Insight into the complete substrate-binding pocket of ThiT by chemical and genetic mutations

Swier, L. J. Y. M.; Monjas, L.; Reeßing, F.; Oudshoorn, R. C.; Aisyah, Aisyah; Primke, T.; Bakker, M. M.; Van Olst, E.; Ritschel, T.; Faustino, I.

Published in:
MedChemCommun

DOI:
10.1039/C7MD00079K

Publication date:
2017

MedChemComm

RESEARCH ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Med. Chem. Commun.*, 2017, **8**, 1121

Insight into the complete substrate-binding pocket of ThiT by chemical and genetic mutations†‡

L. J. Y. M. Swier,§[a] L. Monjas,§[b] F. Reeßing,[b] R. C. Oudshoorn,[a] Aisyah,[a] T. Primke,[a] M. M. Bakker,[b] E. van Olst,[a] T. Ritschel,[c] I. Faustino,[a] S. J. Marrink,[a] A. K. H. Hirsch*[b] and D. J. Slotboom*[a]

Received 17th February 2017,
Accepted 24th March 2017

DOI: 10.1039/c7md00079k

rsc.li/medchemcomm

Energy-coupling factor (ECF) transporters are involved in the uptake of micronutrients in bacteria. The transporters capture the substrate by high-affinity binding proteins, the so-called S-components. Here, we present the analysis of two regions of the substrate-binding pocket of the thiamine-specific S-component in *Lactococcus lactis*, ThiT. First, interaction of the thiazolium ring of thiamine with residues Trp34, His125 and Glu84 by π–π-stacking and cation–π is studied, and second, the part of the binding pocket that extends from the hydroxyl group. We mutated either the transported ligand (chemically) or the protein (genetically). Surprisingly, modifications in the thiazolium ring by introducing substituents with opposite electronic effects had similar effects on the binding affinity. We hypothesize that the electronic effects are superseeded by steric effects of the added substituents, which renders the study of isolated interactions difficult. Amino acid substitutions in ThiT indicate that the electrostatic interaction facilitated by residue Glu84 of ThiT and thiamine is necessary for picomolar affinity. Deazathiamine derivatives that explore the subpocket of the binding site extending from the hydroxyl group of thiamine bind with high affinity to ThiT and may be developed into selective inhibitors of thiamine transport by ECF transporters. Molecular-dynamics simulations suggest that two of these derivatives may not only bind to ThiT, but could also be transported.

Introduction

Energy-coupling factor (ECF) transporters belong to the family of ATP-binding cassette (ABC) transporters.[1–3] A common feature among ABC transporters is their core structure, consisting of two nucleotide-binding domains (NBDs) and two transmembrane domains (TMDs). The NBDs bind and hydrolyze ATP to ADP and P_i, and the energy released in this process is used to drive transport of substrates across the membrane. In classical ABC transporters, the TMDs are identical or highly similar proteins and they form the translocation pathway through the membrane. In the case of ECF transporters, the two TMDs are completely unrelated. One of them is called the S-component and binds the substrate with very high affinity (in the low- to sub-nanomolar range). The second TMD is called EcfT and is the domain that transfers conformational changes between the NBDs and the S-component. Together with the two NBDs (known as EcfA and EcfA′), EcfT forms the ECF module, also known as the energizing module, which provides the energy and conformational changes for completing the transport cycle.

ECF transporters have only been identified in prokaryotes, where they mediate the import of essential micronutrients such as vitamins and their precursors, amino acids and the transition metal ions Co^{2+} and Ni^{2+}.[4] Given that these transporters are essential to many pathogenic bacteria that lack biosynthetic routes to obtain these micronutrients,[2,5] they form a new target for the development of antibiotics. Over the past decade, ECF transporters have been studied extensively, and although some crystal structures of individual substrate-bound S-components and full-length ECF transporters have been reported, the knowledge about the mechanism of substrate-binding and transport is still limited.

ThiT is the S-component for the transport of thiamine (1) in *Lactococcus lactis*.[6] The crystal structure of ThiT in complex with thiamine (PDB ID: 3RLB)[7] shows that thiamine interacts

[a] *Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands. E-mail: d.j.slotboom@rug.nl; Tel: +31 50 363 4187*
[b] *Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands. E-mail: a.k.h.hirsch@rug.nl; Tel: +31 50 363 4275*
[c] *Centre for Molecular and Biomolecular Informatics (CMBI), Radboudumc, 6525 GA Nijmegen, The Netherlands*

† The authors declare no competing interests.

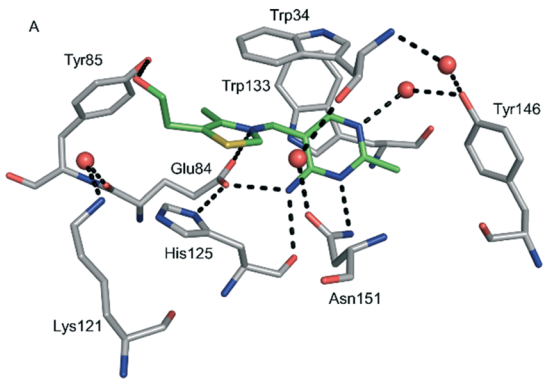
‡ Electronic supplementary information (ESI) available: Synthesis and characterization of compounds, including images of NMR spectra. See DOI: 10.1039/c7md00079k

§ These authors contributed equally to this work.

with numerous residues lining the substrate-binding pocket (Fig. 1A): the aminopyrimidinyl moiety forms hydrogen bonds with Glu84, His125, Tyr146 (*via* an ordered molecule of water) and Asn151, as well as a π–π-stacking interaction with Trp133; the thiazolium ring is involved in π–π-stacking and cation–π interactions with Trp34 and His125; the positively charged nitrogen atom of the same ring is forming an electrostatic interaction with Glu84; and the hydroxyl group is involved in a hydrogen bond with Tyr85.

In our previous work, we described the binding of thiamine analogues with modifications in the thiazolium ring, hydroxyethyl side chain and methyl group of the pyrimidinyl ring, and we showed that ThiT still binds these designed compounds with high affinity (Fig. 1B).[9,10] Here, we expand the exploration of the binding site and rationalize ligand binding to ThiT by chemical modification and amino acid substitution in the ligand and in ThiT, respectively. We focused on the interactions of two parts of the thiamine molecule with ThiT. First, we explored the π–π-stacking interaction of the thiazolium ring of thiamine with residues Trp34 and His125, and the cation–π interaction with Glu84.

Second, we extended the scaffold of compounds **2** and **3** (Fig. 1B) at the hydroxyl group. We chose this scaffold because it lacks the positive charge present in thiamine. Therefore, it may have lower chances of inhibiting human thiamine binding proteins than compound **4** (pyrithiamine).



