDL      47
                            .. *  ..:      *    ::.                     :
```

```
RutG     LMSGIGTLLFFF---------ITGGRVPSYLGSSAAFVGVVIAATGFNG-----------     102
UraA     LFNGIGTLLYLF---------ICKGKIPAYLGSSFAFISPVLLLL-P-------------      84
PlUp     LMNGIGTLIYIF---------LCKGRIPAYLGSSFAFISPVLIVIST------------      85
UapA     IVCGLLSMVQITRFHIYKTPYYIGSGVLSVMGVSFSIISVASGAFNQMYSNGFCQLDEAG     180
XanQ     IASGIGTWLQVN------RYGIVGSGLLSIQSVNFSFVTVMIALGSSMKSDGFHE-----     114
UacT     FCCGIVTLLQCI-----GIGRFMGIRLPVIMSVTFAAVTPMIAIGMNPD----------     115
PlUacP   LTCGIATLLQVW------RNKFFGIGLPVMLGCTFTAVGPMIAIGS--E----------      88
          :   *: : :              :      . . : :
```

TM8

```
RutG     VAVILVAENLGHLKAVAGMTGRN--MD---PYMGRAFVGDGLATMLSGSVGGSGVTTYAE     309
UraA     AALVVIAEHVGHLVVTANIVKKDLLRD---PGLHRSMFANGLSTVISGFFGSTPNTTYGE     290
PlUp     AALVVVAEHIGHLIVTSNIVGKDLSKD---PGLDRSLLGNGVSTVISSFVGSTPNTTYGE     294
UapA     VFIICACECIGDVTATCDVSRLEVRGGTFESRIQGAVLADGINSVVAALATMTPMTTFAQ     408
XanQ     IYLLSVLEAVGDITATAMVSRRPIQGEEYQSRLKGGVLADGLVSVIASAVGSLPLTTFAQ     324
UacT     VLIIVFIESMGMFLALGEIVGRKL----SSHDIIRGLRVDGVGTMIGGTFNSFPHTSFSQ     318
PlUacP   VAIVSVMESTGVFVALGKITEKEL----TADDLKRGYRSEGLASILGSIMNSFPYTTYSQ     291
          ::    *   *  . .   :        :  .   :*: :::..      *:: .:
```

TM10

```
RutG     NIGVMAVTKVYSTLVFVAAAVIAMLLGFSPKFGALIHTIPAAVIGGASIVVFGLIAVAGA     369
UraA     NIGVMAITRVYSTWVIGGAAIFAILLSCVGKLAAAIQMIPLPVMGGVSLLLYGVIGASGI     350
PlUp     NIGVLALTRVYSIWIIGGAAVMAIVLSFVGKLAALIQTIPVPVMGGVSILLFGVIAGSGV     354
UapA     NNVVIALTRCANRWAGYCCCLILIVAGIFAKFAAAIVAIPNSVMGGMKTFLFASVVISGQ     468
XanQ     NNGVIQMTGVASRYVGRTIAVMLVILGLFPMIGGFFTTIPSAVLGGAMTLMFSMIAIAGI     384
UacT     NVGLVSVTRVHSRWVCISSGIILILFGMVPKMAVLVASIPQFVLGGAGLVMFGMVLATGI     378
PlUacP   NVGLIQISKVKSRDVVITAGFILVILGFMPKIAALTLLIPTAVLGGAMIAMFGMVVSSGI     351
```

Figure S31: Putative substrate binding sites of the multifunctional SLC23 transporters PlUP and PlUacP from *Paenibacillus larvae*. Sequence alignements on PlUP (NCBI reference sequence: WP_023482778.1) and PlUacP (NCBI reference sequence: WP_023485062.1) from *Paenibacillus larvae* and *E. coli* UraA, RutG, XanQ and XanP as well as *Aspergillus nidulans* UapA was performed using Clustal Q. Relevant substrate binding site residues, based on the structures of UraA and UapA, are indicated in bold letters.

References

Abbas YM, Toye AM, Rubinstein JL, Reithmeier RAF. 2018. Band 3 function and dysfunction in a structural context. *Curr Opin Hematol* **25**:163–170. doi:10.1097/MOH.0000000000000418

Abbott JA, Livingston NM, Egri SB, Guth E, Francklyn CS. 2017. Characterization of aminoacyl-tRNA synthetase stability and substrate interaction by differential scanning fluorimetry. *Methods* **113**:64–71. doi:10.1016/j.ymeth.2016.10.013

Alguel Y, Amillis S, Leung J, Lambrinidis G, Capaldi S, Scull NJ, Craven G, Iwata S, Armstrong A, Mikros E, Diallinas G, Cameron AD, 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

Alper SL. 2009. Molecular physiology and genetics of Na+-independent SLC4 anion exchangers. *J Exp Biol* **212**:1672–1683. doi:10.1242/jeb.029454

Alper SL, Sharma AK. 2013. The SLC26 Gene Family of Anion Transporters and Channels. *Mol Aspects Med* **34**:494–515. doi:10.1016/j.mam.2012.07.009

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**:3389–3402. doi:10.2503/jjshs.58.977

Andersen PS, Frees D, Fast R, Mygind B. 1995. Uracil Uptake in Escherichia coli K-12: Isolation of uraA Mutants and Cloning of the Gene. *J Bacteriol* **177**:2008–13. doi:10.1128/jb.177.8.2008-2013.1995

Andersen PS, Smith JM, Mygind B. 1992. Characterization of the upp gene encoding uracil phosphoribosyltransferase of Escherichia coli K12. *Eur J Biochem* **204**:51–56. doi:10.1111/j.1432-1033.1992.tb16604.x.

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, Iwata S. 2015. Crystal structure of the anion exchanger domain of human erythrocyte band 3. *Science (80-)* **350**:680–684. doi:10.1126/science.aaa433

Arkhipova V, Guskov A, Slotboom DJ. 2020. Structural ensemble of a glutamate transporter homologue in lipid nanodisc environment. *Nat Commun* **11**. doi:10.1038/s41467-020-14834-8

References

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol Syst Biol* **2**. doi:10.1038/msb4100050

Babu M, Greenblatt JF, Emili A, Strynadka NCJ, Reithmeier RAF, Moraes TF. 2010. Structure of a SLC26 Anion Transporter STAS Domain in Complex with Acyl Carrier Protein: Implications for E. coli YchM in Fatty Acid Metabolism. *Structure* **18**:1450–1462. doi:10.1016/j.str.2010.08.015

Bagnara AS, Finch LR. 1974. The Effects of Bases and Nucleosides on the Intracellular Contents of Nucleotides and 5-Phosphoribosyl 1-pyrophosphate in Escherichia coli. *Eur J Biochem* **41**:421–430. doi:10.1111/j.1432-1033.1974.tb03283.x

Bai N, Roder H, Dickson A, Karanicolas J. 2019. Isothermal Analysis of ThermoFluor Data can readily provide Quantitative Binding Affinities. *Sci Rep* **9**:1–15. doi:10.1038/s41598-018-37072-x

Bai X, Moraes TF, Reithmeier RAF. 2017. Structural biology of solute carrier (SLC) membrane transport proteins. *Mol Membr Biol* **34**:1–32. doi:10.1080/09687688.2018.1448123

Bai X, Moraes TF, Reithmeier RAF. 2016. Effect of SLC26 anion transporter disease-causing mutations on the stability of the homologous STAS domain of E. coli DauA (YchM). *Biochem J* **473**:615–626. doi:10.1042/BJ20151025

Baldwin SA, Beal PR, Yao SYM, King AE, Cass CE, Young JD. 2004. The equilibrative nucleoside transporter family , SLC29. *Pflügers Arch – Eur J Physiol* **447**:735–743. doi:10.1007/s00424-003-1103-2

Barchue J, Symersky J, Narayana SVL, Moore JK, DeLucas LJ, Chattopadhyay D. 1999. Expression, purification, crystallization and preliminary x-ray diffraction analysis of uracil phosphoribosyltransferase of Toxoplasma gondii. *Acta Crystallogr Sect D Biol Crystallogr* **55**:347–349. doi:10.1107/S090744499800821X

Beck CF, Ingraham JL, Neuhard J. 1972. Location on the chromosome of Salmonella typhinurium of genes governing pyrimidine metabolism - II. Uridine kinase, cytosine deaminase and thymidine kinase. *MGG Mol Gen Genet* **115**:208–215. doi:10.1007/BF00268884

Bocian-Ostrzycka KM, Grzeszczuk MJ, Banaś AM, Jagusztyn-Krynicka EK. 2017. Bacterial thiol oxidoreductases — from basic research to new antibacterial strategies. *Appl Microbiol Biotechnol* **101**:3977–3989. doi:10.1007/s00253-017-8291-8

References

Boggavarapu R, Jeckelmann JM, Harder D, Schneider P, Ucurum Z, Hediger M, Fotiadis D. 2013. Expression, Purification and Low-Resolution Structure of Human Vitamin C Transporter SVCT1 (SLC23A1). *PLoS One* **8**:e76427. doi:10.1371/journal.pone.0076427

Bondar A-N, White SH. 2013. Hydrogen Bond Dynamics in Membrane Protein. *Biochim Biophys Acta* **1818**:942–950. doi:10.1016/j.bbamem.2011.11.035.

