

# **Structural Basis of Membrane Transport and Pyrophosphorylation of Thiamine in Humans**

**Dissertation for the degree of Doctor rerum naturalium**



**Florian Gabriel**

**2024**

**European Molecular Biology Laboratory**

**Universität Hamburg**

Institut für Biochemie und Molekularbiologie

Fachbereich Chemie

Fakultät für Mathematik, Informatik und Naturwissenschaften

# **Structural Basis of Membrane Transport and Pyrophosphorylation of Thiamine in Humans**

Dissertation submitted at the University of Hamburg for the degree of Doctor rerum naturalium

Institut für Biochemie und Molekularbiologie  
Fachbereich Chemie  
Fakultät für Mathematik, Informatik und Naturwissenschaften

November 2024

By Florian Gabriel, born in Trier, Germany.

## **Supervisors**

### **Dr. Christian Löw**

Structural Biology Unit Hamburg  
European Molecular Biology Laboratory (EMBL)

### **Prof. Dr. Daniel Wilson**

Institut für Biochemie und Molekularbiologie  
Fachbereich Chemie  
Fakultät für Mathematik, Informatik und Naturwissenschaften  
Universität Hamburg

# **Vorblatt 1: Gutachter\*innen und Prüfungskommissionsmitglieder**

## **Gutachter:**

Prof. Dr. Daniel Wilson

Dr. Christian Löw

## **Prüfungskommissionsmitglieder:**

Prof. Dr. Kay Grünewald

Prof. Dr. Andrew Torda

Dr. Christian Löw

## **Datum der Disputation:**

31. Januar 2025

## **Vorblatt 2: Durchführung der Dissertation**

Die Forschungsarbeit für diese Dissertation wurde von Oktober 2020 bis November 2024 in der Forschungsgruppe von Dr. Christian Löw am Europäischen Labor für Molekularbiologie (EMBL) Hamburg erbracht. Die Selektion von Nanobodies sowie die Aufnahme bestimmter Cryo-EM Datensätze habe ich während Forschungsaufenthalten an der Vrije Universiteit Brussel und dem EMBL Heidelberg durchgeführt. Die Dissertation selbst wurde von Juni 2021 bis November 2024 durch Prof. Dr. Daniel Wilson an der Universität Hamburg betreut.



## Preface

In the beginning, I would like to make a few statements regarding this Dissertation in general. Parts of the Results, specifically the ones described Sections in 3.1-3.5, were previously published in the journal Nature Communications (Gabriel et al., 2024). The corresponding sections are thus largely inspired by this paper. I also want to highlight the contributions of the following people to the presented research. A comprehensive list of acknowledgements follows the Appendix.

**Lea Spriestersbach** was a master student under my supervision from February 2022 until January 2023. She contributed significantly to the presented work. In particular, she cloned and purified the human SLC19A2, performed a thorough biophysical characterisation of the two human thiamine transporters SLC19A2 and SLC19A3, and cloned and purified SLC19A3-BRIL constructs. This work was published in her Master's Thesis (Spriestersbach 2022, University of Regensburg).

**Antonia Fuhrmann** was a trainee under my supervision from September through November 2023. She made important contributions to the investigation of drug interactions of the human SLC19A3 by thermal shift assays. She furthermore purified human SLC19A3 and nanobodies for the preparation of cryo-EM samples.

**Alison Pirro-Lundqvist**, a technical assistant at the Vrije Universiteit Brussel, performed essential work for the nanobody discovery process. Importantly, she prepared VHH phage display libraries from the B cell repertoire of immunised llamas and trained me in the use of phage display for the subsequent discovery process.

**Björn Windshügel**, Professor at the Constructor University Bremen and researcher at the Fraunhofer Institute for Translational Medicine, performed the structure-based computational screens for the discovery SLC19A3 inhibitors.

## I. List of publications

Over the course of this Dissertation, I published the following papers as first or co-author. Only the first of these is thematically connected to the Dissertation.

1. **Structural basis of thiamine transport and drug recognition by SLC19A3**  
Gabriel, F., Spriestersbach, L., Fuhrmann, A., Jungnickel, K. E. J., Mostafavi, S., Pardon, E., Steyaert, J. & Löw, C.  
*Nature Communications* (2024)
2. **MFSD1 in complex with its accessory subunit GLMP functions as a general dipeptide uniporter in lysosomes**  
Jungnickel, K. E. J., Guelle, O., Iguchi, M., Dong, W., Kotov, V., Gabriel, F., Debacker, C., Dairou, J., McCort-Tranchepain, I., Laqtom, N. N., Chan, S. H., Ejima, A., Sato, K., Massa López, D., Saftig, P., Mehdipour, A. R., Abu-Remaileh, M., Gasnier, B., Löw, C. & Damme, M.,  
*Nature Cell Biology* (2024)
3. **Crystal structure of the Na<sup>+</sup>/H<sup>+</sup> antiporter NhaA at activating pH reveals the mechanistic basis for pH sensing**  
Winkelmann, I., Uzdavinys, P., Kenney, I. M., Brock, J., Meier, P. F., Wagner, L., Gabriel, F., Jung, S., Matsuoka, R., von Ballmoos, C., Beckstein, O. & Drew, D.  
*Nature Communications* (2022)

The following manuscripts are under preparation at the time of writing. Their content is presented in this Dissertation.

4. **Structure-based discovery of inhibitors of the human SLC19A3**  
Gabriel, F., Windshügel, B. & Löw, C. (to be submitted)
5. **Cryo-EM structure of the human TPK1**  
Gabriel, F. & Löw, C. (to be submitted)
6. **Strukturbiologische Grundlage des Membrantransports und der Pyrophosphorylierung von Thiamin im menschlichen Körper**  
Gabriel, F. & Löw, C. (invited submission to *Biospektrum*)

## II. Table of content

<b>Content</b>	<b>Page</b>
Vorblatt 1: Gutachter*innen und Prüfungskommissionsmitglieder	1
Vorblatt 2: Durchführung der Dissertation	2
<b>Preface</b>	<b>3</b>
<b>I. List of publications</b>	<b>4</b>
<b>II. Table of content</b>	<b>5</b>
<b>III. List of abbreviations</b>	<b>8</b>
<b>Zusammenfassung</b>	<b>10</b>
<b>Abstract</b>	<b>11</b>
<b>Chapter 1: Introduction</b>	<b>12</b>
1.1 Chemical biology of thiamine	12
1.2 Thiamine transport across cellular membranes	16
1.3 Thiamine deficiencies in humans	18
1.4 Cryogenic electron microscopy of small proteins	19
1.5 Structural fiducials for small protein targets	26
<b>Aim of the Study</b>	<b>30</b>
<b>Chapter 2: Methods</b>	<b>31</b>
2.1 Cloning of mammalian vitamin transporters	31
2.2 Expression and purification of mammalian vitamin transporters	33
2.3 Cloning, expression, and purification of the human TPK1	34
2.4 Sybody discovery	36
2.5 Nanobody discovery and production	38
2.6 Biolayer interferometry (BLI)	40
2.7 Thermal stability assays (nanoDSF)	40
2.8 Mass spectrometry-based cellular thiamine uptake assays	41
2.9 Cryo-EM sample preparation and data collection	42