B

	K_D ± S.D. (nM)		K_D ± S.D. (nM)
1: A = N^+, *n* = 2	0.122 ± 0.013	**4**: A = N^+, *n* = 2, R = CH_3	0.180 ± 0.070
2: A = C, *n* = 1	5.21 ± 3.19	**5**: A = C, *n* = 1, R = H	528 ± 135
3: A = C, *n* = 2	4.23 ± 1.69	**6**: A = C, *n* = 2, R = H	266 ± 131

Fig. 1 A) Binding mode of thiamine in the binding pocket of ThiT (PDB ID: 3RLB).[7] Thiamine and the residues involved in binding are shown in stick representation, with the following color code: C: green (thiamine) and gray (ThiT), O: red, N: blue and S: yellow. This color code will be maintained throughout the text. The red spheres represent water molecules and the dashed lines indicate hydrogen bonds (distance 2.7 to 3.1 Å) and an electrostatic interaction (distance 3.2 Å). This figure and other figures throughout the article were generated with PyMOL.[8] B) Structure of thiamine (**1**) and five of the previously reported thiamine analogues (**2**–**6**) with their binding affinities for ThiT.[6,9]

Even though compound **4** binds more tightly to ThiT, compounds **2** and **3** still have binding affinities in the low nanomolar range. In addition, this scaffold is not charged, which simplifies the synthesis and may provide access to more selective compounds over other thiamine- or thiamine-diphosphate-dependent proteins. This part of the substrate-binding pocket of ThiT has not been explored before and may be crucial for substrate transport. The new compounds could provide us with information about the mechanism of substrate binding and transport. In addition, they hold the potential to block thiamine transport and could therefore be developed into antibiotics. We adopted a two-pronged approach in order to obtain the new compounds: we performed *de novo* design and used the KRIPO (Key Representation of Interactions in POckets) software.[11] This software creates a pharmacophore fingerprint of the binding site to identify similar pockets in unrelated proteins that are available in the PDB, resulting in the proposal of the cocrystallized ligands as potential binders of our protein of interest. Molecular-dynamics simulations enabled us to predict whether binding of two of these new compounds would induce conformational changes within ThiT.

Results and discussion

Chemical mutations

When the thiazolium ring of thiamine (**1**) was modified, the binding affinities of the thiamine derivatives changed: in the case of a thiophenyl ring (compounds **2** and **3**), the K_D values increased about 30-fold, and in the case of a phenyl ring (compounds **5** and **6**) about 2000- to 4000-fold in comparison to thiamine (Fig. 1B).[9] In previously reported crystal structures,[9] we observed that the π–π-stacking interactions of these aromatic rings with Trp34 were slightly different. In order to understand these differences in affinity, we investigated if the electronic properties of the aromatic ring could be responsible for the differences observed in affinity. To do so, we designed two derivatives of compound **5** with modifications in the phenyl ring based on the crystal structure of ThiT in complex with this thiamine analogue (**5**) (Fig. 2A).

Inspection of the binding mode of **5** in the pocket of ThiT (PDB ID: 4N4D)[9] revealed that most of the interactions observed for thiamine (PDB ID: 3RLB)[7] are maintained (Fig. 2A *vs.* Fig. 1A). The only difference is that in compound **5**, the phenyl ring has flipped over in such a way that the hydroxymethyl substituent is pointing towards Asn151 and is now forming a hydrogen bond with Asn151 instead of with Tyr85. In this way, there is space to extend the molecule in the direction of Tyr85 (from the carbon atom highlighted with an arrow in Fig. 2A), which is usually occupied by the hydroxyethyl group of thiamine, thereby avoiding steric hindrance with the residues in the pocket. We introduced two substituents with opposite electronic effects: an electron-donating methyl group and an electron-withdrawing trifluoromethyl group (compounds **7** and **8**, Fig. 2B).

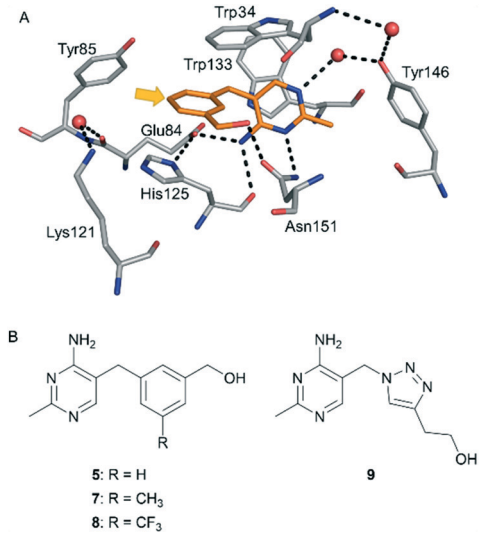
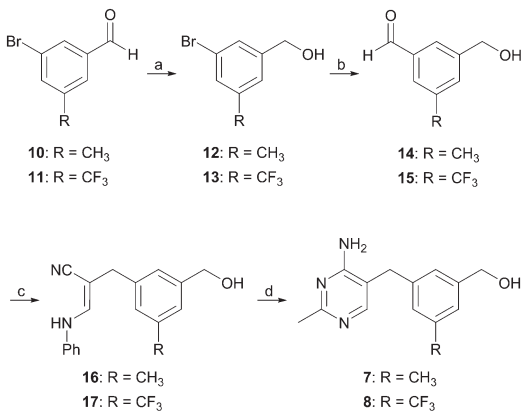


Fig. 2 A) Binding mode of compound **5** in the binding pocket of ThiT (PDB ID: 4N4D),[9] shown as in Fig. 1A, with the C atoms of compound **5** colored orange. The arrow indicates the potential extension point. B) Chemical structure of compound **5**, its derivatives **7** and **8** and the triazole derivative **9**.

We synthesized compounds **7** and **8** by following the same route used to synthesize compound **5**[9] with an additional step (Scheme 1). Commercially available aldehydes **10** and **11** were reduced to the corresponding alcohols using $NaBH_4$, in 91% and 87% yield for **12** and **13**, respectively. Treatment with ^{i}PrMgCl to achieve the *in situ* protection of the hydroxyl group, followed by formylation, afforded aldehydes **14** and **15** in 65% and 66% yield, respectively. These aldehydes were condensed with 3-anilinopropionitrile under basic conditions to form enamines **16** and **17** in 28% and 13% yield, respectively. In the final step, reaction with acetamidine under basic conditions afforded the products **7** and **8** in 87% and 30% yield, respectively.