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* **108**:204–219. doi:10.1111/mmi.13931

Boudker O, Ryan RM, 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. doi:10.1038/nature05455

Burton K. 1977. Transport of adenine, hypoxanthine and uracil into Escherichia coli. *Biochem J* **168**:195–204. doi:10.1042/bj1680195

Burtscher V, Schicker K, Freissmuth M, Sandtner W. 2019. Kinetic models of secondary active transporters. *Int J Mol Sci* **20**. doi:10.3390/ijms20215365

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

Călinescu O, Dwivedi M, Patiño-Ruiz M, Padan E, Fendler K. 2017. Lysine 300 is essential for stability but not for electrogenic transport of the Escherichia coli NhaA Na+/H+ antiporter. *J Biol Chem* **292**:7932–7941. doi:10.1074/jbc.M117.778175

Campagnaro GD, de Koning HP. 2020. Purine and pyrimidine transporters of pathogenic protozoa – conduits for therapeutic agents. *Med Res Rev* **40**:1679–1714. doi:10.1002/med.21667

Carro MD, Falkenstein E, Blemings KP, Klandorf H. 2009. Determination of xanthine oxidoreductase activity in broilers: Effect of pH and temperature of the assay and distribution in tissues. *Poult Sci* **88**:2406–2414. doi:10.3382/ps.2009-00278

References

Carter D, Donald RGK, Roos D, Ullman B. 1997. Expression, purification, and characterization of uracil phosphoribosyltransferase from Toxoplasma gondii. *Mol Biochem Parasitol* **87**:137–144. doi:10.1016/S0166-6851(97)00058-3

Casadaban MJ, Cohen SN. 1980. Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *J Mol Biol* **138**:179–207. doi:10.1016/0022-2836(80)90283-1

Casiraghi A, Bensimon A, Superti-Furga G. 2021. Recent developments in ligands and chemical probes targeting solute carrier transporters. *Curr Opin Chem Biol* **62**:53–63. doi:10.1016/j.cbpa.2021.01.012

César-Razquin A, Snijder B, Frappier-Brinton T, Isserlin R, Gyimesi G, Bai X, Reithmeier RA, Hepworth D, Hediger MA, Edwards AM, Superti-Furga G. 2015. A Call for Systematic Research on Solute Carriers. *Cell* **162**:478–487. doi:10.1016/j.cell.2015.07.022

Chaliotis A, Vlastaridis P, Ntountoumi C, Botou M, Yalelis V, Lazou P, Tatsaki E, Mossialos D, Frillingos S, Amoutzias GD. 2018. NAT/NCS2-hound: a webserver for the detection and evolutionary classification of prokaryotic and eukaryotic nucleobase-cation symporters of the NAT/NCS2 family. *Gigascience* **7**:1–8. doi:10.1093/gigascience/giy133

Chang Y-N, Geertsma ER. 2017. The novel class of seven transmembrane segment inverted repeat carriers. *Biol Chem* **398**:165–174. doi:10.1515/hsz-2016-0254

Chang Y-N, Jaumann EA, Reichel K, Hartmann J, Oliver D, Hummer G, Joseph B, Geertsma ER. 2019. Structural basis for functional interactions in dimers of SLC26 transporters. *Nat Commun* **10**:2032. doi:10.1038/s41467-019-10001-w

Cherest H, Davidian J, Benes V, Ansorge W, Surdin-kejan Y. 1997. Molecular Characterization of Two High Affinity Sulfate Transporters in Saccharomyces cerevisiae. *Genetics* **145**:627–635.

Chernova MN, Jiang L, Crest M, Hand M, Vandorpe DH, Strange K, Alper SL. 1997. Electrogenic sulfate/chloride exchange in Xenopus oocytes mediated by murine AE1 E699Q. *J Gen Physiol* **109**:345–360. doi:10.1085/jgp.109.3.345

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

References

Choi I. 2012. SLC4A Transporters. *Curr Top Membr* **70**:77–103. doi:10.1016/B978-0-12-394316-3.00003-X.

Christopherson RI, Finch LR. 1978. Response of the Pyrimidine Pathway of Escherichia coli K12 to Exogenous Adenine and Uracil. *Eur J Biochem* **90**:347–358. doi:10.1111/j.1432-1033.1978.tb12611.x

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

Compton ELR, Karinou E, Naismith JH, Gabel F, Javelle A. 2011. Low resolution structure of a bacterial SLC26 transporter reveals dimeric stoichiometry and mobile intracellular domains. *J Biol Chem* **286**:27058–27067. doi:10.1074/jbc.M111.244533

Corpe CP, Eck P, Wang J, Al-Hasani H, Levine M. 2013. Intestinal dehydroascorbic acid (DHA) transport mediated by the facilitative sugar transporters, GLUT2 and GLUT8. *J Biol Chem* **288**:9092–9101. doi:10.1074/jbc.M112.436790

Corpe CP, Tu H, Eck P, Wang J, Faulhaber-Walter R, Schnermann J, Margolis S, Padayatty S, Sun H, Wang Y, Nussbaum RL, Espey MG, Levine M. 2010. Vitamin C transporter Slc23a1 links renal reabsorption, vitamin C tissue accumulation, and perinatal survival in mice. *J Clin Invest* **120**:1069–1083. doi:10.1172/JCI39191

Coudray N, L. Seyler S, Lasala R, Zhang Z, Clark KM, Dumont ME, Rohou A, Beckstein O, Stokes DL. 2017. Structure of the SLC4 transporter Bor1p in an inward-facing conformation. *Protein Sci* **26**:130–145. doi:10.1002/pro.3061

Dagert M, Ehrlich SD. 1979. Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. *Gene* **6**:23–28. doi:10.1016/0378-1119(79)90082-9

Dallos P, Fakler B. 2002. Prestin, a new type of motor protein. *Nat Rev Mol Cell Biol* **3**:104–111. doi:10.1038/nrm730

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**:6640–6645. doi:10.1073/pnas.120163297

References

De Koning H, Diallinas G. 2000. Nucleobase transporters. *Mol Membr Biol* **75**:75–94. doi:10.1080/09687680050117101

De Koning HP, Jarvis SM. 1998. A highly selective , high-affinity transporter for uracil in Trypanosoma brucei brucei : evidence for proton-dependent transport. *Biochem Cell Biol* **76**:853–858. doi:10.1080/09687680050117101

Del Arco J s, Fernández-Lucas J. 2017. Purine and Pyrimidine Phosphoribosyltransferases: A Versatile Tool for Enzymatic Synthesis of Nucleoside-5'-monophosphates. *Curr Pharm Des* **23**:6898–6912. doi:10.2174/1381612823666171017165707

Detro-Dassen S, Schänzler M, Lauks H, Martin I, Meyer zu Berstenhorst S, Nothmann D, Torres-Salazar D, Hidalgo P, Schmalzing G, Fahlke C. 2008. Conserved dimeric subunit stoichiometry of SLC26 multifunctional anion exchangers. *J Biol Chem* **283**:4177–4188. doi:10.1074/jbc.M704924200

Deutscher J, Francke C, Postma PW. 2006. How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. *Microbiol Mol Biol Rev* **70**:939–1031. doi:10.1128/MMBR.00024-06

Diallinas G. 2016. Dissection of Transporter Function: From Genetics to Structure. *Trends Genet* **32**:576–590. doi:10.1016/j.tig.2016.06.003

Diallinas G. 2013. Allopurinol and xanthine use different translocation mechanisms and trajectories in the fungal UapA transporter. *Biochimie* **95**:1755–1764. doi:10.1016/j.biochi.2013.05.013

Diallinas G, Gorfinkiel L, Arst HN, Cecchetto G, Scazzocchio C. 1995. Genetic and molecular characterization of a gene encoding a wide specificity purine permease of Aspergillus nidulans reveals a novel family of transporters conserved in prokaryotes and eukaryotes. *J Biol Chem* **270**:8610–8622. doi:10.1007/BF02814078.