2.10 Cryo-EM data processing	43
2.11 Model building and refinement	44
<b>Chapter 3: Results</b>	<b>45</b>
3.1 Expression, purification, and biophysical characterisation of hSLC19A3	45
3.2 Development of structural fiducials for hSLC19A3	50
3.3 Structural insights into the transport cycle of hSLC19A3	60
3.4 Exploration of drug interactions of hSLC19A3	63
3.5 Structural basis of drug interactions of hSLC19A3	65
3.6 Structure-based discovery of thiamine uptake inhibitors	68
3.7 Biophysical and structural work on SLC19A2	73
3.8 Cryo-EM structure of hTPK1	78
<b>Chapter 4: Discussion and Conclusion</b>	<b>92</b>
4.1 Limitations of the study	92
4.2 Drug interactions of vitamin transporters	96
4.3 Structure determination of hSLC19A3 - a Sino-European tale	97
4.4 Structure determination of small proteins by cryo-EM	98
<b>References</b>	<b>100</b>
<b>Appendix</b>	
Appendix 1: Supplementary Figures	
Appendix 2: List of Hazardous Substances	
Appendix 3: Materials	
<b>Acknowledgements</b>	
<b>Eidesstattliche Versicherung</b>	

## List of Figures

Figure No.	Title	Page
1	Biosynthesis of thiamine in bacteria	13
2	General reaction mechanism of thiamine-dependent enzymes	14
3	Overview of thiamine's membrane transport and metabolism in the cell	15
4	Cryo-EM micrographs and their CTFs at different defoci	23
5	Structural fiducials in single particle cryo-EM	27
6	FSEC of mammalian vitamin transporters	46
7	Purification of hSLC19A3 in LMNG/CHS	48
8	Biophysical characterisation of hSLC19A3	50
9	Low resolution reconstruction of the fiducial-free thiamine-bound hSLC19A3	51
10	Design, purification, and cryo-EM analysis of hSLC19A3-BRIL fusions	53
11	Sybody discovery campaign for hSLC19A3	55
12	Nanobody discovery campaign for hSLC19A3	57
13	Structure and function of hSLC19A3	59
14	Thiamine binding site of hSLC19A3	61
15	Drug interactions of hSLC19A3	64
16	Cryo-EM structures of hSLC19A3 in complex with high-affinity inhibitors	67
17	Experimental validation of hSLC19A3 docking screen hits	70
18	Cryo-EM data processing for hSLC19A3:Nb3.3:domperidone	72
19	Cryo-EM structure of domperidone-bound hSLC19A3	73
20	Purification and drug interactions of hSLC19A2	74
21	Biophysical characterisation of hSLC19A2-specific nanobodies	75
22	Cryo-EM analysis and AlphaFold3 models of hSLC19A2-nanobody complexes	76
23	AlphaFold2 model and cryo-EM analysis of the hSLC19A2-Nb <sub>ALFA</sub> complex	77
24	Purification and thermal shift assay of hTPK1	79
25	Cryo-EM data processing for hTPK1	81
26	Cryo-EM structure of hTPK1	82
27	Fold of the hTPK1 homodimer	83
28	Domains of hTPK1	84
29	Substrate binding site of hTPK1	85
30	Untargeted mass spectrometry analysis of purified hTPK1	86
31	Untargeted MS of purified hTPK1 with a focus on thiamine-pp	87
32	Comparison of crystal and cryo-EM structures of hTPK1	88
33	Sequence and structure conservation of TPK1 in human parasites	91
34	Speculative model of proton-coupling in hSLC19A3	94

## List of Tables

Table No.	Title	Page
1	PCR reaction mixture	31
2	PCR thermocycler program	32
3	Colony PCR thermocycler program	33
4	PCR thermocycler program for the amplification of hTPK1	35
5	Overview of the screened mammalian vitamin transporters (MVTs).	45
6	Overview of the selected hit compounds from the virtual screen	69
7	Crystal structures of TPK1	88
8	Overview of disease-causing point mutations of hTPK1	90

### III. List of abbreviations

<b>Abbreviation</b>	<b>Explanation</b>
2D	Two-dimensional
3D	Three-dimensional
Ab	Antibody
ABC transporter	ATP binding cassette transporter
BLI	Biolayer interferometry
CCCP	Carbonyl cyanide m-chlorophenyl hydrazone
CMC	Critical micelle concentration
Cryo-EM	Cryogenic electron microscopy
CV	Column volume
ddH <sub>2</sub> O	Double-distilled water
DDM	n-Dodecyl- $\beta$ -D-maltopyranoside
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
<i>E. coli</i>	<i>Escherichia coli</i>
ELISA	Enzyme-linked immunosorbent assay
H <sup>+</sup>	Proton
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl $\beta$ -d-1-thiogalactopyranoside
K <sup>+</sup>	Potassium ion
K <sub>d,app</sub>	Apparent dissociation constant
kDa	Kilodalton
LB	Lysogeny broth medium
LMNG	Lauryl maltose neopentyl glycol
mAU	Milli absorption units
MES	2-(N-morpholino)ethanesulfonic acid
MFS	Major facilitator superfamily
MWCO	Molecular weight cut-off
nanoDSF	Nano differential scanning fluorometry
Nb	Nanobody
Ni-NTA	Nickel nitrilotriacetic acid
NU refinement	Non-uniform refinement
OD <sub>600</sub>	Optical density at 600 nm
PCR	Polymerase chain reaction
PDB	Protein data bank

PEG	Polyethylene glycol
psi	Pound-force per square inch
RT	Room temperature
Sb	Sybody
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SLC	Solute carrier
TAE	Tris-acetate-EDTA
TB	Terrific broth medium
TBS	Tris-buffered saline
TCEP	Tris(2-carboxyethyl)phosphine
Thiamine-p	Thiamine monophosphate
Thiamine-pp	Thiamine pyrophosphate
TM	Transmembrane helix
$T_m$	Melting temperature
TPK1	Thiamine pyrophosphokinase 1
Tris	Tris(hydroxymethyl)aminomethane
TUI	Thiamine uptake inhibitor
v/v	Volume per volume
w/v	Weight per volume
w/w	Weight per weight
wt	Wildtype
x g	Multiplicity of the standard acceleration of gravity ( $\sim 9.81 \text{ m/s}^2$ )
$\Delta T_m$	Melting temperature difference
$\lambda$	Wavelength