Scheme 1 Synthesis of compounds **7** and **8**. Reagents and conditions: a) $NaBH_4$, MeOH, room temperature, 30–60 min, 91% (for compound **12**) and 87% (for compound **13**). b) i: ^{i}PrMgCl, THF, 0 °C, 10 min, ii: ^{n}BuLi, −78 to −40 °C, 1 h, iii: DMF, −78 °C to room temperature, 18 h, 65% (for compound **14**) and 66% (for compound **15**); c) 3-anilinopropionitrile, NaOMe, DMSO, microwave, 100 W, 70 °C, 30 min, 28% (for compound **16**) and 13% (for compound **17**); d) acetamidine·HCl, NaOMe, MeOH, reflux, 2 d, 87% (for compound **7**) and 30% (for compound **8**).

We determined the binding affinities of ThiT for compounds **7** and **8** by isothermal titration calorimetry (ITC). We expected one of the new compounds to be a stronger binder than compound **5**, and the other one to be a weaker binder based on the different electronic properties of the new aromatic substituents. Both compounds, however, turned out to be weaker binders than compound **5**, and displayed nearly identical K_D values of 1.99 ± 0.636 μM and 2.30 ± 1.81 μM for **7** and **8**, respectively (the errors represent the standard deviation from four experiments). Therefore, electronic effects cannot explain why a thiamine analogue bearing a thiophenyl ring binds stronger to ThiT than a phenyl derivative. Possibly, steric effects of the substituents play an important role, making it difficult to study one non-covalent interaction in isolation.

To further study this effect, we designed a new thiamine analogue bearing a triazolyl ring instead of the thiazolium ring (**9**, Fig. 2B). We synthesized compound **9** following a previously described route,[12] and we determined the binding affinity of ThiT for compound **9** by ITC, yielding a K_D value of 28.6 ± 10.1 μM (the error represents the standard deviation from three experiments). Taken together, we conclude that the high binding affinity of the thiophenyl derivatives **2** and **3** are most likely ascribed to the sulfur atom, which engages in favorable S–π interactions with the aromatic residues Trp34 and His125.

Biochemical mutations

Mutagenesis of four residues lining the binding site of ThiT (Trp34, Trp133, Tyr146 and Asn151), which are all involved in the interactions with the pyrimidinyl ring of thiamine, had been reported previously.[6] Here, we are interested in the residues that interact with the thiazolium ring of thiamine (**1**): Trp34, Glu84 and His125 (Fig. 1A). Previously, the Trp34Ala mutant proved to affect binding only to a minimal extent (Table 1).[6] We now focused on the binding of thiamine to six new ThiT mutants, in which we have modified residues Glu84 and His125 (Table 1).

Glu84 forms an electrostatic interaction with the positively charged nitrogen atom of the thiazolium ring of thiamine, as well as hydrogen bonds with the amino group of the

Table 1 Binding affinities of different ThiT mutants for thiamine, with the errors indicating the standard deviations

Mutation	K_D ± S.D. (nM)
–(Wild type)[6]	0.122 ± 0.013
Trp34Ala[6]	0.35 ± 0.05
Glu84Asp	2.21 ± 0.78[a]
Glu84Gln	(13.6 ± 3.7) × 10^3[b]
Glu84Ala	(3.65 ± 1.71) × 10^3[b]
His125Asn	13.6 ± 8.1[b]
His125Ala	14.9 ± 5.1[a]
His125Phe	No binding observed

[a] The error represents the standard deviation obtained from 4 experiments. [b] The error represents the standard deviation obtained from 5 experiments.

pyrimidinyl ring and with His125 (Fig. 1A). Replacing Glu84 by an aspartate shortens the amino acid side chain by one carbon atom, possibly increasing the distance to thiamine and thereby weakening the interaction, which could explain the 18-fold increase in K_D value. However, when mutating into a glutamine or alanine, the interaction with thiamine becomes much weaker, leading to K_D values in the micromolar range. These results highlight the importance of the negative charge on this residue and show that the resulting electrostatic interaction with thiamine appears to be the main contributor to the picomolar affinity.

Within the thiamine-binding site, the side chain of His125 can form cation–π and π–π-stacking interactions with the thiazolium ring of thiamine as well as a hydrogen bond with the carbonyl side chain of Glu84 (Fig. 1A). In case of the His125Asn mutant, the hydrogen bond with the carbonyl group in the side chain of Glu84 can still be conserved, while this hydrogen bond is disrupted in the His125Ala mutant. The similar K_D values for thiamine in both mutants (low-nanomolar range), suggest little or no contribution of this hydrogen bond to substrate binding (Table 1). In both mutants the K_D values for thiamine are 100-fold higher compared to wild-type ThiT. The loss in affinity seems to be due to the disruption of the cation–π and π–π-stacking interactions with the thiazolium ring of thiamine. When mutating His125 to a phenylalanine, no binding was observed. Compared to histidine, phenylalanine is unable to form hydrogen bonds. Besides, phenylalanine is bigger and could partially occupy the thiazolium-binding pocket, thereby preventing thiamine from binding. We tried to confirm these hypotheses by structural characterization but co-crystallization attempts were unsuccessful.

From these mutagenesis studies, we conclude that Glu84 is very important for high-affinity binding. Modification of the interaction with Trp34 or His125 has a smaller effect, most probably because these residues both form cation–π and π–π-stacking interactions with the thiazolium ring, enabling one to compensate for loss of interaction by the other one. These results support the hypothesis that thiamine binds to ThiT with high affinity when the loop L1 is 'open'.[13] As a result, even if the π–π-stacking interaction with Trp34 stabilizes the ligand in the substrate-binding pocket, the interaction is not crucial for subnanomolar affinity given that His125 can form cation–π and π–π-stacking interactions with the thiazolium ring as well. When thiamine is bound and the loop L1 closes over the substrate-binding pocket, Trp34 now participates in this cation–π and π–π-stacking interaction network too, possibly helping to keep the S-component in the closed state until it has toppled over and substrate release on the other side of the membrane is triggered.

Thiamine derivatives to explore an unoccupied subpocket

When we analyzed the substrate-binding pocket of ThiT, we observed that the large pocket is only partially occupied by thiamine and extends in the direction of the hydroxyethyl substituent of thiamine (Fig. 3). It is of interest to know whether this empty subpocket plays an important role in substrate transport, given that molecules occupying this part of the substrate-binding pocket may block transport. Considering that this subpocket is unique for ThiT, new compounds that extent into it may be selective for the S-component, and may not interact with human thiamine transporters or thiamine-dependent enzymes. Therefore, such compounds could be the basis for development of new antimicrobial agents.