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

Diasio RB, Harris BE. 1989. Clinical Pharmacology of 5-fluorouracil. *Clin Pharmacokinet* **16**:215–237. doi:10.2165/00003088-198916040-00002

References

Dorwart MR, Shcheynikov N, Baker JMR, Forman-Kay JD, Muallem S, Thomas PJ. 2008. Congenital chloride-losing diarrhea causing mutations in the STAS domain result in misfolding and mistrafficking of SLC26A3. *J Biol Chem* **283**:8711–8722. doi:10.1074/jbc.M704328200

Drew D, Boudker O. 2016. Shared Molecular Mechanisms of Membrane Transporters. *Annu Rev Biochem* **85**:543–572. doi:10.1146/annurev-biochem-060815-014520

Drew D, Slotboom D-J, Friso G, Reda T, Genevaux P, Rapp M, Meindl-Beinker NM, Lambert W, Lerch M, Daley DO, Van Wijk K-J, Hirst J, Kunji E, De Gier J-W. 2005. A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein Sci* **14**:2011–2017. doi:10.1110/ps.051466205

El-Awady RA, Saleh EM, Dahm-Daphi J. 2010. Targeting DNA double-strand break repair: Is it the right way for sensitizing cells to 5-fluorouracil? *Anticancer Drugs* **21**:277–287. doi:10.1097/CAD.0b013e328334b0ae

Erkens GB, Hänelt I, Goudsmits JMH, Slotboom DJ, Van Oijen AM. 2013. Unsynchronised subunit motion in single trimeric sodium-coupled aspartate transporters. *Nature* **502**:119–123. doi:10.1038/nature12538

Ficici E, Faraldo-Gómez JD, Jennings ML, Forrest LR. 2017. Asymmetry of inverted-topology repeats in the AE1 anion exchanger suggests an elevator-like mechanism. *J Gen Physiol* **149**:1149–1164. doi:10.1085/jgp.201711836 1

Forloni M, Liu AY, Wajapeyee N. 2019. Megaprimer polymerase chain reaction (PCR)-based mutagenesis. *Cold Spring Harb Protoc* **2019**:513–515. doi:10.1101/pdb.prot097824

Forrest LR. 2015. Structural Symmetry in Membrane Proteins. *Annu Rev Biophys* **44**:311–337. doi:10.1146/annurev-biophys-051013-023008.

Fredriksson R, Nordström KJV, Stephansson O, Hägglund MGA, Schiöth HB. 2008. The solute carrier (SLC) complement of the human genome: Phylogenetic classification reveals four major families. *FEBS Lett* **582**:3811–3816. doi:10.1016/j.febslet.2008.10.016

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.

References

Gagnon KB, Delpire E. 2021. Sodium Transporters in Human Health and Disease. *Front Physiol* **11**:1–18. doi:10.3389/fphys.2020.588664

Galperin MY, Kristensen DM, Makarova KS, Wolf YI, Koonin E V. 2018. Microbial genome analysis: The COG approach. *Brief Bioinform* **20**:1063–1070. doi:10.1093/bib/bbx117

Garaeva AA, Guskov A, Slotboom DJ, Paulino C. 2019. A one-gate elevator mechanism for the human neutral amino acid transporter ASCT2. *Nat Commun* **10**:1–8. doi:10.1038/s41467-019-11363-x |

Garaeva AA, Slotboom DJ. 2020. Elevator-type mechanisms of membrane transport. *Biochem Soc Trans* **48**:1227–1241. doi:10.1042/BST20200290

Geertsma ER. 2013. FX cloning: A versatile high-throughput cloning system for characterization of enzyme variants. *Methods Mol Biol* **978**:133–148. doi:10.1007/978-1-62703-293-3_10

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

Geertsma ER, Dutzler R. 2011. A versatile and efficient high-throughput cloning tool for structural biology. *Biochemistry* **50**:3272–3278. doi:10.1021/bi200178z

Geertsma ER, Duurkens RH, Poolman B. 2005. The activity of the lactose transporter from Streptococcus thermophilus is increased by phosphorylated IIA and the action of β-galactosidase. *Biochemistry* **44**:15889–15897. doi:10.1021/bi051638w

Geertsma ER, Groeneveld M, Slotboom DJ, Poolman B. 2008a. Quality control of overexpressed membrane proteins. *Proc Natl Acad Sci U S A* **105**:5722–5727. doi:10.1073/pnas.0802190105

Geertsma ER, Mahmood NNAB, Schuurman-Wolters GK, Poolman B. 2008b. Membrane reconstitution of ABC transporters and assays of translocator function. *Nat Protoc* **3**:256–266. doi:10.1038/nprot.2007.519

Ghim SY, Neuhard J. 1994. The pyrimidine biosynthesis operon of the thermophile Bacillus caldolyticus includes genes for uracil phosphoribosyltransferase and uracil permease. *J Bacteriol* **176**:3698–3707. doi:10.1128/jb.176.12.3698-3707.1994

References

Gmeiner W. 2012. Novel Chemical Strategies for Thymidylate Synthase Inhibition. *Curr Med Chem* **12**:191–202. doi:10.2174/0929867053363432

Godoy A, Ormazabal V, Moraga-Cid G, Zúñiga FA, Sotomayor P, Barra V, Vasquez O, Montecinos V, Mardones L, Guzmán C, Villagrán M, Aguayo LG, Oñate SA, Reyes AM, Cárcamo JG, Rivas CI, Vera JC. 2007. Mechanistic insights and functional determinants of the transport cycle of the ascorbic acid transporter SVCT2: Activation by sodium and absolute dependence on bivalent cations. *J Biol Chem* **282**:615–624. doi:10.1074/jbc.M608300200

Gong J, Chen Y, Pu F, Sun P, He F, Zhang L, Li Y, Ma Z, Wang H. 2019. Understanding Membrane Protein Drug Targets in Computational Perspective. *Curr Drug Targets* **20**:551–564. doi:10.2174/1389450120666181204164721 CITATIONS

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. doi:10.1080/09687860500093016

Greenlee L, Handler P. 1964. Xanthine Oxidase. Vi. Influence of Ph on Substrate Specificity. *J Biol Chem* **239**:1090–1095. doi:10.1016/S0021-9258(18)91395-5

Gudin S, Quashie NB, Candlish D, Al-salabi MI, Jarvis SM, Ranford-cartwright LC, Koning HP De. 2006. Trypanosoma brucei : A survey of pyrimidine transport activities. *Exp Parasitol* **114**:118–125. doi:10.1016/j.exppara.2006.02.018

Hallworth R, Stark K, Zholudeva L, Currall B, Nichols MG. 2013. The Conserved Tetrameric Subunit Stoichiometry of SLC26 Proteins. *Microsc Microanal* **19**:799–807. doi:10.1017/S1431927613000457. The

Hawkesford MJ. 2003. Transporter gene families in plants: The sulphate transporter gene family - Redundancy or specialization? *Physiol Plant* **117**:155–163. doi:10.1034/j.1399-3054.2003.00034.x

Hazes B, Dijkstra BW. 1988. Model building of disulfide bonds in proteins with known three-dimensional structure. *Protein Eng Des Sel* **2**:119–125. doi:10.1093/protein/2.2.119

Hediger MA, Clémençon B, Burrier RE, Bruford EA. 2013. The ABCs of membrane transporters in health and disease (SLC series): Introduction. *Mol Aspects Med* **34**:95–107. doi:10.1016/j.mam.2012.12.009

References

Hochstadt-Ozer J, Stadtman ER. 1971. The Regulation of Purine Utilization in Bacteria. *J Biol Chem* **246**:5294–5303. doi:10.1016/s0021-9258(18)61906-4

Höglund PJ, Nordström KJV, Schiöth HB, Fredriksson R. 2011. The solute carrier families have a remarkably long evolutionary history with the majority of the human families present before divergence of Bilaterian species. *Mol Biol Evol* **28**:1531–1541. doi:10.1093/molbev/msq350

Holman GD. 2020. Structure, function and regulation of mammalian glucose transporters of the SLC2 family. *Pflugers Arch Eur J Physiol* **472**:1155–1175. doi:10.1007/s00424-020-02411-3

Holzhüter K, Geertsma ER. 2020. Functional (un)cooperativity in elevator transport proteins. *Biochem Soc Trans*. doi:10.1042/BST20190970

Houssett P, Nagy M. 1977. Study of the Role of Purine Phosphoribosyltransferases in the Uptake of Adenine and Guanine by Schizosaccharomyces pombe Cells. *Eur J Biochem* **73**:99–105. doi:10.1111/j.1432-1033.1977.tb11295.x

Hu C, Tao L, Cao X, Chen L. 2020. The solute carrier transporters and the brain: Physiological and pharmacological implications. *Asian J Pharm Sci* **15**:131–144. doi:10.1016/j.ajps.2019.09.002

Hu N-J, Iwata S, Cameron AD, Drew D. 2012. Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT. *Nature* **478**:408–411. doi:10.1038/nature10450. Europe

Huynh KW, Jiang J, Abuladze N, Tsirulnikov K, Kao L, Shao X, Newman D, Azimov R, Pushkin A, Zhou ZH, Kurtz I. 2018. CryoEM structure of the human SLC4A4 sodium-coupled acid-base transporter NBCe1. *Nat Commun* **9**:900. doi:10.1038/s41467-018-03271-3

Jaehme M, Singh R, Garaeva AA, Duurkens RH, Slotboom DJ. 2018. PnuT uses a facilitated diffusion mechanism for thiamine uptake. *J Gen Physiol* **150**:41–50. doi:10.1085/jgp.201711850

Jaehme M, Slotboom DJ. 2015. Structure, function, evolution, and application of bacterial Pnu-type vitamin transporters. *Biol Chem* **396**:955–966. doi:10.1515/hsz-2015-0113

Jardetzky O. 1966. Simple Allosteric Model for Membrane Pumps. *Nature* **211**:969–970.