## Zusammenfassung

Thiamin (Vitamin B<sub>1</sub>) ist ein essenzielles Coenzym für prokaryotische und eukaryotische Zellen. In Verbindung mit thiaminabhängigen Enzymen katalysiert das Vitamin Reaktionen zentraler Stoffwechselwege wie des Citratzyklus und des Pentosephosphatwegs. Menschen, sowie die Mehrzahl aller Metazoen, haben die Fähigkeit verloren, Thiamin *de novo* zu synthetisieren. Daher muss jede Körperzelle dieses Vitamin über ihre Plasmamembran aufnehmen. Beim Menschen wird dieser Membrantransport hauptsächlich durch die Solute Carrier-Proteine SLC19A2 und SLC19A3 vermittelt. Störungen dieses Transportsystems haben schwerwiegende Auswirkungen auf die menschliche Gesundheit. Mutationen in SLC19A2 und SLC19A3 führen zu schweren bis tödlich verlaufenden Erkrankungen. Darüber hinaus wird die Inhibition von SLC19A3 durch den Januskinase-Inhibitor Fedratinib mit gravierenden neurologischen Nebenwirkungen in Form von Wernicke-Enzephalopathie in Verbindung gebracht. Trotz der physiologischen und pharmakologischen Bedeutung des Thiamintransports im Menschen war die zugrunde liegende molekulare Maschinerie auf struktureller und biophysikalischer Ebene bislang nur unzureichend verstanden. In dieser Dissertation präsentiere ich hochaufgelöste cryo-EM Strukturen des menschlichen SLC19A3 (hSLC19A3) im Komplex mit seinem endogenen Substrat Thiamin sowie den bekannten Thiaminaufnahmehemmern Fedratinib, Amprolium und Hydroxychloroquin. Die Strukturbestimmung wurde durch die Entwicklung von drei SLC19A3-spezifischen Nanobodies ermöglicht. Diese Nanobodies erlaubten mir, die Struktur des hSLC19A3 in seinen nach außen und nach innen geöffneten Konformationszuständen zu bestimmen. Mithilfe von Thermostabilisationsassays und zellulären Thiaminaufnahmeassays erweiterte ich das bekannte Spektrum der Arzneimittelinteraktionen von hSLC19A3. Dieser Ansatz führte zu der Entdeckung der thiamintransportinhibierenden Wirkung von sieben auf dem Markt befindlichen Medikamenten. Basierend auf den hier präsentierten cryo-EM Strukturen führten wir eine computerbasierte Docking-Analyse durch, um weitere Medikamente hinsichtlich ihres Potentials zur Hemmung des Thiamintransports im Menschen zu screenen. Hit Compounds aus diesem virtuellen Screening wurden experimentell auf biophysikalischer, funktioneller und struktureller Ebene validiert. Die biologische Funktion von Thiamin als Coenzym ist entscheidend von seiner intrazellulären Pyrophosphorylierung durch das Enzym TPK1 abhängig. In der vorgelegten Dissertation präsentiere ich eine cryo-EM Struktur der menschlichen TPK1 bei einer Auflösung von 2,1 Å. Diese Arbeit demonstriert das Vermögen der cryo-EM, kleine Proteinstrukturen mit Auflösungen zu bestimmen, die für eine klare biochemische Interpretation und für die strukturbasierte Arzneimittelentwicklung nützlich sind.



## Abstract

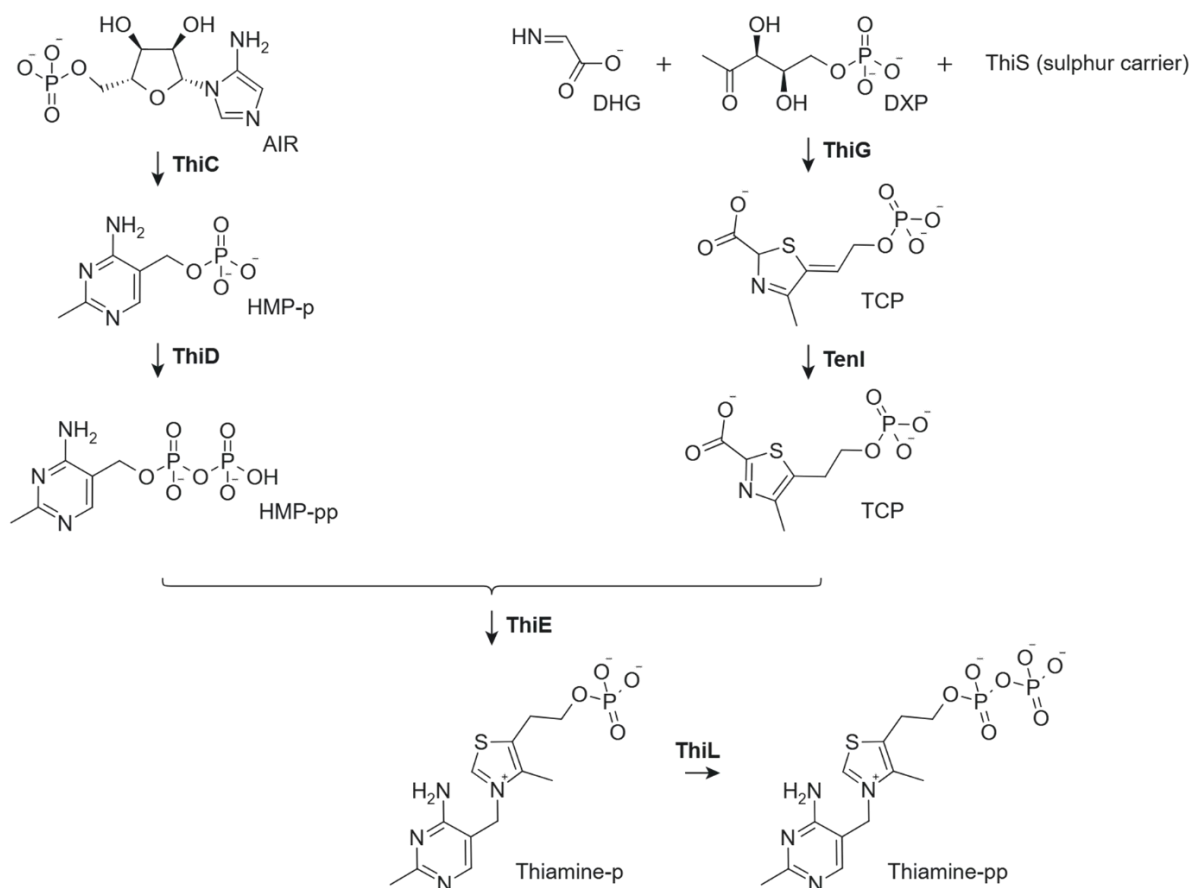
Thiamine (vitamin B<sub>1</sub>) is an essential coenzyme for prokaryotic and eukaryotic cells alike. In conjunction with thiamine-dependent enzymes, the vitamin catalyses reactions of central metabolic pathways, like the citric acid cycle and the pentose phosphate pathway. Metazoans, including humans, have lost the ability to synthesise thiamine *de novo*. Hence, every cell in these organisms needs to import the vitamin over its plasma membrane. In humans, this membrane transport is mainly mediated by the solute carrier proteins SLC19A2 and SLC19A3. Perturbation of this transport system has severe effects on human health. Loss of function mutations in SLC19A2 and SLC19A3 cause detrimental diseases. Furthermore, the inhibition of SLC19A3 by the Janus kinase inhibitor fedratinib is associated with severe neurological side effects in the form of Wernicke's encephalopathy. Despite the physiological and pharmacological importance of thiamine transport in humans, the underlying molecular machinery remained poorly understood on a structural and biophysical level. In this Dissertation, I present high-resolution cryo-EM structures of the human SLC19A3 (hSLC19A3) in complex with its endogenous substrate thiamine, as well as the known thiamine uptake inhibitors fedratinib, amprolium, and hydroxychloroquine. The structure determination was enabled by the discovery of three unique SLC19A3-specific nanobodies. The nanobodies allowed me to determine the structure of hSLC19A3 in its outward- and inward-open conformational states. Using thermal shift and cellular uptake assays, I explored the wider drug interaction spectrum of hSLC19A3. This brought about the discovery of seven inhibitors of the human thiamine transporter. Based on the determined cryo-EM structures, we performed computational docking to screen medications for their propensity to block thiamine transport in humans. Hit compounds from this virtual screen were validated experimentally on a biophysical, functional, and structural level. The biological function of thiamine as a coenzyme in the cell is critically dependent on its pyrophosphorylation by the enzyme TPK1. In this Dissertation, I present a cryo-EM structure of the human TPK1 at a resolution of 2.1 Å. This work showcases the ability of cryo-EM to determine small protein structures at resolutions useful for clear biochemical interpretation and structure-based drug design.