Design of thiamine derivatives to occupy the subpocket. In order to design compounds that would form additional interactions with this subpocket, we used the software MOLOC[14] for molecular modeling and the FlexX docking module[15] and HYDE scoring function[16,17] of the LeadIT suite to predict binding poses and to estimate the Gibbs free energies of binding of such compounds. In our design, we maintained the thiophenyl ring of our previously reported best binders (compounds **2** and **3**, Fig. 1B). ThiT has high binding affinity for thiamine diphosphate (K_D = 1.6 ± 0.0 nM),[6] which carries two additional phosphate groups at the end of the original hydroxyethyl side chain of thiamine. When adding these phosphate groups to our compound **3** to form deazathiamine diphosphate **18** (Fig. 4), the binding affinity decreases 50-fold compared to the affinity for compound **3**, to a K_D value of 209 ± 98 nM (Table 2). This result suggests that extension of the thiophenyl ring into the large subpocket is possible, and even though we do not improve the affinity, the compounds might interfere with transport and be more selective. Our first approach to develop extended molecules with nanomolar affinity was *de novo* design, by performing modeling and docking using the crystal structure of ThiT in complex with its natural substrate thiamine (**1**) (PDB ID: 3RLB).[7] Given that the unexplored subpocket is rather big, we first investigated if one aromatic ring would improve binding or interfere with the loop L1. This led to the design of compounds **19** and **20** (Fig. 4), bearing a phenyl or an indolyl ring. These rings engage in π–π-stacking interactions with Trp63 and Tyr85, or with Trp34, respectively, as predicted by modeling and docking (Table 2, ΔG_{est} = −55 and −56 kJ mol^{-1}, respectively). Careful visual inspection revealed several small cavities in the pocket. In order to expand into two of those

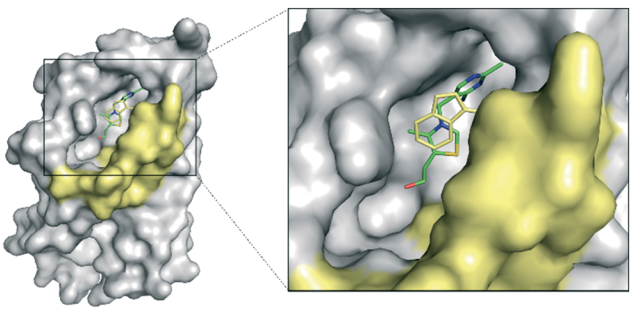


Fig. 3 Structure of ThiT in complex with thiamine (PDB ID: 3RLB).[7] The surface of ThiT is shown in gray, with the surface of the loop L1 (residues Leu26–Ile39) shown in pale yellow and Trp34 shown in stick representation to facilitate visualization of the substrate-binding pocket.

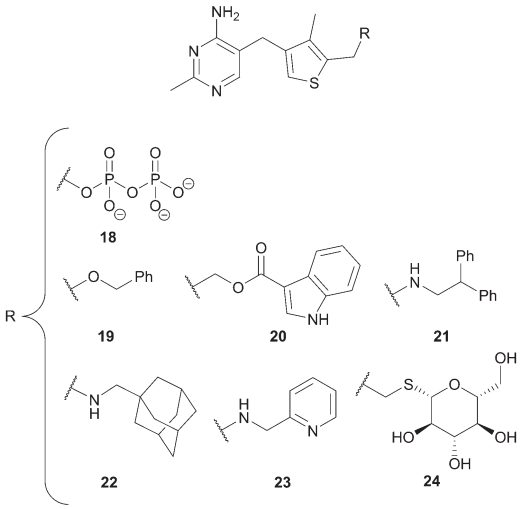


Fig. 4 Structure of deazathiamine diphosphate (**18**) and the compounds chosen to interact with the subpocket of ThiT obtained by *de novo* design (**19**–**21**) or by using KRIPO (**22**–**24**).

cavities at the same time within one molecule, we designed compound **21** (ΔG_{est} = −61 kJ mol^{-1}).

With the aim of improving our initial design, we used the software KRIPO[11] to identify fragments that could bind with high affinity in the unexplored subpocket. We performed three individual searches with KRIPO by dividing the unoccupied part of the pocket into three smaller subpockets (Fig. 5).

In case of subpocket A, mainly hydrophobic fragments were found, and we selected an adamantyl moiety as the best fitting fragment. For subpocket B, we observed that a lot of hits were heterocycles containing nitrogen atoms and sugar moieties. As a result, we selected a pyridyl ring and a glucose moiety as representative fragments. The search with subpocket C led to a large variety of compounds, but when we modeled and docked these compounds into ThiT, we observed numerous repulsions with several residues. Therefore, we decided to focus on subpockets A and B. Taking into account the distance between our deazathiamine scaffold (compounds **2** and **3**) and the fragments identified by KRIPO, we designed synthetically accessible linkers, resulting in compound **22** (for pocket A, ΔG_{est} = −57 kJ mol^{-1}) as well as compounds **23** and **24** (both for pocket B, with ΔG_{est} = −48 and −59 kJ mol^{-1}, respectively) (Fig. 4).

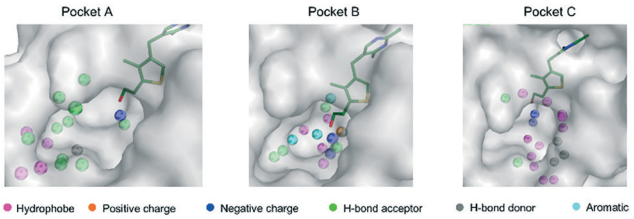


Fig. 5 Pharmacophore features of the three small subpockets A, B and C within ThiT.

Synthesis. We synthesized the designed compounds **19**–**24** using **2** and **3** as starting materials, which were synthesized following previously reported procedures.[9] In all cases, we performed the reactions once in order to obtain sufficient amount of compounds for the biological evaluation and the conditions are not optimized.

We synthesized compounds **19** and **21**–**23** starting from compound **2** (Scheme 2). Reaction of **2** and benzyl bromide using NaH afforded compound **19** in 27% yield (*N*-benzylated and dibenzylated compounds were also obtained as side products). Aldehyde **25** was obtained by oxidation of **2** with MnO_2, as previously described.[9] This aldehyde was used for the synthesis of compounds **21**–**23**, which were synthesized by reductive amination, using the corresponding amine followed by treatment with $NaBH_4$, in 32–66% yield.