Jennings ML. 2005. Evidence for a second binding/transport site for chloride in erythrocyte anion transporter AE1 modified at glutamate 681. *Biophys J* **88**:2681–2691. doi:10.1529/biophysj.104.056812

References

Jensen KF, Arent S, Larsen S, Schack L. 2005. Allosteric properties of the GTP activated and CTP inhibited uracil phosphoribosyltransferase from the thermoacidophilic archaeon Sulfolobus solfataricus. *FEBS J* **272**:1440–1453. doi:10.1111/j.1742-4658.2005.04576.x

Jensen KF, Mygind B. 1996. Different oligomeric states are involved in the allosteric behavior of uracil phosphoribosyltransferase from Escherichia coli. *Eur J Biochem* **240**:637–645. doi:10.1111/j.1432-1033.1996.0637h.x

Jeschke G. 2018. MMM: A toolbox for integrative structure modeling. *Protein Sci* **27**:76–85. doi:10.1002/pro.3269

Kantrowitz ER. 2012. Allostery and cooperativity in Escherichia coli Aspartate Transcarbamoylase. *Arch Biochem Biophys* **519**:81–90. doi:10.1016/j.abb.2011.10.024

Karatza P, Frillingos S. 2005. Cloning and functional characterization of two bacterial members of the NAT/NCS2 family in Escherichia coli. *Mol Membr Biol* **22**:251–261. doi:10.1080/09687860500092927

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 α-helix. *J Biol Chem* **281**:39881–39890. doi:10.1074/jbc.M605748200

Karena E, Frillingos S. 2011. The Role of Transmembrane Segment TM3 in the Xanthine Permease XanQ of Escherichia coli. *J Biol Chem* **286**:39595–39605. doi:10.1074/jbc.M111.299164

Karena E, Tatsaki E, Lambrinidis G, Mikros E, Frillingos S. 2015. Analysis of conserved NCS2 motifs in the Escherichia coli xanthine permease XanQ. *Mol Microbiol* **98**:502–517. doi:10.1111/mmi.13138

Karinou E, Compton ELR, Morel M, Javelle A. 2013. The Escherichia coli SLC26 homologue YchM (DauA) is a C4-dicarboxylic acid transporter. *Mol Microbiol* **87**:623–640. doi:doi:10.1111/mmi.12120

Karniski LP. 2004. Functional expression and cellular distribution of diastrophic dysplasia sulfate transporter (DTDST) gene mutations in HEK cells. *Hum Mol Genet* **13**:2165–2171. doi:10.1093/hmg/ddh242

References

Kawate T, Gouaux E. 2006. Fluorescence-Detection Size-Exclusion Chromatography for Precrystallization Screening of Integral Membrane Proteins. *Structure* **14**:673–681. doi:10.1016/j.str.2006.01.013

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**:845–858. doi:10.1038/nprot.2015.053

Kim JW, Kim Subin, Kim Songwon, Lee H, Lee JO, Jin MS. 2017. Structural insights into the elevator-like mechanism of the sodium/citrate symporter CitS. *Sci Rep* **7**:2548. doi:10.1038/s41598-017-02794-x

Kim KS, Pelton JG, Inwood WB, Andersen U, Kustu S, Wemmer DE. 2010. The Rut pathway for pyrimidine degradation: Novel chemistry and toxicity problems. *J Bacteriol* **192**:4089–4102. doi:10.1128/JB.00201-10

Knapik A, Petkowski J, Otwinowski Z, Cymborowski MT, Cooper DR, Majorek KA, Chruszcz M, Krajewska WM, Minor W. 2012a. A multi-faceted analysis of RutD reveals a novel family of α/β hydrolases. *Proteins* **80**:2359–2368. doi:10.1002/prot.24122.A

Knapik A, Petkowski JJ, Otwinowski Z, Cymborowski MT, Cooper DR, Chruszcz M, Krajewska WM, Minor W. 2012b. Structure of Escherichia coli RutC, a member of the YjgF family and putative aminoacrylate peracid reductase of the rut operon. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **68**:1294–1299. doi:10.1107/S1744309112041796

Ko SBH, Zeng W, Dorwart MR, Luo X, Kim KH, Millen L, Goto H, Naruse S, Soyombo A, Thomas PJ, Muallem S. 2004. Gating of CFTR by the STAS domain of SLC26 transporters. *Nat Cell Biol* **6**:343–350. doi:10.1038/ncb1115

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. doi:10.1371/journal.pone.0041939

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. doi:10.1016/j.jmb.2010.02.037

References

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. doi:10.1016/j.jmb.2005.04.076

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

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* **433**:166814. doi:10.1016/j.jmb.2021.166814

Lee C, Yashiro S, Dotson DL, Uzdavinys P, Iwata S, Sansom MSP, von Ballmoos C, Beckstein O, Drew D, Cameron AD. 2014. Crystal structure of the sodium-proton antiporter NhaA dimer and new mechanistic insights. *J Gen Physiol* **144**:529–544. doi:10.1085/jgp.201411219

Li J, Huang S, Chen J, Yang Z, Fei X, Zheng M, Ji C, Xie Y, Mao Y. 2007. Identification and characterization of human uracil phosphoribosyltransferase (UPRTase). *J Hum Genet* **52**:415–422. doi:10.1007/s10038-007-0129-2

Liang WJ, Johnson D, Jarvis SM. 2001. Vitamin C transport systems of mammalian cells. *Mol Membr Biol* **18**:87–95. doi:10.1080/09687680110033774

Linde L, Jensen KF. 1996. Uracil phosphoribosyltransferase from the extreme thermoacidophilic archaebacterium Sulfolobus shibatae is an allosteric enzyme, activated by GTP and inhibited by CTP. *Biochim Biophys Acta* **1296**:16–22. doi:10.1016/0167-4838(96)00045-3.