# 1. Introduction

## 1.1 Chemical biology of thiamine

### Biosynthesis of thiamine

Thiamine can be synthesised *de novo* by prokaryotes and many eukaryotes, including fungi, plants, and protists (Jurgenson et al., 2009, Helliwell et al., 2013). Most metazoans have, however, lost the capability of producing this essential molecule (Bettendorf et al., 2012). In bacteria, thiamine synthesis requires the orchestrated action of at least seven enzymes, and many more delivering the necessary substrates. In *E. coli*, these enzymes are encoded in the *thiCEFSGH* operon (Jurgenson et al., 2009). The aminopyrimidine and thiazole moieties of the vitamin are in this synthesised separately (Fig. 1). The substrate for the synthesis of the aminopyrimidine moiety of thiamine is 5-aminoimidazole ribotide (AIR). This molecule is converted to hydroxymethyl pyrimidine phosphate (HMP-p) by the enzyme ThiC (Webb et al., 2007). Subsequently, HMP-p is further phosphorylated by ThiD to form hydroxymethyl pyrimidine pyrophosphate (HMP-pp). At the beginning of thiazole synthesis, ThiG catalyses the addition of deoxy-D-xylulose-5-phosphate (DXP), dehydroglycine (DHG), and a sulphur atom, delivered by the sulphur carrier protein ThiS, to form the thiazole phosphate carboxylate (TPC) tautomer. The aromatisation of the TPC tautomer is subsequently catalysed by TenI. TPC and HMP-pp are then ligated by ThiE, resulting in thiamine monophosphate (thiamine-p). TMP is eventually converted to the biochemically active coenzyme thiamine pyrophosphate (thiamine-pp) by the enzyme ThiL (Fig. 1). Thiamine biosynthesis in eukaryotic organisms, like baker's yeast (*Saccharomyces cerevisiae*), occurs in an analogous manner, in the sense that the aminopyrimidine and thiazole groups are produced first and then chemically linked in a subsequent reaction step (Jurgenson et al., 2009). In yeast, the enzyme THI5 catalyses the generation of HMP from histidine and pyridoxine. HMP is then phosphorylated by THI3 or THI20 to form HMP-pp. The synthesis of the thiazole moiety from glycine, cysteine and NAD is catalysed by THI4. The product of the corresponding reaction is an ADP-ribosylated thiazole carboxylate, which is subsequently chemically conjugated with HMP-pp by the enzyme THI6 to form thiamine-p. Thiamine-p is then phosphorylated by THI80 to form the coenzyme thiamine-pp. As the brief summary of thiamine biosynthesis above shows, the *de novo* generation requires an entire enzymatic machinery dedicated to this sole purpose and consumes plenty of energy and metabolites in the form of ATP, NAD, and amino acids. This might be the primary reason why metazoans lost their ability to synthesise thiamine and adapted to extract thiamine from exogenous sources, i.e. their meals, instead. In humans and other mammals, thiamine is primarily absorbed in its un-phosphorylated form. To fulfil its biochemical role in cells, it is converted to thiamine-pp by the enzyme thiamine pyrophosphokinase 1 (TPK1).

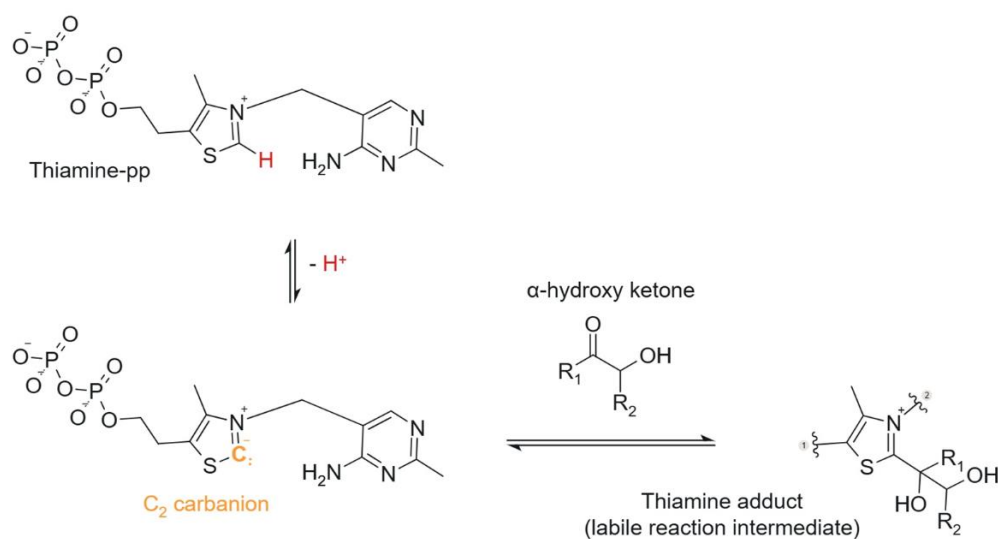


**Fig. 1: Biosynthesis of thiamine in bacteria.** The aminopyrimidine and thiazole moieties of thiamine are synthesised separately and later linked by the enzyme ThiE. Figure designed in Marvin (Chemaxon).