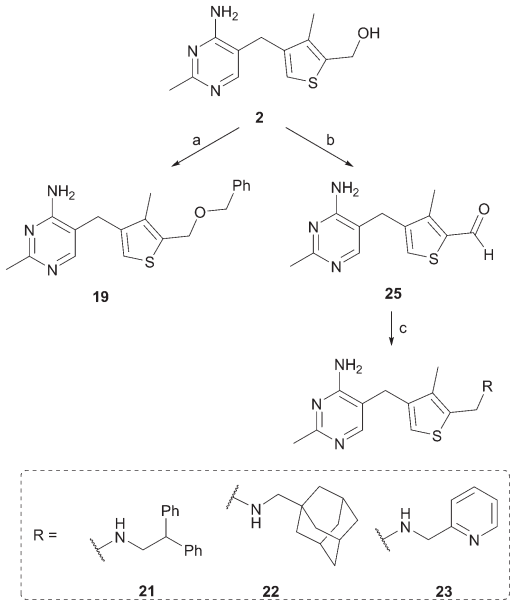
We synthesized compounds **20** and **24** starting from deazathiamine (**3**) (Scheme 3). For the synthesis of compound **20**, tosylation of **3** as described to obtain tosylate **26**,[18] followed by reaction of the resulting compound with indole-3-carboxylic acid using 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) and tetrabutylammonium iodide (TBAI), afforded **20** in 6% yield after purification by preparative HPLC. We synthesized the glucose derivative **24** using the same intermediate **26**, by reaction with 1-thio-β-D-glucose tetraacetate, DBU and TBAI, which afforded compound **27** in 32% yield. After deprotection of the acetate groups under basic conditions, we obtained compound **24** in 45% yield.

Binding-affinity determination. We determined the binding affinities of ThiT for the new compounds by an intrinsic-protein-fluorescence titration assay (compounds

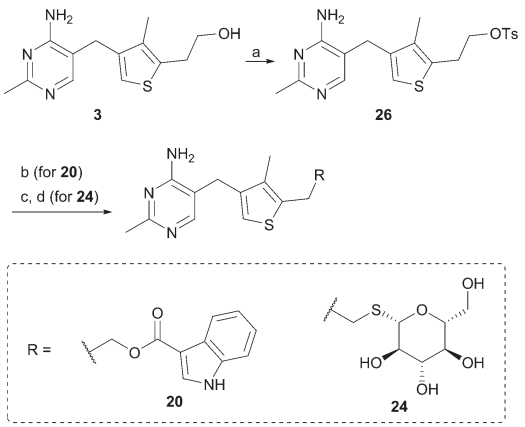
Table 2 Binding affinities of ThiT for compounds **18**–**24**, with the errors indicated as standard deviations, together with the experimental (ΔG_{exp}) and the estimated Gibbs free energies of binding (ΔG_{est}). The estimated values are based on the scoring function HYDE

Compound	K_D ± S.D. (nM)	ΔG_{exp} (kJ mol^{-1})	ΔG_{est} (kJ mol^{-1})
18	209 ± 98[a,e]		
19	40.3 ± 15.7[a,c]	−42	−55
20	20.1 ± 3.3[a,c]	−44	−56
21	168 ± 113[a,d]	−39	−61
22	(49.8 ± 27.3) × 10^3[b,c]	−25	−57
23	96.7 ± 23.7[a,d]	−40	−48
24	183 ± 17[a,d]	−38	−59

[a] Binding affinity measured by intrinsic-protein-fluorescence titration assay. [b] Binding affinity measured by ITC. [c] The error represents the standard deviation obtained from 3 experiments. [d] The error represents the standard deviation obtained from 4 experiments. [e] The error represents the standard deviation obtained from 5 experiments.



Scheme 2 Synthesis of compounds **19** and **21**–**23**, using **2** as starting material. Reagents and conditions: a) i: NaH, DMF, room temperature, 1 h; ii: BnBr, DMF, room temperature, 1 h, 27%; b) MnO_2, DMF, room temperature, 2 h, 70%; c) i: amine = 2,2-diphenylethylamine (for **21**), 1-adamantanemethylamine (for **22**) or 2-(aminomethyl)pyridine (for **23**); DMF, MeOH, $MgSO_4$, 60 °C, 20 h; ii: $NaBH_4$, room temperature, 20 h, 61% (for **21**), 32% (for **22**), 66% (for **23**).



Scheme 3 Synthesis of compounds **20** and **24**, using **3** as starting material. Reagents and conditions: a) *p*-TsCl, pyridine, 0 °C, 1 h, 73%; b) indole-3-carboxylic acid, DBU, TBAI, CH_3CN, room temperature, 3 d, 6%; c) 1-thio-β-D-glucose tetraacetate, DBU, TBAI, CH_3CN, room temperature, 4 d, 32%; d) NaOMe, MeOH, room temperature, 17 h, 45%.

18–21, 23 and 24) or by ITC (compound 22) (Table 2). Most of the compounds bind to ThiT with high affinity, featuring K_D values in the nanomolar range, despite the large modifications introduced with respect to the structure of thiamine. The binding affinities observed experimentally, however, do not match our prediction as well as in our previous work. A possible explanation for the observed discrepancy is that upon ligand binding, the explored subpockets may display greater flexibility compared with the rigid thiamine-binding pocket.

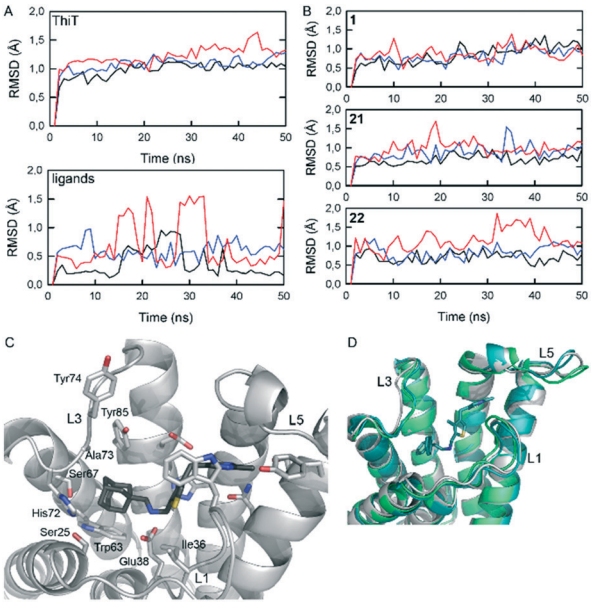


Fig. 6 A) Dynamics of ThiT and the bound ligands (thiamine (**1**): black, **21**: blue, **22**: red), with RMSD variations of the backbone atoms of ThiT and the ligand atoms plotted as a function of the simulation time of 50 ns. B) Dynamics of the residues in loops L1 (black), L3 (red) and L5 (blue), with the RMSD variations plotted as in A. C) Final snapshot of the simulation of **22** bound to ThiT, with the residues predicted to be involved in hydrogen bonding with **22** shown in sticks. The residues involved in interactions with the adamantyl moiety of **22** are indicated. D) Overlay of the most abundant poses of thiamine (**1**)-, **21**- and **22**-bound to ThiT observed during the simulations, shown in ribbon representation and colored lime green, teal and gray, respectively. Compound **21** is shown in sticks and its carbon atoms are colored dark teal. The loops L1, L3 and L5 are indicated.