Liu SW, Milman G. 1983. Purification and characterization of Escherichia coli guanine-xanthine phosphoribosyltransferase produced by a high efficiency expression plasmid utilizing a lambda PL promoter and CI857 temperature-sensitive repressor. *J Biol Chem* **258**:7469–7475. doi:10.1016/s0021-9258(18)32201-4

Loh KD, Gyaneshwar P, Papadimitriou EM, Fong R, Kim KS, Parales R, Zhou Z, Inwood W, Kustu S. 2006. A previously underscribed pathway for pyrimidine catabolism. *Proc Natl Acad Sci U S A* **103**:5114–5119. doi:10.1073/pnas.0600521103

Longley DB, Harkin DP, Johnston PG. 2003. 5-fluorouracil: Mechanisms of Action and Clinical Strategies. *Nat Rev Cancer* **3**:330–8. doi:10.1038/nrc1074

References

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**:243–247. doi:10.1038/nature09885

Lundegaard C, Jensen KF. 1999. Kinetic mechanism of uracil phosphoribosyltransferase from Escherichia coli and catalytic importance of the conserved proline in the PRPP binding site. *Biochemistry* **38**:3327–3334. doi:10.1021/bi982279q

Luo P, Dai S, Zeng J, Duan J, Shi H, Wang J. 2018. Inward-facing conformation of L-ascorbate transporter suggests an elevator mechanism. *Cell Discov* **4**:1–9. doi:10.1038/s41421-018-0037-y Cell

Maes M, Rimon A, Kozachkov-Magrisso L, Friedler A, Padan E. 2012. Revealing the ligand binding site of NhaA Na+/H+ antiporter and its pH dependence. *J Biol Chem* **287**:38150–38157. doi:10.1074/jbc.M112.391128

Mani C, Pai S, Papke CM, Palle K, Gmeiner WH. 2018. Thymineless Death by the Fluoropyrimidine Polymer F10 Involves Replication Fork Collapse and Is Enhanced by Chk1 Inhibition. *Neoplasia (United States)* **20**:1236–1245. doi:10.1016/j.neo.2018.10.006

Marino J, Holzhüter K, Kuhn B, Geertsma ER. 2017. Efficient Screening and Optimization of Membrane Protein Production in Escherichia coli. *Methods Enzymol* **594**:139–164. doi:10.1016/bs.mie.2017.05.011

Martinez KA, Kitko RD, Mershon JP, Adcox HE, Malek KA, Berkmen MB, Slonczewski JL. 2012. Cytoplasmic pH response to acid stress in individual cells of Escherichia coli and Bacillus subtilis observed by fluorescence ratio imaging microscopy. *Appl Environ Microbiol* **78**:3706–3714. doi:10.1128/AEM.00354-12

Martzoukou O, Karachaliou M, Yalelis V, Leung J, Byrne B, Amillis S, Diallinas G. 2015a. Oligomerization of the UapA Purine Transporter Is Critical for ER-Exit, Plasma Membrane Localization and Turnover. *J Mol Biol* **427**:2679–2696. doi:10.1016/j.jmb.2015.05.021

Martzoukou O, Karachaliou M, Yalelis V, Leung J, Byrne B, Amillis S, Diallinas G. 2015b. Oligomerization of the UapA Purine Transporter Is Critical for ER-Exit, Plasma Membrane Localization and Turnover. *J Mol Biol* **427**:2679–2696. doi:10.1016/j.jmb.2015.05.021

Masrati G, Mondal R, Rimon A, Kessel A, Padan E, Lindahl E, Ben-Tal N. 2020. An angular motion of a conserved four-helix bundle facilitates alternating access transport in the TtNapA and EcNhaA transporters. *Proc Natl Acad Sci U S A* **117**:31850–31860. doi:10.1073/pnas.2002710117

References

May JM. 2011. The SLC23 family of ascorbate transporters: Ensuring that you get and keep your daily dose of vitamin C. *Br J Pharmacol* **164**:1793–1801. doi:10.1111/j.1476-5381.2011.01350.x

Maynes JT, Yuan RG, Snyder FF. 2000. Identification, expression, and characterization of Escherichia coli guanine deaminase. *J Bacteriol* **182**:4658–4660. doi:10.1128/JB.182.16.4658-4660.2000

Meintanis C, Karagouni AD, Diallinas G. 2000. Amino acid residues N450 and Q449 are critical for the uptake capacity and specificity of UapA, a prototype of a nucleobase-ascorbate transporter family. *Mol Membr Biol* **17**:47–57. doi:10.1080/096876800294489

Mermelekas G, Georgopoulou E, Kallis A, Botou M, Vlantos V, Frillingos S. 2010. Cysteine-scanning analysis of helices TM8, TM9a, and TM9b and intervening loops in the YgfO xanthine permease: A carboxyl group is essential at ASP-276. *J Biol Chem* **285**:35011–35020. doi:10.1074/jbc.M110.170415

Michael AJ. 2017. Evolution of biosynthetic diversity. *Biochem J* **474**:2277–2299. doi:10.1042/BCJ20160823

Migneault I, Dartiguenave C, Bertrand MJ, Waldron KC. 2004. Glutaraldehyde: Behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. *Biotechniques* **37**:790–802. doi:10.2144/04375rv01

Moraes TF, Reithmeier RAF. 2012. Membrane transport metabolons. *Biochim Biophys Acta - Biomembr* **1818**:2687–2706. doi:10.1016/j.bbamem.2012.06.007

Moy BE, Seshu J. 2021. STAS Domain Only Proteins in Bacterial Gene Regulation. *Front Cell Infect Microbiol* **11**:1–8. doi:10.3389/fcimb.2021.679982

Mueckler M, Thorens B. 2013. The SLC2 (GLUT) Family of Membrane Transporters. *Mol Aspects Med* **34**:121–138. doi:10.1038/jid.2014.371

Mulligan C, Fenollar-Ferrer C, Fitzgerald GA, Vergara-Jaque A, Kaufmann D, Li Y, Forrest LR, Mindell JA. 2016. The bacterial dicarboxylate transporter, VcINDY, uses a two-domain elevator-type mechanism. *Nat Struct Mol Biol* **23**:256–263. doi:10.1085/jgp.201311141

Najafi SMA, Harris DA, Yudkin MD. 1996. The SpoIIAA protein of Bacillus subtilis has GTP-binding properties. *J Bacteriol* **178**:6632–6634. doi:10.1128/jb.178.22.6632-6634.1996

References

Niopek-Witz S, Deppe J, Lemieux MJ, 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 - Biomembr* **1838**:3025–3035. doi:10.1016/j.bbamem.2014.08.013

Padan E, Danieli T, Keren Y, Alkoby D, Masrati G, Haliloglu T, Ben-Tal N, Rimon A. 2015. NhaA antiporter functions using 10 helices, and an additional 2 contribute to assembly/stability. *Proc Natl Acad Sci U S A* **112**:E5575–E5582. doi:10.1073/pnas.1510964112

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. doi:10.1016/j.jmb.2008.08.005

Papageorgiou IG, Yakob L, Salabi MIAL, Diallinas G. 2005. Identification of the first pyrimidine nucleobase transporter in Leishmania : similarities with the Trypanosoma brucei U1 transporter and antileishmanial activity of uracil analogues. *Parasitology* **130**:275–283. doi:10.1017/S0031182004006626

Papakostas K, Botou M, Frillingos S. 2013. Functional identification of the hypoxanthine/guanine transporters YjcD and YgfQ and the adenine transporters PurP and YicO of Escherichia coli K-12. *J Biol Chem* **288**:36827–36840. doi:10.1074/jbc.M113.523340

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

Papakostas K, Georgopoulou E, Frillingos S. 2008. Cysteine-scanning Analysis of Putative Helix XII in the YgfO Xanthine Permease. *J Biol Chem* **283**:13666–13678. doi:10.1074/jbc.m800261200

Pasqualetto E, Aiello R, Gesiot L, Bonetto G, Bellanda M, Battistutta R. 2010. Structure of the Cytosolic Portion of the Motor Protein Prestin and Functional Role of the STAS Domain in SLC26/SulP Anion Transporters. *J Mol Biol* **400**:448–462. doi:10.1016/j.jmb.2010.05.013

Perland E, Fredriksson R. 2017. Classification Systems of Secondary Active Transporters. *Trends Pharmacol Sci* **38**:305–315. doi:10.1016/j.tips.2016.11.008

Pivovarov AS, Calahorro F, Walker RJ. 2019. Na+/K+-pump and neurotransmitter membrane receptors. *Invertebr Neurosci* **19**:1–16. doi:10.1007/s10158-018-0221-7

References

Porra V, Bernier-Valentin F, Trouttet-Masson S, Berger-Dutrieux N, Peix JL, Perrin A, Selmi-Ruby S, Rousset B. 2002. Characterization and semiquantitative analyses of pendrin expressed in normal and tumoral human thyroid tissues. *J Clin Endocrinol Metab* **87**:1700–1707. doi:10.1210/jcem.87.4.8372

Printz RL, Magnuson MA, Granner DK. 1993. Mammalian Glucokinase. *Annu Rev Nutr* **13**:463–496. doi:10.1146/annurev.nu.13.070193.002335

Putman M, Van Veen HW, Poolman B, Konings WN. 1999. Restrictive use of detergents in the functional reconstitution of the secondary multidrug transporter LmrP. *Biochemistry* **38**:1002–1008. doi:10.1021/bi981863w

Rapp C, Bai X, Reithmeier RAF. 2017. Molecular analysis of human solute carrier SLC26 anion transporter disease-causing mutations using 3-dimensional homology modeling. *Biochim Biophys Acta - Biomembr* **1859**:2420–2434. doi:10.1016/j.bbamem.2017.09.016

Rasmussen UB, Mygind B, Nygaard P. 1986. Purification and some properties of uracil phosphoribosyltransferase from Escherichia coli K12. *Biochim Biophys Acta* **881**:268–275. doi:10.1016/0304-4165(86)90013-9

Reang J, Sharma PC, Thakur VK, Majeed J. 2021. Understanding the therapeutic potential of ascorbic acid in the battle to overcome cancer. *Biomolecules* **11**. doi:10.3390/biom11081130

Rimon A, Tzubery T, Padan E. 2007. Monomers of the NhaA Na+/H+ Antiporter of Escherichia coli Are Fully Functional yet Dimers Are Beneficial under Extreme Stress Conditions at Alkaline pH in the Presence of Na+ or Li+. *J Biol Chem* **282**:26810–26821. doi:10.1074/jbc.M704469200

Robertson JL, Kolmakova-Partensky L, Miller C. 2010. Design, function and structure of a monomeric ClC transporter. *Nature* **468**:844–847. doi:10.1038/nature09556

Ruan Y, Miyagi A, Wang X, Chami M, Boudker O, Scheuring S. 2017. Direct visualization of glutamate transporter elevator mechanism by high-speed AFM. *Proc Natl Acad Sci U S A* **114**:1584–1588. doi:10.1073/pnas.1616413114

Sadee W, Wong CG. 1977. Pharmacokinetics of 5-fluorouracil: Inter-relationship with Biochemical Kinetics in Monitoring Therapy. *Clin Pharmacokinet* **2**:437–450. doi:10.2165/00003088-197702060-00004

References

Saier MH, Reddy VS, Tsu B V., Ahmed MS, Li C, Moreno-Hagelsieb G. 2016. The Transporter Classification Database (TCDB): Recent advances. *Nucleic Acids Res* **44**:D372–D379. doi:10.1093/nar/gkv1103

Saier MH, Tran C V., Barabote RD. 2006. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res* **34**:181–186. doi:10.1093/nar/gkj001

Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa G, Karlsson A, Al-lazikani B, Hersey A, Oprea TI. 2019. A comprehensive map of molecular drug targets. *Nat Rev Drug Discov* **16**:19–34. doi:10.1038/nrd.2016.230.A

Sauer DB, Trebesch N, Marden JJ, Cocco N, Song J, Koide A, Koide S, Tajkhorshid E, Wang DN. 2020. Structural basis for the reaction cycle of dass dicarboxylate transporters. *Elife* **9**:1–74. doi:10.7554/eLife.61350

Savini I, Rossi A, Pierro C, Avigliano L, Catani M V. 2008. SVCT1 and SVCT2: Key proteins for vitamin C uptake. *Amino Acids* **34**:347–355. doi:10.1007/s00726-007-0555-7

Schindelin J, Arganda-Carrera I, Frise E, Verena K, Mark L, Tobias P, Stephan P, Curtis R, Stephan S, Benjamin S, Jean-Yves T, Daniel JW, Volker H, Kevin E, Pavel T, Albert C. 2013. Fiji - an Open platform for biological image analysis. *Nat Methods* **9**. doi:10.1038/nmeth.2019.Fiji

Schneider TD, Stephens RM. 1990. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18**:6097–6100.