## Thiamine-dependent enzymes (TDEs)

Thiamine pyrophosphate (thiamine-pp) functions as a coenzyme in thiamine-dependent enzymes (TDEs). TDEs are part of central metabolic pathways and are described in more detail below. For high-affinity binding of thiamine to TDEs, its pyrophosphorylation is essential (Walsh & Tang, 2019). The pyrophosphate, together with the aminopyrimidine ring of thiamine, act as molecular anchors tethering the coenzyme stably, but non-covalently to the respective TDE (Fullam et al., 2012). The catalytic activity of thiamine-pp resides in the C<sub>2</sub> atom of its thiazolium ring (Fig. 2). This position in the heterocycle was found to undergo proton exchange, which requires the presence of a C<sub>2</sub> carbanion intermediate species in aqueous solution. This C<sub>2</sub> carbanion is electronically stabilised by the adjacent cationic nitrogen and through its resonance with a more stable carbene configuration of the thiazolium ring. The C<sub>2</sub>-H of free thiamine-pp is only weakly acidic, with a proton exchange rate of about 1 ms<sup>-1</sup>. When thiamine-pp is bound to a TDE, however, the pK<sub>a</sub> of the C<sub>2</sub>-H drops significantly and its proton exchange rate increases to about 314,000 ms<sup>-1</sup> (Walsh & Tang, 2019). Therefore, the C<sub>2</sub> carbanion is readily available as a reactive species in TDEs. In the enzymatic context, the C<sub>2</sub> carbanion functions as a nucleophile reacting mostly with α-hydroxy ketones and α-keto acids. In the corresponding reactions, labile thiamine-ketone adducts are established through C-C bond formation between the coenzyme C<sub>2</sub>

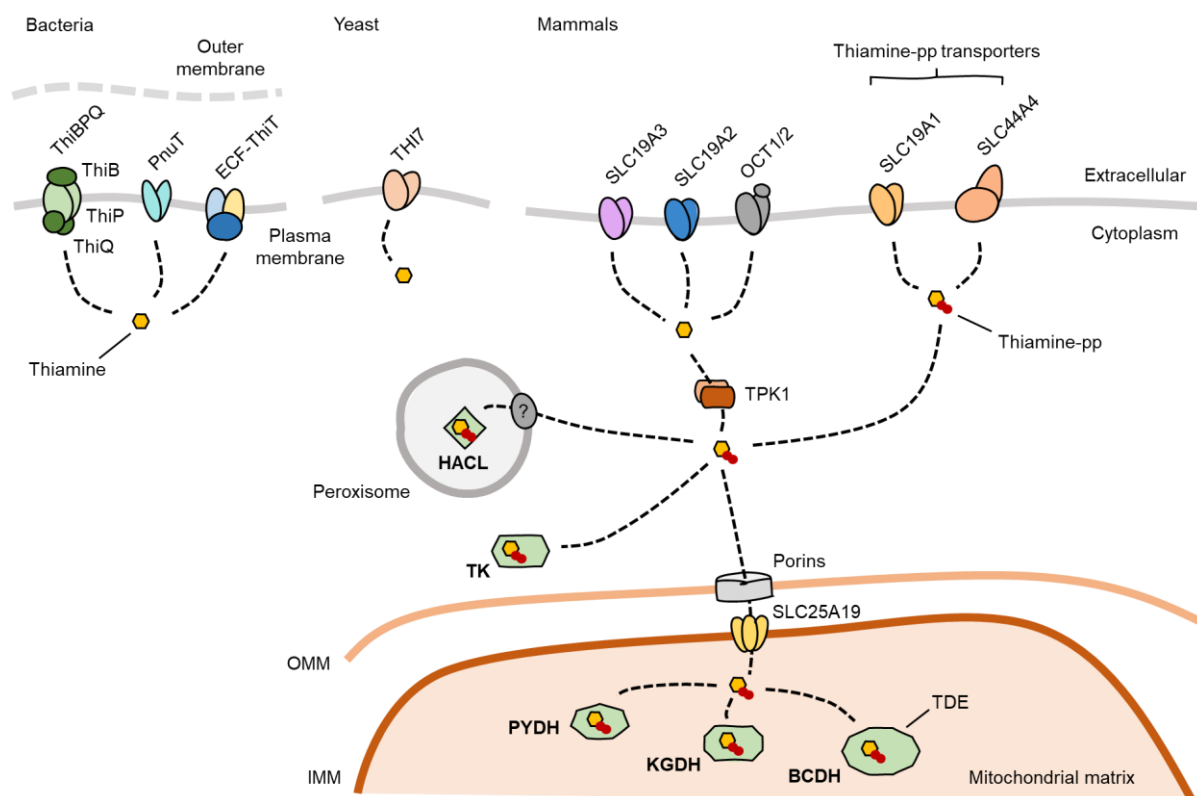
and the substrate carbonyl carbon. These adducts subsequently decay, typically under the elimination of one of the substituents of the substrate, and the return of the coenzyme thiamine-pp to its C<sub>2</sub> carbanion state (Fig. 2).



**Fig. 2: General reaction mechanism of thiamine-dependent enzymes.** The C<sub>2</sub> atom of the thiazole ring of thiamine can be deprotonated. The resulting carbanion serves as the reactive species in the active site of thiamine-dependent enzymes (TDEs). Most TDEs catalyse the chemical conversion of α-hydroxy ketones or α-keto acids. Figure designed in Marvin (Chemaxon).

The human proteome comprises five known TDEs (Fig. 3, Walsh & Tang, 2019). Three of them catalyse the decarboxylation of α-keto acids in mitochondria. These are pyruvate dehydrogenase (PYDH), α-ketoglutarate dehydrogenase (KGDH), and branched-chain amino acid dehydrogenase (BCDH) (Fig. 3). The thiamine-pp carbanion acts as a nucleophile, attacking at the carbonyl carbon of α-keto acids, and thereby establishing a C-C bond to the substrate. The resulting labile reaction intermediate decays releasing CO<sub>2</sub> (decarboxylation), while a stable hydroxyethyl carbanion (HEC) stays connected to thiamine-pp through a C-C bond. This reactive species serves then as a nucleophilic substrate in subsequent enzymatic steps. PYDH, KGDH, and BCDH are all part of E1-E2-E3 multi-enzyme complexes. The TDEs respectively constitute the E1 enzyme in these complexes, and hand their HEC product over to different E2 and E3 enzymes. These enzymes link the HEC molecule covalently to coenzyme A (CoA). The products of these enzyme complexes are important metabolites. The product of the PYDH complex is acetyl-CoA, the raw material for many metabolic pathways, but in particular for the citric acid cycle. The product of the KGDH complex is succinyl-CoA, a metabolite within the citric acid cycle. The products of the BCDH complex are acyl-CoA derivatives of the three branched chain amino acids (BCAAs) valine, leucine, and isoleucine, which are further degraded by β-oxidation. Transketolase (TK) is another type of TDE, which mediates the ketol transfer between aldose sugars in the cytoplasm. This is an important reaction step in the pentose phosphate pathway. This pathway is crucial for the generation of NADPH and ribose-5-phosphate as the sugar core for the synthesis of nucleotides. TK catalyses the transfer of 2-hydroxyethanal (HEA) from xylulose-5-phosphate to either