Molecular-dynamics simulations. We have performed molecular dynamics (MD) simulations of thiamine (**1**) and compounds **21** and **22** in the substrate-binding pocket of ThiT to gain further insight into the ligand–protein interactions. We observed that during the simulation time (50 ns), the protein structure remains stable when thiamine or **21** are bound. However, **22** generates changes in the conformation of ThiT with respect to the original crystal structure (PDB ID: 3RLB) (Fig. 6A). The same trend is observed when we look at the RMSD variations of the ligands (Fig. 6A). While thiamine and **21** show on average relative stability, **22** shows higher dynamic behavior compared to thiamine and compound **21**. These results indicate that in the case of compound **22**, both the protein and the ligand need some time to rearrange. In fact, when we look at the dynamics of loops L1, L3 and L5 of ThiT (Fig. 6B), we see some movement of loop L3 in the simulation of **22**-bound ThiT. This observation can be explained by the fact that the adamantyl fragment mainly interacts with residues in helix 3 (Trp63, Ser67), helix 4 (Glu84, Tyr85) and loop L3 (His72, Ala73, Tyr74) (Fig. 6C). Evaluation of the MD simulation suggests that the main difference after 40 ns is a conformational change of the side chain of Tyr74 moving away from **22**. Interestingly, a conformational change of this loop L3 plays a role during the transport cycle of the ECF-FolT transporter for folate from *Lactobacillus delbrueckii*.[19] The fact

that it seems to be more difficult for 22 to accommodate in the substrate-binding pocket of ThiT might explain the weaker affinity of ThiT for this compound (K_D value in the upper micromolar range).

In agreement with our predicted binding poses from modeling and docking, compounds 21 and 22 display a similar binding mode in the simulations, maintaining the thiamine scaffold in the same conformation as in the case of the MD simulation and crystal structure with thiamine. The new fragments (two phenyl rings and one adamantyl moiety for 21 and 22, respectively) occupy the new subpocket (Fig. 6C and D). We have tried to confirm the binding pose of compound 21 by co-crystallizing of ThiT with this compound, but no crystals were obtained. Co-crystallization have been performed with compounds 19, 20, 23 and 24 too, but only in the case of compounds 20 and 23, crystals have been obtained, which diffracted up to 8 Å resolution. Attempts to improve this resolution have not been successful.

Conclusions

The chemical modifications of thiamine indicate that the electronic properties of the aromatic ring occupying the thiazolium-binding pocket are not the only crucial factor for high binding affinity, given that analogues bearing groups with opposite electronic properties (compounds 7 and 8) display nearly the same binding affinity for ThiT. The fact that these compounds are bulkier than our reference compound 5 cannot be the only reason why their binding affinities decrease, given that an analogue of thiamine bearing a triazolyl ring (compound 9), also shows a binding affinity in the micromolar range. Therefore, the sulfur atom of the thiophenyl ring (compounds 2 and 3) appears to be crucial for high-affinity binding, probably due to S–π interactions that it undergoes with Trp34 and His125.

The amino acid substitutions in ThiT indicate that Glu84 is crucial for high-affinity binding. We observed that when the electrostatic interaction with the positively charged nitrogen atom of thiamine is disrupted or weakened, the binding affinity decreases by one to five orders of magnitude. This result is consistent with the binding affinity of wild-type ThiT for deazathiamine (3), in which the electrostatic interaction is not possible either.[9] His125 can, however, be replaced by small residues without aromatic properties, although the binding affinity decreases about 18-fold when the cation–π and π–π-stacking interactions are abolished. To study if any aromatic residue in this position would maintain the subnanomolar affinity, we mutated His125 to a phenylalanine, but no binding was observed, probably because phenylalanine is bigger and does not allow thiamine to fit into the pocket.

We have shown that *de novo* design and the use of KRIPO are efficient methods for the design of ligands for ThiT. The new compounds were predicted to bind in the thiamine-binding pocket, and at the same time, the substituents were appended onto the deazathiamine scaffold to occupy a subpocket of ThiT that had not been studied up to now. Most of the novel compounds bind to ThiT in the nanomolar range, except for compound 22 that has an adamantyl group occupying the new subpocket and binds in the upper micromolar range. It is remarkable that these extended thiamine derivatives are still able to bind with such a high affinity. Given that the concentration of thiamine in many natural environments of prokaryotes is rather low, it might be possible for the thiamine analogues to efficiently compete for binding to the S-component, despite their lower affinity.

The MD simulations provided insight into the binding of 22 to ThiT, showing that ThiT has to be more flexible in order to accommodate 22 in the binding pocket compared to the simulations with thiamine and 21. Comparing the superimposition, however, revealed no significant differences, suggesting that the ECF-ThiT transporter should be able to transport 22 into the cell.

To conclude, we have shown that it is possible to extend molecules into the new subpocket of ThiT, and we already obtained K_D values in the nanomolar range for compounds that explore this new pocket. These results show that with the information on substrate binding and our workflow of design, synthesis, biochemical evaluation and MD simulations of the thiamine derivatives, we have obtained tools to successfully extend thiamine derivatives in such a way that substrate binding and probably transport are still possible. These tool compounds, including the thiamine derivatives that we have described, are useful for further studies on vitamin transport, not only by ECF transporters but also by unrelated protein families.[20,21]

Experimental part

Modeling and docking

The crystal structure of ThiT in complex with thiamine (PDB ID: 3RLB) was used.[7] Thiamine derivatives were designed by using the program MOLOC,[14] and the energy of the system was minimized by use of the MAB force field implemented in this software, while the protein coordinates and the crystallographically localized water molecule (HOH196) were kept fixed. Hydrogen bonds and hydrophobic interactions were measured in MOLOC. The designed thiamine derivatives were subsequently docked into the binding pocket of ThiT with the aid of the FlexX docking module in the LeadIT suite.[15] During docking, the binding site in the protein was restricted to 8.0 Å (for compounds 7–9) or to 12.5 Å (for compounds 19–24) around the cocrystallized thiamine, the 30 top-scored solutions were retained and subsequently post-scored with the scoring function HYDE.[16,17] After careful visualization to exclude poses with significant inter- or intramolecular clash terms or unfavorable conformations, the resulting solutions were subsequently ranked according to their binding energies.