Schramm VL, Grubmeyer C. 2004. Phosphoribosyltransferase Mechanisms and Roles in Nucleic Acid Metabolism. *Prog Nucleic Acid Res Mol Biol* **78**:261–304. doi:10.1016/S0079-6603(04)78007-1

Schumann T, König J, Henke C, Willmes DM, Bornstein SR, Jordan J, Fromm MF, Birkenfeld AL. 2020. Solute carrier transporters as potential targets for the treatment of metabolic disease. *Pharmacol Rev* **72**:343–379. doi:10.1124/pr.118.015735

Schushan M, Xiang M, Bogomiakov P, Padan E, Rao R, Ben-Tal N. 2010. Model-guided mutagenesis drives functional studies of human NHA2, implicated in hypertension. *J Mol Biol* **396**:1181–1196. doi:10.1016/j.jmb.2009.12.055.Model-guided

References

Seavers PR, Lewis RJ, Brannigan JA, Verschueren KHG, Murshudov GN, Wilkinson AJ. 2001. Structure of the Bacillus cell fate determinant SpoIIAA in phosphorylated and unphosphorylated forms. *Structure* **9**:605–614. doi:10.1016/S0969-2126(01)00623-2

Sharma AK, Rigby AC, Alper SL. 2011. STAS Domain Structure and Function. *Cell Physiol Biochem* **28**:407–422. doi:10.1159/000335104

Shibagaki N, Grossman AR. 2006. The role of the STAS domain in the function and biogenesis of a sulfate transporter as probed by random mutagenesis. *J Biol Chem* **281**:22964–22973. doi:10.1074/jbc.M603462200

Shibagaki N, Grossman AR. 2004. Probing the function of STAS domains of the Arabidopsis sulfate transporters. *J Biol Chem* **279**:30791–30799. doi:10.1074/jbc.M403248200

Shibata T, Kawasaki S ya, Fujita J ya, Kabashima T, Kai M. 2010. A novel and specific fluorescence reaction for uracil. *Anal Chim Acta* **674**:234–238. doi:10.1016/j.aca.2010.06.028

Shimada T, Hirao K, Kori A, Yamamoto K, Ishihama A. 2007. RutR is the uracil/thymine-sensing master regulator of a set of genes for synthesis and degradation of pyrimidines. *Mol Microbiol* **66**:744–757. doi:10.1111/j.1365-2958.2007.05954.x

Shimosaka M, Fukuda Y, Murata K, Kimura A. 1985. Purification and Properties of Orotate Phosphoribosyltransferases from Escherichia coli K-12 , and Its Derivative Purine-Sensitive Mutant. *J Biochem* **98**:1689–1697.

Sievers F, Higgins DG. 2018. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* **27**:135–145. doi:10.1002/pro.3290

Skerra A. 1994. Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in Escherichia coli. *Gene* **151**:131–135. doi:10.1016/0378-1119(94)90643-2

Song Y, Mao J, Gunner MR. 2009. MCCE2: Improving Protein pKa Calculations with Extensive Side Chain Rotamer Sampling. *J Comput Chem* **30**:2231–2247. doi:10.1002/jcc.21222.MCCE2

Sotiriou S, Gispert S, Cheng J, Wang Y, Chen A, Hoogstraten-Miller S, Miller GF, Kwon O, Levine M, Guttentag SH, Nussbaum RL. 2002. Ascorbic-acid transporter Slc23a1 is essential for vitamin C transport into the brain and for perinatal survival. *Nat Med* **8**:514–517. doi:10.1038/0502-514

References

Soysa R, Wilson ZN, Elferich J, Forquer I, Shinde U, Riscoe MK, Yates PA, Ullman B. 2013. Substrate inhibition of uracil phosphoribosyltransferase by uracil can account for the uracil growth sensitivity of leishmania donovani pyrimidine auxotrophs. *J Biol Chem* **288**:29954–29964. doi:10.1074/jbc.M113.478826

Springer S, Malkus P, Borchert B, Wellbrock U, Duden R, Schekman R. 2014. Regulated oligomerization induces uptake of a membrane protein into COPII vesicles independent of its cytosolic tail. *Traffic* **15**:531–545. doi:10.1111/tra.12157

Stamm M, Staritzbichler R, Khafizov K, Forrest LR. 2014. AlignMe - A membrane protein sequence alignment web server. *Nucleic Acids Res* **42**:246–251. doi:10.1093/nar/gku291

Sternisha SM, Miller BG. 2019. Molecular and Cellular Regulation of Human Glucokinase. *Arch Biochem Biophys* **663**:199–213. doi:10.1016/j.abb.2019.01.011

Stoffer-Bittner AJ, Alexander CR, Dingman DW, Mourad GS, Schultes NP. 2018. The solute transport and binding profile of a novel nucleobase cation symporter 2 from the honeybee pathogen Paenibacillus larvae. *FEBS Open Bio* **8**:1322–1331. doi:10.1002/2211-5463.12488

Stoffer AJ, Alexander CR, Dingman DW, Mourad GS, Schultes NP. 2018. Functional characterization of the uracil transporter from honeybee pathogen Paenibacillus larvae. *Microb Pathog* **124**:305–310. doi:10.1016/j.micpath.2018.08.059

Tang Y, Cao Z, Livoti E, Krauss U, Jaeger KE, Gärtner W, Losi A. 2010. Interdomain signalling in the blue-light sensing and GTP-binding protein YtvA: A mutagenesis study uncovering the importance of specific protein sites. *Photochem Photobiol Sci* **9**:47–56. doi:10.1039/b9pp00075e

Thurtle-Schmidt BH, Stroud RM. 2016. Structure of Bor1 supports an elevator transport mechanism for SLC4 anion exchangers. *Proc Natl Acad Sci* **113**:10542–10546. doi:10.1073/pnas.1612603113

Tsukaguchi H, Tokui T, Mackenzie B, Berger U V, Chen X-Z, Wang X, Brubaker RF, Hediger MA. 1999a. A family of mammalian Na+-dependent L-ascorbic acid transporters. *Nature* **399**:70–75.