ribose-5-phosphate (C<sub>5</sub> sugar) or erythrose-4-phosphate (C<sub>4</sub> sugar). This generates sedoheptulose-7-phosphate (C<sub>7</sub> sugar) and fructose-6-phosphate (C<sub>6</sub> sugar), respectively, as well as glyceraldehyde-3-phosphate (C<sub>3</sub> sugar) in both cases. Fructose-6-phosphate and glyceraldehyde-3-phosphate are also metabolites of the glycolysis. Therefore, the role of TK is in essence to metabolically connect the pentose phosphate pathway and glycolysis. The fifth known human TDE, 2-hydroxyacyl-CoA-lyase (HACL), is important for the degradation of 2-hydroxy fatty acids in peroxisomes (Mori et al., 2023). 2-hydroxy fatty acids are degradation products of cerebrosides and other sphingolipids, which are particularly enriched in the brain (Foulon et al., 2005). The thiamine-pp carbanion acts as a nucleophile to attack – in this case – a thioester, forming a C-C bond with the carbonyl carbon of the 2-hydroxy-acyl-CoA substrates of HACL. The labile reaction intermediate decays and releases the n-1 acyl chain, which leaves the reaction as an aldehyde. This n-1 fatty aldehyde can subsequently be oxidised to form an n-1 fatty acid, which can then be degraded in standard  $\beta$ -oxidation. Meanwhile, methanol-CoA stays covalently connected to thiamine-pp in the form of an enamine adduct. Deprotonation of the methanol hydroxy group destabilises this enamine intermediate and leads to its dissociation into formyl-CoA and the thiamine-pp carbanion. Formyl-CoA can be repurposed by other cellular enzymes.



**Fig. 3: Overview of thiamine's membrane transport and metabolism in the cell.** Explanation of abbreviations: TDE (thiamine-dependent enzyme), OMM (outer mitochondrial membrane), IMM (inner mitochondrial membrane). For the individual transporters and enzymes, consult the main text.

## 1.2 Thiamine transport across cellular membranes

### Membrane transport of thiamine in bacteria

Many bacteria are able to synthesise thiamine *de novo*. As this biosynthesis is energetically costly, most bacteria still have dedicated molecular transport machineries for the uptake of thiamine from their environment (Jaehme et al., 2018). A common solution for thiamine import across the plasma membrane is the use of primary active transporters (Webb et al., 1998). One example for these is the ABC transporter ThiBPQ (Jurgenson et al., 2009). This transporter consists of three distinct subunits, namely ThiB (periplasmic binding protein), ThiP (transmembrane permease), and ThiQ (ATPase) (Webb et al., 1998, Jaehme et al., 2015). There are furthermore ECF-type primary active transport systems, which use the thiamine-binding S component ThiT for the translocation of the vitamin (Slotboom, 2014). Another set of bacterial thiamine transporters are PnuT carrier proteins (Jaehme et al., 2018). PnuTs show some structural resemblance of the SWEET transporter family, in that their functional core is formed by two symmetrically related 3-helix bundles (Jaehme et al., 2018). These are thought to move in a rocker-switch type manner to provide alternating access to the substrate binding site (Xu et al., 2014). The PnuT from *Shewanella woodyi* (PnuT<sub>sw</sub>) has been shown to operate in a facilitated diffusion transport mode (Jaehme et al., 2018). The vast evolutionary diversity of bacteria likely brought about many more membrane transport systems for thiamine. The discovery and characterisation of these is an ongoing scientific endeavour (Jaehme et al., 2015, Jaehme et al., 2018).

### Membrane transport of thiamine in yeast

In baker's yeast (*S. cerevisiae*), THI7 (formerly THI10, Uniprot ID: A7A1C0), was identified as a membrane transporter for thiamine (Enjo et al., 1997, Singleton, 1997). Interestingly, THI7 is inhibited by chloroquine (Huang et al., 2012). This is a feature it shares with its human counterpart SLC19A3 (Vora et al., 2020). Other quinoline-containing drugs inhibiting thiamine transport in yeast include amodiaquine, quinacrine, mefloquine, primaquine, quinine, and quinidine (Huang et al., 2012).

### Thiamine transport in humans

In humans and other mammals, thiamine transport across the cellular plasma membrane is mainly mediated by SLC19A2 and SLC19A3 (Fig. 3, Dutta et al., 1999, Rajgopal et al., 2001). They are members of the solute carrier protein family (SLC) and mediate the cellular uptake of thiamine with high affinity ( $K_m = 2-7 \mu\text{M}$ , Giacomini et al., 2017, Yamashiro et al., 2020). Due to their highly similar substrate spectrum and transport affinities, the two transporters are thought to be biochemically rather identical (Yamashiro et al., 2020). SLC19A2 and SLC19A3 are, however, differentially expressed in different tissues and cell types (Uhlén et al., 2015, Karlsson et al., 2021). SLC19A2 is relatively ubiquitously expressed but can be found in its highest quantities in skeletal muscles and for some

potentially interesting reason in the tongue (The Human Protein Atlas, Uhlén et al., 2015). SLC19A3 expression appears to be more restricted to specific tissues and cell types. It shows its strongest expression in the placenta and the breast, i.e. organs concerned with the nutritional supply of the foetus and new-borns (The Human Protein Atlas, Uhlén et al., 2015). SLC19A3, but not SLC19A2, is expressed in the endothelial cells of the blood-brain barrier (transcriptomics datasets: Human Multiple Cortical Areas SMART-seq, 2018, Human M1 10x, 2020, Human MTG 10x (SEA-AD); Allen Brain Map). This makes SLC19A3 a vital entry gate for thiamine into the brain (Kono et al., 2009, Suzuki et al., 2017, Giacomini et al., 2017, Aburezq et al., 2023). At the subcellular level, SLC19A3 was found to localise specifically to the apical membrane of polarised cells of the intestinal wall (Said et al., 2004). As a result of this differential expression and localisation, the two thiamine transporters are functionally non-redundant on a physiological level. This aspect becomes particularly conspicuous in patients carrying mutations in one of the transporter genes, which I elaborate on in more detail in Section 1.3. OCT1 (SLC22A1) and OCT2 (SLC22A2), which are known to be promiscuous transporters for organic cations can also in principle transport thiamine (Fig. 3, Chen et al., 2014, Kato et al., 2015). The affinity of this transport is, however, about two orders of magnitude lower, compared to SLC19A2 and SLC19A3 ( $K_m^{OCT1} \approx 780 \mu\text{M}$ , Chen et al., 2014). It is questionable whether OCT1 and OCT2 mediate thiamine transport under standard diet conditions. But they likely contribute to thiamine transport, when the vitamin is supplied in high doses (Kato et al., 2015, Aburezq et al., 2023). SLC35F3 is an intracellular thiamine transporter (Zhang et al., 2014). As a member of the SLC35 nucleotide sugar transporter family, it is localising to the endoplasmic reticulum (ER) and the Golgi apparatus (Ishida et al., 2004, Thul et al., 2017). It has not been thoroughly studied, neither on a biochemical nor a biological level. Some studies have found a genetic association of polymorphisms in SLC35F3 with hypertension (Zhang et al., 2014, Zang et al., 2016). These findings are, however, under debate (Seo & Choi., 2021). In humans, thiamine pyrophosphate (thiamine-pp) has its own distinct membrane transportome. Plasma membrane translocation of thiamine-pp can be mediated by the reduced folate carrier SLC19A1 (Fig. 3, Zhao & Goldman, 2008). This solute carrier operates as an antiporter for organic anions and is closely related to the thiamine transporters SLC19A2 and SLC19A3 (35% sequence identity, 48% sequence similarity, Wright et al., 2022, Zhang et al., 2022, Dang et al., 2022). Thiamine-pp can also be transported across the plasma membrane by SLC44A4, a member of the choline-like transporter SLC44 family (Fig. 3, Nabokina et al., 2014). Within the cell, thiamine-pp has to be transported across organellar membranes as well. In particular, it has to be imported into the mitochondrial matrix, where the three thiamine-dependent enzymes (TDEs) PYDH, KGDH, and BCDH are localised. As a small solute, thiamine-pp can likely diffuse through porins of the outer mitochondrial membrane (Kühlbrandt, 2015). The transport across the inner mitochondrial membrane, however, needs to be facilitated by the specific thiamine-pp transporter SLC25A19 (Fig. 3, Bottega et al., 2019, Ruprecht & Kunji, 2020). SLC19A1-3 are MFS transporters, following the canonical MFS fold on a structural level (Quistgaard et al., 2016). SLC35F3 and SLC44A4 are structurally distinct and fall into solute carrier families that