KRIPO search

For the KRIPO search,[11] the March 2013 version of the KRIPO database of target–ligand complexes was used, which

contained all protein–ligand complexes deposited in the PDB at that time. The pocket of ThiT was divided in three subpockets, and the pharmacophore used by KRIPO in the three queries consisted of the following features:

- Subpocket A: hydrophobic contacts from Phe22 and Leu26; negative charge from Glu38; hydrogen bond acceptors from Ser25 (OH side chain and CO backbone), Glu38, Ser67, His72 (N imidazole side chain and CO backbone), Tyr74 and Tyr85; and hydrogen bond donors from Trp63.

- Subpocket B: hydrophobic contacts from Trp63, Ala88 and Pro89; positive charge from Lys121; negative charge from Glu38; hydrogen bond acceptors from Glu38, Ser67 and Glu85 (CO backbone); and aromatic interactions from Trp63, Tyr85 and His125.

- Subpocket C: hydrophobic contacts from Trp34, Ile36, Ile39, Ala40, Trp63 and Tyr74; negative charge from Glu38; hydrogen bond acceptors from Ser25 (CO backbone) and Tyr85; hydrogen bond donors from Asn29 and Ala47 (NH backbone).

Synthesis

General methods. All reagents were purchased from Sigma-Aldrich, Acros Organics, TCI Europe or Fluorochem, and were used without further purification unless noted otherwise. All solvents were reagent-grade, and if necessary, dried and distilled prior to use. All reactions were carried out under a nitrogen atmosphere (if not otherwise indicated), with use of dried glassware. Reactions were monitored either by GC-MS (GCMS-QP2010 Shimadzu) with a HP-5 column (Agilent Technologies) or by thin-layer chromatography (TLC) on silica-gel-coated aluminum foils (silica gel 60/Kieselguhr F254, Merck). Flash-column chromatography was performed on silica gel (SiliCycle 40–63 μm). Melting points were determined with a Büchi B-545 apparatus. NMR spectra were recorded on a Varian AMX400 spectrometer at 25 °C. Chemical shifts (δ) are reported in ppm relative to the residual solvent peak. Splitting patterns are indicated as (s) singlet, (d) doublet, (t) triplet, (q) quartet, (m) multiplet and (br) broad. Coupling constants (J) are reported in Hertz (Hz). FT-IR spectra (neat) were recorded on a Perkin Elmer FT-IR spectrometer. High-resolution mass spectra (HRMS) were recorded on a Thermo Scientific LTQ Orbitrap-XL mass spectrometer. Compounds **2**,[9] **3**,[9] **9**,[12] **25**[9] and **26**[18] were synthesized according to reported procedures and their spectroscopic data are in agreement. Compound **18** was kindly provided by Prof. F. J. Leeper (University of Cambridge, UK) and it was synthesized according to a reported procedure.[18] Details of the synthesis and characterization of compounds **7**, **8**, **19–24** and their intermediates are described in the ESI.‡

Mutagenesis of wild-type ThiT

For cloning and mutagenesis purposes, the gene encoding ThiT was placed in the pREnHis plasmid containing an *N*-terminal His_8-tag.[6] In order to introduce the desired mutations, site-directed mutagenesis was performed with the primers given in Table 3. After verification by DNA sequencing (Seqlab and GATC, Germany), the mutated pREnHis ThiT plasmids were converted to mutated pNZnHis ThiT plasmids using the vector backbone exchange protocol in order to be used as expression vectors in *L. lactis*.[22]

Expression and purification of wild-type ThiT and ThiT mutants

The expression and purification of substrate-free wild-type ThiT and the ThiT mutants were performed as described previously,[9] with the cultivation and expression of the ThiT mutants being performed semi-anaerobically in 2 L of chemically defined medium[23] without thiamine and supplemented with glucose (2.0%, w/v) and chloramphenicol (5 mg L^{-1}) in a 3 L bioreactor (Applikon) instead of a 10 L bioreactor.

Binding-affinity determination

For compounds **18–21**, **23** and **24**, the binding affinity was determined using the intrinsic-protein-fluorescence titration assay described previously,[9] with 50 nM of ThiT in a final volume of 1000 μL of buffer (KP_i (pH 7.0, 50 mM), KCl (150 mM), *n*-decyl-β-D-maltopyranoside (DM, Anatrace, 0.15%, w/v)). For compounds **7**, **8**, **9** and **22**, as well as for the ThiT mutants, the binding affinity was determined by ITC as described previously[10] with 4.80–26.0 μM of ThiT. Various concentrations of the compounds, or thiamine in case of the ThiT mutants, were added in steps of 1 μL (Table 4), while maintaining equal volume percentages of DMSO in both the protein and the substrate solution when determining the binding affinities for the compounds.

Molecular-dynamics simulations

Atomistic MD simulations were performed with three ThiT ligands (thiamine (**1**), **21** and **22**). The crystal structure PDB ID: 3RLB (thiamine) and the predicted docking binding poses of compounds **21** and **22** were used as initial structures and

Table 3 Primers for mutagenesis of wild type ThiT, with the mutated bases underlined

Mutation	Forward primer	Reverse primer
Glu84Asp	tcccaagctttccttgattatcttg	caagataatcaaggaaagcttggga
Glu84Gln	tcccaagctttccttcaatatcttg	caagatattgaaggaaagcttggga
Glu84Ala	tcccaagctttccttgcatatcttg	caagatatgcaaggaaagcttggga
His125Phe	cttaaatacttttttcttttcattgccggaattattttctggagcc	ggctccagaaaataattccggcaatgaaaaagaaaaagtatttaag
His125Asn	cttaaatacttttttcaatttcattgccggaattattttctggagcc	ggctccagaaaataattccggcaatgaaattgaaaaagtatttaag
His125Ala	cttaaatacttttttcgctttcattgccggaattattttctggagcc	ggctccagaaaataattccggcaatgaaagcgaaaaagtatttaag