Tsukaguchi H, Tokui T, Mackenzie M, Berger U, Chen X-Z, Wang Y, Brubaker RF, Hediger MA. 1999b. A family of mammalian Na + -dependent. *Nature* **399**:2–7. doi:10.1038/19986

References

Tu A-HT, Turnbough CL. 1997. Regulation of upp Expression in Escherichia coli by UTP-Sensitive Selection of Transcriptional Start Sites Coupled with UTP-Dependent Reiterative Transcription. *J Bacteriol* **179**:6665–6673. doi:10.1128/jb.179.21.6665-6673.1997

Turner RJ, Bonner ER, Grabner GK, Switzer RL. 1998. Purification and Characterization of Bacillus subtilis PyrR, a Biofunctional pyr mRNA-binding Attenuation Protein/Uracil Phosphoribosyltransferase. *J Biol Chem* **273**:5932–5938. doi:10.1016/S0969-2126(98)00036-7

Uldry M, Thorens B. 2004. The SLC2 family of facilitated hexose and polyol transporters. *Pflugers Arch Eur J Physiol* **447**:480–489. doi:10.1007/s00424-003-1085-0

Vastermark A, Driker A, Weng J, Li X, Wang J, Saier Jr. MH. 2016. The V-motifs facilitate the substrate capturing step of the PTS elevator mechanism. *J Struct Biol* **196**:496–502. doi:10.1126/science.Liquids

Veenhoff LM, Poolman B. 1999. Substrate recognition at the cytoplasmic and extracellular binding site of the lactose transport protein of Streptococcus thermophilus. *J Biol Chem* **274**:33244–33250. doi:10.1074/jbc.274.47.33244

Verdon G, Oh SC, Serio R, Boudker O. 2014. Coupled ion binding and structural transitions along the transport cycle of glutamate transporters. *Elife* **2014**:1–23. doi:10.7554/eLife.02283

Vergara-Jaque A, Fenollar-Ferrer C, Kaufmann D, Forrest LR. 2015. Repeat-swap homology modeling of secondary active transporters: Updated protocol and prediction of elevator-type mechanisms. *Front Pharmacol* **6**. doi:10.3389/fphar.2015.00183

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. doi:10.1016/j.jmb.2005.12.070

Wagner S, Baarst L, Ytterberg AJ, Klussmerer A, Wagner CS, Nord O, Nygren PÅ, Van Wijks KJ, De Gier JW. 2007. Consequences of membrane protein overexpression in Escherichia coli. *Mol Cell Proteomics* **6**:1527–1550. doi:10.1074/mcp.M600431-MCP200

Waldo GS, Standish BM, Berendzen J, Terwilliger TC. 1999. Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol* **17**:691–695. doi:10.1038/10904

References

Walter JD, Sawicka M, Dutzler R. 2019. Cryo-EM structures and functional characterization of murine Slc26a9 reveal mechanism of uncoupled chloride transport. *Elife* **8**:1–31. doi:10.7554/eLife.46986

Wang C, Sun B, Zhang W, Huang X, Zhang M, Guo H, Chen X, Huang F, Chen T, Mi H, Yu F, Liu LN, Zhang P. 2019. Structural mechanism of the active bicarbonate transporter from cyanobacteria. *Nat Plants* **5**:1184–1193.

Wang Liguang, Feng X, Zhao H, Wang Lidong, An L, Qiu QS. 2014. Functional analysis of the Na+,K+/H+ antiporter PeNHX3 from the tree halophyte Populus euphratica in yeast by model-guided mutagenesis. *PLoS One* **9**:29–32. doi:10.1371/journal.pone.0104147

Wang Y, Mackenzie B, Tsukaguchi H, Weremowicz S, Morton CC, Hediger MA. 2000. Human vitamin C (L-Ascorbic acid) transporter SVCT1. *Biochem Biophys Res Commun* **267**:488–494. doi:10.1006/bbrc.1999.1929

Wilks JC, Slonczewski JL. 2007. pH of the cytoplasm and periplasm of Escherichia coli: Rapid measurement by green fluorescent protein fluorimetry. *J Bacteriol* **189**:5601–5607. doi:10.1128/JB.00615-07

Winkelmann I, Matsuoka R, Meier PF, Shutin D, Zhang C, Orellana L, Sexton R, Landreh M, Robinson C V, Beckstein O, Drew D. 2020. Structure and elevator mechanism of the mammalian sodium/proton exchanger NHE9. *EMBO J* **39**:4541–4559. doi:10.15252/embj.2020105908

Wohlrab C, Phillips E, Dachs GU. 2017. Vitamin c transporters in cancer: Current understanding and gaps in knowledge. *Front Oncol* **7**:5–10. doi:10.3389/fonc.2017.00074

Wright NJ, Lee SY. 2019. Structures of human ENT1 in complex with adenosine reuptake inhibitors. *Nat Struct Mol Biol* **26**:599–606. doi:10.1038/s41594-019-0245-7

Xie H. 2008. Activity assay of membrane transport proteins. *Acta Biochim Biophys Sin (Shanghai)* **40**:269–277. doi:10.1111/j.1745-7270.2008.00400.x

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

References

Yamamoto S, Inoue K, Murata T, Kamigaso S, Yasujima T, Maeda J, Yoshida Y, Ohta K, Yuasa H. 2010b. Identification and Functional Characterization of the First Nucleobase Transporter in Mammals. *J Biol Chem* **285**:6522–6531. doi:10.1074/jbc.m109.032961

Yang Liu, Yang Lianjuan, Yu H, Liu L, Zhao X, Huang X. 2018. Conformational transitions of uracil transporter UraA from Escherichia coli: a molecular simulation study. *J Biomol Struct Dyn* **36**:3398–3410. doi:10.1080/07391102.2017.1388288

Yanisch-Perron C, Vieira J, Messing J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene* **33**:103–119. doi:10.1016/0378-1119(85)90120-9

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

Yata VK, Sen K, Kumar MVS, Ghosh SS. 2012. Interaction studies of E. coli uracil phosphoribosyltransferase with 5-fluorouracil for potent anti cancer activity. *Med Chem Res* **21**:1149–1155. doi:10.1007/s00044-011-9627-z

Yin H, Flynn AD. 2016. Drugging Membrane Protein Interactions. *Annu Rev Biomed Eng* **18**:51–76. doi:10.1146/annurev-bioeng-092115-025322

Young JD. 2016. The SLC28 (CNT) and SLC29 (ENT) nucleoside transporter families: A 30-year collaborative odyssey. *Biochem Soc Trans* **44**:869–876. doi:10.1042/BST20160038

Yu X, Yang G, Yan C, Baylon JL, 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. doi:10.1038/cr.2017.83

Yuasa H, Yasujima T, Inoue K. 2020. Current Understanding of the Intestinal Absorption of Nucleobases and Analogs. *Biol Pharm Bull* **43**:1293–1300. doi:10.1248/bpb.b20-00342

Zhang D, Kiyatkin A, Bolin JT, Low PS. 2000. Crystallographic structure and functional interpretation of the cytoplasmic domain of erythrocyte membrane band 3. *Blood* **96**:2925–2933. doi:10.1182/blood.v96.9.2925

References

Zheng J, Du GG, Matsuda K, Orem A, Aguiñaga S, Deák L, Navarrete E, Madison LD, Dallos P. 2005. The C-terminus of prestin influences nonlinear capacitance and plasma membrane targeting. *J Cell Sci* **118**:2987–2996. doi:10.1242/jcs.02431

Zhou F, Zhu L, Wang K, Murray M. 2017. Recent advance in the pharmacogenomics of human Solute Carrier Transporters (SLCs) in drug disposition. *Adv Drug Deliv Rev* **116**:21–36. doi:10.1016/j.addr.2016.06.004

Zhou W, Fiorin G, Anselmi C, Karimi-Varzaneh HA, Poblete H, Forrest LR, Faraldo-Gómez JD. 2019. Large-scale state-dependent membrane remodeling by a transporter protein. *Elife* **8**:1–32. doi:10.7554/eLife.50576

Zhou X, Levin EJ, Pan Y, McCoy JG, Sharma R, Kloss B, Bruni R, Quick M, Zhou M. 2014. Structural basis of the alternating-access mechanism in a bile acid transporter. *Nature* **505**:569–573. doi:10.1038/nature12811.Structural

Zhou Y, Bushweller JH. 2018. Solution Structure and Elevator Mechanism of the Membrane Electron Transporter CcdA. *Nat Struct Mol Biol* **25**:163–169. doi:10.1038/s41594-018-0022-z

List of abbreviations

1MU	1-methyluracil
2-HCT	2-Hydroxycarboxylate Transporter
3MU	3-methyluracil
5FU	5-fluorouracil
6MU	6-methyluracil
7TMIR	seven transmembrane segments inverted repeat
aa	amino acids
AbgT	p-Aminobenzoyl glutamate transporter
ACP	acyl carrier protein
AE1	anion exchanger 1
AFM	atomic force microscopy
amp	ampicillin
ara	arabinose
ASBT	apical sodium dependent bile acid transporter
atc	anhydrotetracycline
BAO	benzamide oxime
BME	ß-mercaptoethanol
CCCP	carbonyl cyanide m-chlorophenyl hydrazone
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CNT	concentrative nucleoside transporter
COG	Cluster of orthologous genes
CPA	Cation-Proton Antiporter
CPM	7-Diethylamino-3-(4'-Maleimidylphenyl)-4-Methylcoumarin
CuPhen	copper 1,10-phenanthroline
CW	continuous wavelength
cysl	cysteine-free
CYT	cytosine
DDM	n-Dodecyl-ß-maltoside
DM	n-Decyl-ß-maltoside (DM)
DMSO	dimethyl sulfoxide
DSF	differential scanning fluorimetry
DTT	dithiothreitol
E. coli	*Escherichia coli*
ECF	Energy-coupling factor
EDTA	ethylene diamine tetraacetic acid
ENT	Equilibrative Nucleoside Transporter
EPR	electron paramagnetic resonance
ER	endoplasmatic reticulum
FDA	food and drug administration
FRET	Förster resonance energy transfer
FSEC	fluorescence-detection size exclusion chromatography
GA	glutaraldehyde
GFP	green fluorescent protein
Glc	glucose-glucoside
GPCR	G-protein coupled receptor
GS-linker	glycine-serine-linker
HP	helical hairpin

HPLC	high-pressure liquid chromatography
HRV	Human rhinovirus
IMAC	immobilized metal affinity chromatography
IVS	intervening sequence
KOH	kalium hydroxide
KPi	potassium phosphate
Lac	Lactose-N,N'-Diacetylchitobiose-β-glucoside
L-asc	L-ascorbate
LB	lysogeny broth
LOV	light-oxygen-voltage sensing
	(1-Acetoxy-2,2,5,5-tetramethyl-δ-3-pyrroline-3-methyl)
MTSL	methanethiosulfonate
MWCO	molecular weight cutoff
NaCT	Na+-citrate cotransporter
NaPi	sodium phosphate
NCS1	nucleobase-cation symporter 1
NiNTA	Nickel nitrilotriacetic acid
OG	n-Octyl-ß-D-glucoside
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PDB	protein data bank
PDVF	Polyvinylidene fluoride
PEG	polyethylene glycol
PELDOR	pulsed electron-electron double resonance
PMSF	phenylmethylsulfonyl fluoride
PRPP	Phosphoribosyl pyrophosphate
PTS	phosphotransferase
PYR	pyrimidine
RFP	ref fluorescent protein
RMSD	root mean square deviation
SDS	sodium dodecyl sulfate
SEC	size-exlcusion chromatography
SLC	solute carrier
soy PC	L-α-phosphatidylcholine from soy bean
SPA	scintillation proximity assay
STAS	sulphate transporter and antisigma factor antagonist
TB	terrific broth
TCA	trichloroacetic acid
TCDB	Transporter Classification Database
THY	thymine
TM	transmembrane
Tm	melting temperature
UC	ultracentrifugation
UMP	uridine monophosphate
UPRT	uracil phosphoribosyl transferase
URA	uracil

Declaration of collaborative work

Except where stated otherwise by reference or acknowledgement, the work presented here was generated by myself under the supervision of my advisor during my doctoral studies. All contributions from colleagues are explicitly referenced in the thesis. The material listed below was obtained in the context of collaborative research with the following people:

Figure 13: **Uracil uptake by UraA(E241A).** The UraA(E241A) mutant was generated and kindly provided by Benedikt Kuhn, Institute of Biochemistry, Goethe University Frankfurt a. M.. I performed all transformations and uptake experiments.