have not been thoroughly studied in terms of their structures and transport mechanisms. SLC25A19 is part of the SLC25 family of mitochondrial carrier proteins, which are 6-pass transmembrane proteins sharing a characteristic fold and presumably also transport mechanisms (Ruprecht & Kunji, 2020). The mammalian thiamine transporters have no obvious homology to bacterial or yeast thiamine transporters. This appears to represent an example of convergent evolution, with membrane proteins of different genetic origin having evolved to fulfil the function of thiamine transport.

### **1.3 Thiamine deficiencies in humans**

In humans, thiamine deficiencies can have nutritional, drug-induced, or genetic causes.

#### **Nutritional thiamine deficiencies**

A nutritional thiamine deficiency can be caused by anything that interferes with the dietary intake of thiamine. As humans have only limited storage capacities for thiamine, a prolonged nutritional deficiency in the uptake of this vitamin can lead to organism-wide thiamine deficiencies. The recommended amount of dietary thiamine is 1.2 mg per person per day (NIH, 2024). Nutritional thiamine deficiencies can have different origins but are ordinarily caused by a thiamine-deficient diet. Other causes include inflammatory bowel diseases, like Crohn's disease (Oudman et al., 2021), excessive vomiting in the context of eating disorders or chemotherapy (Gomes et al., 2021, Smith et al., 2021), and alcoholism (Langlais, 1995). Nutritional thiamine deficiencies cause a group of disorders commonly known under the collective term Beriberi (Smith et al., 2021), as well as Wernicke's encephalopathy. Beriberi can present as Wet or Dry Beriberi. While Wet Beriberi is primarily characterised by defects of the cardiovascular system, patients with Dry Beriberi show primarily neurological symptoms. Wernicke's encephalopathy typically presents with ophthalmoplegia (weakness of the extraocular muscles), ataxia, and confusion (Smith et al., 2021). Wernicke's encephalopathy is most commonly observed in the form of the Wernicke-Korsakoff syndrome (WKS) in long-term alcoholics (Thomson & Marshall, 2006). In addition to the symptoms of Wernicke's encephalopathy described above, patients with WKS present with neurocognitive disorders including amnesia and confabulation (Butterworth, 1989, Rensen et al., 2021). The causal connection between alcohol abuse and thiamine deficiency can be largely explained by poor eating habits paired with an alcohol-induced damaging of the intestinal wall (Langlais, 1995).

#### **Drug-induced thiamine deficiencies**

Medication-induced development of Wernicke's encephalopathy is known in the case of the Janus-kinase (JAK) inhibitor fedratinib. In a phase III clinical trial testing fedratinib for the treatment of myelofibrosis (NCT01437787), several patients presented with Wernicke's encephalopathy. Some of these patients died presumably in connection to this neuropathological presentation (Pardanani et al.,



2015). Trying to find the underlying cause of this drug-induced phenotype, researchers found that fedratinib is inhibiting thiamine transport through SLC19A3 with high affinity (Zhang et al., 2014).

### **Genetic thiamine deficiencies**

The synthesis and membrane transport of thiamine is dependent on the concerted activity of many proteins. Loss-of-function-mutations in any of these can lead to thiamine deficiencies, often with fatal consequences for the organism. In humans, at least four proteins are known to be essential in maintaining the necessary levels of thiamine and thiamine-pp throughout the body. SLC19A2 and SLC19A3 transport thiamine over cellular plasma membranes and consequently also across tissue barriers. The expression of the two carrier proteins is cell type specific. Therefore, loss-of-function mutations of either SLC19A2 or SLC19A3 have different effects on the body. Malfunction of SLC19A2 causes thiamine-responsive megaloblastic anaemia (TRMA), which is characterised by sensorineural hearing loss and diabetes mellitus (Labay et al., 1999, Diaz et al., 1999, Lagarde et al., 2004, Liberman et al., 2006). Loss-of-function mutations in SLC19A3 cause biotin- and thiamine-responsive basal ganglia disease (BTBGD) and Wernicke's like encephalopathy (WLE), which present with severe neurological symptoms (Supplementary Fig. 18, Ozand et al., 1998, Zeng et al., 2005, Subramanian et al., 2006, Kono et al., 2009, Schänzer et al., 2014, Alfadhel et al., 2019, Wang et al., 2021, Aburezq et al., 2023). TPK1 is the enzyme catalysing conversion of thiamine into its coenzyme form thiamine-pp. Mutations in the *TPK1* gene cause thiamine metabolism dysfunction syndrome 5 (THMD5), which presents with acute encephalopathic episodes, which progressively develop into severe neurologic dysfunctions (Mayr et al., 2011, Huang et al., 2019). SLC25A19 is mediating the import of thiamine-pp into the mitochondrial matrix. Defects in this transporter can cause thiamine metabolism dysfunction syndrome 4 (THMD4) or Amish lethal microcephaly, a severe neurodevelopmental disorder. In both cases, the central nervous system is heavily affected (Rosenberg et al., 2002, Spiegel et al., 2009, Gowda et al., 2019, Bottega et al., 2019, Chen et al., 2021). Inheritance of these genetic thiamine deficiency syndromes is autosomal recessive (Labay et al., 1999, Aburezq et al., 2023, Mayr et al., 2011, Chen et al., 2021). This implies that one functional copy of the respective genes is sufficient to provide thiamine and thiamine-pp at the necessary concentrations throughout the body. However, heterozygous carriers of the mutations might have a higher susceptibility to nutritional or drug-induced thiamine deficiencies.