Table 4 Conditions under which ITC measurements were performed

Compound/mutation	[protein] (μM)	[substrate] (μM)	[substrate]:[protein] ratio
7	9.02–14.7	250–375	20–30
8	10.2–16.8	204–322	20
9	3.54–9.42	354–952	100–101
22	16.1	1.67×10^3	103
Glu84Asp	4.80–9.04	9.6–120	2–13.3
Glu84Gln	8.06–13.8	806–1.38×10^3	100
Glu84Ala	6.19–10.7	200–350	20–35
His125Phe	9.48–16.2	162–974	10–100
His125Asn	12.1–26.0	121–260	10
His125Ala	9.28–9.52	95.2–140	10–15.1

embedded in a lipid bilayer. To this purpose, the CHARMM Membrane Builder GUI[24] was used to build the systems and converted to the AMBER format using the charmmlipid2amber.x script.[25] The lipid bilayer was composed of 150 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) lipids and the systems used TIP3P[26] water model and had 0.15 M NaCl concentration added to the water layer. The ligands were parameterized using the GAFF force field and the antechamber module included the AmberTools package (version 15).[27,28] The prepared systems were minimized and equilibrated by initially fixing the coordinates of the complex (10 kcal mol^{-1} $Å^{-1}$) and reducing the constraints gradually until complete relaxation.[29] Both the equilibration and production phases were run with AMBER14[25] using the NPT ensemble (1 atm, 303 K) with periodic boundary conditions and the Langevin thermostat and semi-isotropic pressure scaling. Bonds involving hydrogens were constrained using the SHAKE algorithm,[30] allowing a 2 fs time step. PME was used to account for the electrostatic interactions beyond a cutoff of 10 Å.

Author contributions

L. J. Y. M. S, L. M., A. K. H. H and D. J. S. designed the experiments. L. M., F. R. and M. M. B performed modeling, docking and synthesis. L. J. Y. M. S., A. S. and R. C. O. performed mutagenesis experiments. L. J. Y. M. S., R. C. O., A. S., T. P. and E. O. performed protein expression and purification, and determination of binding affinities. T. R. performed the KRIPO search. I. F. and S. J. M. performed MD simulations. L. J. Y. M. S, L. M., A. K. H. H. and D. J. S. analyzed the data and wrote the manuscript.

Acknowledgements

The authors acknowledge Prof. F. J. Leeper (University of Cambridge, UK) for providing compound **18**. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grant agreement no. 283570), the Ministry of Education, Culture and Science (Gravitation program 024.001.035), the Netherlands Organisation for Scientific Research (NWO) (NWO ChemThem grant 728.011.104, NWO Vidi grant 723.014.008 and NWO Vici grant 865.11.001) and the European Research Council (ERC) (ERC Starting Grant 282083).

References

1 D. A. Rodionov, P. Hebbeln, M. S. Gelfand and T. Eitinger, *J. Bacteriol.*, 2006, **188**, 317–327.
2 D. A. Rodionov, P. Hebbeln, A. Eudes, J. ter Beek, I. A. Rodionova, G. B. Erkens, D. J. Slotboom, M. S. Gelfand, A. L. Osterman, A. D. Hanson and T. Eitinger, *J. Bacteriol.*, 2009, **91**, 42–51.
3 D. J. Slotboom, *Nat. Rev. Microbiol.*, 2014, **12**, 79–87.
4 T. Eitinger, D. A. Rodionov, M. Grote and E. Schneider, *FEMS Microbiol. Rev.*, 2011, **35**, 3–67.
5 C. T. Jurgenson, T. P. Begley and S. E. Ealick, *Annu. Rev. Biochem.*, 2009, **78**, 569–603.
6 G. B. Erkens and D. J. Slotboom, *Biochemistry*, 2010, **49**, 3203–3212.
7 G. B. Erkens, R. P.-A. Berntsson, F. Fulyani, M. Majsnerowska, A. Vujičić-Žagar, J. ter Beek, B. Poolman and D. J. Slotboom, *Nat. Struct. Mol. Biol.*, 2011, **18**, 755–760.
8 W. L. DeLano, 2002, http://www.pymol.org.
9 L. J. Y. M. Swier, L. Monjas, A. Guskov, A. R. de Voogd, G. B. Erkens, D. J. Slotboom and A. K. H. Hirsch, *ChemBioChem*, 2015, **16**, 819–826.
10 L. Monjas, L. J. Y. M. Swier, A. R. de Voogd, R. C. Oudshoorn, A. K. H. Hirsch and D. J. Slotboom, *Med. Chem. Commun.*, 2016, **7**, 966–971.
11 D. J. Wood, J. de Vlieg, M. Wagener and T. Ritschel, *J. Chem. Inf. Model.*, 2012, **52**, 2031–2043.
12 K. M. Erixon, C. L. Dabalos and F. J. Leeper, *Org. Biomol. Chem.*, 2008, **6**, 3561–3572.
13 M. Majsnerowska, I. Hänelt, D. Wunnicke, L. V. Schäfer, H. J. Steinhoff and D. J. Slotboom, *Structure*, 2013, **21**, 861–867.
14 P. R. Gerber and K. Müller, *J. Comput.-Aided Mol. Des.*, 1995, **9**, 251–268.
15 *LeadIT (version 2.1.8)*, BioSolveIT GmbH, An Der Ziegelei 79, 53757 St. Augustin, Germany, 2014.
16 I. Reulecke, G. Lange, J. Albrecht, R. Klein and M. Rarey, *ChemMedChem*, 2008, **3**, 885–897.
17 N. Schneider, S. Hindle, G. Lange, R. Klein, J. Albrecht, H. Briem, K. Beyer, H. Claußen, M. Gastreich, C. Lemmen and M. Rarey, *J. Comput.-Aided Mol. Des.*, 2012, **26**, 701–723.
18 S. Mann, C. Perez Melero, D. Hawksley and F. J. Leeper, *Org. Biomol. Chem.*, 2004, **2**, 1732–1741.
19 L. J. Y. M. Swier, A. Guskov and D. J. Slotboom, *Nat. Commun.*, 2016, **7**, 11072.
20 H. J. Genee, A. P. Bali, S. D. Petersen, S. Siedler, M. T. Bonde, L. S. Gronenberg, M. Kristensen, S. J. Harrison and M. O. A. Sommer, *Nat. Chem. Biol.*, 2016, **12**, 1015.
21 M. Jaehme and D. J. Slotboom, *Biochim. Biophys. Acta*, 2015, **1850**, 565–576.

22 E. R. Geertsma and B. Poolman, *Nat. Methods*, 2007, **4**, 705–707.
23 R. P.-A. Berntsson, N. Alia Oktaviani, F. Fusetti, A.-M. W. H. Thunnissen, B. Poolman and D. J. Slotboom, *Protein Sci.*, 2009, **18**, 1121–1127.
24 S. Jo, J. B. Lim, J. B. Klauda and W. Im, *Biophys. J.*, 2009, **97**, 50–58.
25 D. A. Case, *et al.*, *AMBER14*, Univ. California, S. F.
26 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
27 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graphics Modell.*, 2006, **25**, 247–260.
28 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
29 C. J. Dickson, B. D. Madej, A. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould and R. C. Walker, *J. Chem. Theory Comput.*, 2014, **10**, 865–879.
30 J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.