Figure 15: Background uptake in *E. coli* BW25113(*ΔuraAΔrutG*). The knockout of *rutG* in *E. coli* BW25113(*ΔuraA*) was performed by Leonard Präve as part of his master thesis under my supervision. I performed transformations and all uptake experiments.

Figure 23: Co-expression of UPRT-GFP and UraA-RFP in *E. coli* BW25113(*ΔuraAΔrutG*). The empty p7XC3RH expression vector used to generate pTX-XH_UraA-RFP was generated and kindly provided by Dr. Alina Ornik-Cha, Institute of Biochemistry, Goethe University Frankfurt a. M.. All subsequent cloning and experiments were performed by myself.

Figure 33: Uracil uptake of substrate binding site mutants of UraA, 51ThXi and 55PyPi. The UraA(E241A) mutant was generated and kindly provided by Benedikt Kuhn, Institute of Biochemistry, Goethe University Frankfurt a. M.. All other mutants were generated by myself. I performed all transformations and uptake experiments.

Figure 59: Cysteine mutants of SLC26Dg. SLC26Dg^{cysl} (SLC26Dg(C282S)) in pINIT vector was kindly provided by Dr. Yung-Ning Chang, Institute of Biochemistry, Goethe University Frankfurt a. M..

Figure 61: PELDOR measurements on SLC26Dg. PELDOR measurements and CW-EPR, data analysis as well as preparation of figure 61C were done by Dr. Benesh Joseph, Institute of Biophysics, Goethe University Frankfurt a. M.. Sample preparation and spin labeling was performed by myself.

Figure 64: PELDOR measurements on the linker mutant of SLC26Dg. PELDOR measurements and CW-EPR, data analysis as well as preparation of figures 64B and 64C were done by Dr. Benesh Joseph, Institute of Biophysics, Goethe University Frankfurt a. M.. Sample preparation and spin labeling was performed by myself.

Figure 65: **Purification and reconstitution of SLC26Si[IL].** SLC26Si[cysl] (SLC26Si(C578S)) in pINIT vector was kindly provided by Dr. Yung-Ning Chang, Institute of Biochemistry, Goethe University Frankfurt a. M..

Table 29: Theoretical pK$_a$ of substrate binding site residues E241, H245 and E290 of UraA. The calculation of theoretical pKa of UraA mutants was performed by Dr. Ahmadreza Mehdipour from the MPI for Biophysics, Frankfurt a. M..

Figure S27: SEC-MALS on 51ThXi and 51ThXi tryptophan mutants. SEC-MALS experiments were performed and chromatograms were generated by Dr. Rupert Abele, Institute of Biochemistry, Goethe University Frankfurt a. M.. Samples were prepared by myself.

List of publications

Holzhüter K, Geertsma ER. 2020. Functional (un)cooperativity in elevator transport proteins. *Biochem Soc Trans*. doi:10.1042/BST20190970

Marino J, Holzhüter K, Kuhn B, Geertsma ER. 2017. Efficient Screening and Optimization of Membrane Protein Production in Escherichia coli. *Methods Enzymol* **594**:139–164. doi:10.1016/bs.mie.2017.05.011

Acknowledgements

Special thanks to 'the gang' (that would be Selena, Julia, and Alina): I would probably also have managed somehow without you, but with you it was so much more enjoyable. I guess you know I consider you family.

I also would like to thank my grandmother, who really has no clue about what we are doing, but occasionally sent chocolate.

Ultimately, I thank Dr. Reinke Müller, for being rock-solid like a mountain to which I can cling in stormy weather. I know together we are better than on our own, and I hope we will also share whatever step now lies ahead of me.

Curriculum vitae

Name: Katharina Holzhüter

Birthday: 26.02.1991

Birthplace: Köln

Education

02/2016 – present: **PhD candidate** in the group of Ass. Prof. Dr. Eric Geertsma, Goethe University Frankfurt a. M.
Thesis: "Substrate translocation in SLC23 and SLC26 transporters"

10/2013 – 12/2015:   **M.Sc. Biochemistry**, Goethe University Frankfurt a. M.
Average grade: 1.8
Thesis: "Structural and Functional Characterization of UraA-like Transporters", supervised by Ass. Prof. Dr. Eric Geertsma

10/2010 – 03/2014:   **B.Sc. Biochemistry**, Goethe University Frankfurt a. M.
Average grade: 1.3
Thesis: "Spectroscopic study of natural and unnatural derivatives of the pH-responsive cytosine-rich human telomeric DNA for nanodevice insight", supervised by Prof. Dr. Harald Schwalbe

08/2002-03/2010:   Final secondary-school examination, Martin-Butzer-Gymnasium Dierdorf

Experimental Skills

Molecular biology   Conventional cloning, FX cloning, InFusion cloning
Mutagenesis (Quikchange, Inverse PCR, Overlap Extension PCR)
Genomic knockouts
Vector construction
Nanobody library construction (RNA extraction, RT-PCR)

Protein expression and purification   Heterologous protein expression in *E. coli*
Purification of Nanobodies, soluble proteins and membrane proteins (IMAC, SEC)
Membrane protein reconstitution
In gel fluorescence, Western blotting, silver staining

Experimental Skills

Protein characterization	High-throughput homolog screening
	Differential scanning fluorimetry
	Fluorescence-detection size exclusion chromatography
	Protein crystallization (excluding data processing)
	Oxidative and unspecific crosslinking
Assay development and functional studies	Radioactive transport assays in whole cells
	Proteoliposome-based radioactive transport assays
	Plate dilution and minimal inhibitory concentration assays
	Development of an activity assay for a soluble *E. coli* protein

Conferences and Poster Presentation

09/2021	Plenary talk on "Uniport in non-mammalian SLC23 transporters", Transporttage 2021, Greifswald (hybrid online conference)
11/2019	Plenary talk on "The STAS domain in SLC26 Transporters", Joint winter school in Eberbach, Germany
10/2017	Poster presentation at the CRC807 International Symposium "New Horizons in Membrane Transport and Communication" Poster title: "Structural and Functional Characterization of SLC23 Proteins", Katharina Holzhüter, Benedikt Kuhn and Eric R. Geertsma
10/2016	Short talk on "Binding vs. transport – uracil uptake in *E. coli*", CRC807 Retreat, Bad Muenster am Stein, Germany
08/2016	Plenary talk on "Structural and Functional Characterization of NCS2 Proteins", Joint summer school in Groningen, Netherlands
05/2016	Plenary talk on "Functional Characterization of NCS2 Proteins", Retreat of the CRC807 Integrated Research Group Edersee, Germany