## **1.4 Cryogenic electron microscopy of small proteins**

Traditionally, X-ray crystallography has been the technique of choice for determining high-resolution structures of biological macromolecules. Over the past decade, however, single particle cryogenic electron microscopy (cryo-EM) has become an important technique for structure determination. This development was enabled by major technical and computational innovations. Nowadays, cryo-EM is competitive with X-ray crystallography in terms of achievable resolutions, and superior when it comes

to challenging targets, such as membrane proteins or large macromolecular complexes. Membrane proteins make particularly difficult samples for X-ray crystallography, as they can rarely be purified in sufficient quantities and are inherently difficult to crystallize. Cryo-EM requires only small amounts of protein sample and does not require crystallisation of the target. However, one clear limitation of cryo-EM is the size of the macromolecules that can be structurally resolved. Below a certain molecular mass, biomolecules are not accessible to structure determination by cryo-EM. The reason for this is primarily that high-resolution reconstructions strictly require distinguishable geometry in the particles images in order to unambiguously assign and refine their tomographic poses at high angular resolutions (Fig. 4, Lander & Glaeser, 2021, Wentinck et al., 2022). In the beginnings of the so-called resolution revolution of cryo-EM around the year 2014, the practical lower size limit was at about 100 kDa (Kühlbrandt, 2014, Lander & Glaeser, 2021). Over the past ten years, optimisation of sample preparation, microscope instrumentation, and software development have enabled structure determination of ever smaller targets (Kimanius et al., 2024). The size of proteins that can be effectively resolved at high resolution lies currently, as I will also demonstrate in this thesis, well within the 40-60 kDa range. This makes a large fraction of the proteome accessible for structure determination by cryo-EM (Wentinck et al., 2022, Nevers et al., 2023, Kimanius et al., 2024). In the following, I will provide a concise overview of the sample preparation for cryo-EM, the physical basis of image formation for biomolecular specimens in an electron microscope, and the particular challenges small targets pose for data processing.

### **Sample preparation for cryo-EM**

For single particle cryo-EM, the macromolecular target of interest is ideally produced at high purity. The target is then concentrated to commonly between 1-10 mg/mL. This concentrated sample is subsequently applied to small metal grids carrying a thin holey film made of carbon or gold. After application of the sample, the grid is blotted with absorptive paper. This leaves an ideally 30-500 nm thick film of the aqueous sample behind (Naydenova et al., 2020, Neselu et al., 2023). The grid is then plunged at high speed into a liquid ethane/propane mixture, which is kept below -150 °C. At these temperatures, the thin sample film cools so rapidly that it is fixed in a glass-like, vitreous, state. If the vitrified sample is thin enough, it is sufficiently transparent for electrons at the energies used in cryo-EM (200-300 keV) (Martynowycz et al., 2021). Importantly, the vitrification prevents the formation of ice crystals, which can destroy the structural integrity of biomolecules and interfere heavily with image formation (Dubochet et al., 1988). The cryogenic temperatures furthermore prevent molecular motion and significantly reduce radiation damage during imaging in the electron microscope (Naydenova et al., 2020).

### **Image formation in cryo-EM**

The technical premise of single particle cryo-EM is to image biological macromolecules at close to atomic resolutions, i.e. somewhere between 1-3 Å. For classical light microscopes, which use photons

for imaging, this resolution range is not accessible. The maximal resolution of an optical system is ultimately capped by the diffraction limit of the particles used for imaging. This diffraction limit is approximated by the following equation (Eq. 1), developed by Ernst Karl Abbe:

Eq. 1:

$$d = \frac{\lambda}{2 * n * \sin\theta}$$

$d$  is the size of the smallest resolvable distance between two features,  $\lambda$  is the wavelength of the particles used for imaging,  $n$  is the refractive index of the medium of the specimen, and  $\theta$  the aperture half-angle of the imaging lens.

In the case of classical light microscopy, the photons have wavelengths from 400-700 nm, and cannot surpass a diffraction limit of about 200 nm. A theoretical solution to this problem would be to use photons with smaller wavelengths. For the atomic resolution range, these would be hard X-rays with wavelengths of down to 60 pm and corresponding photon energies up of several keV (e.g. Beamline P14, PETRA III, DESY, Hamburg, Germany). The problem here is, however, that there is currently no sufficiently refractive material for these high-energy X-rays, which makes it impossible to build proper lenses for atomic resolutions (Rösner et al., 2020). The only way to do photon-based structure determination today is therefore X-ray crystallography. In transmission electron microscopy (TEM), we shift the diffraction limit by using electrons instead of photons. The wavelength of electrons can be easily tuned to subatomic dimensions by increasing their energy. This relationship is defined in the de Broglie wavelength of an electron (Eq. 2):

Eq. 2:

$$\lambda_e = \frac{h}{p_e}$$

With  $\lambda_e$  being the de Broglie wavelength,  $h$  the Planck constant ( $\approx 6.626 \times 10^{-34}$  J-s), and  $p_e$  the momentum of the electron, which is approximately proportional to the used acceleration voltage.

At energies of 200 keV and 300 keV conventionally used for cryo-EM, the effective de Broglie wavelength of the electrons is about 2 pm and 1 pm, respectively (Kohl & Reimer, 2008). This shifts the diffraction limit to subatomic dimensions, wherefore it is not a resolution limiting factor anymore. At the same time, the movement of electrons can be controlled by magnetic fields. This allows for the construction of effective and tuneable electromagnetic lenses, which can be used to build complex optical systems for electron microscopes.

While electrons are excellent in terms of their high-resolution optical properties, their interaction with biological molecules provides hardly any direct contrast. Biological molecules are largely made of the light elements hydrogen, carbon, nitrogen, oxygen, and sulphur and are embedded in vitreous H<sub>2</sub>O for

cryo-EM. Electrons interacting with light atoms are mostly elastically scattered. This means that they keep their energy through the scattering event, and just experience a phase shift relative to the unscattered electron wave front (Kohl & Reimer, 2008). Elastically scattered electrons consequently reach the detector at their fully intensity, which is why cryo-EM samples are almost contrast-free when imaged in focus. One way to chemically introduce contrast, is to embed biological molecules in a solution of heavy atoms. A frequently used reagent for this so-called negative-stain electron microscopy (ns-EM) is uranyl ( $^{238}\text{U}$ ) acetate. Electrons passing through heavy atoms are typically inelastically scattered, which means that they deposit part of their energy in the specimen. In ns-EM, electrons experience therefore a clear difference between the biological molecules (elastic scattering) and the surrounding staining reagent (inelastic scattering), which translates into sharp contrast on the electron detector. Heavy metal salts can, however, denature biological samples. Furthermore, the maximum resolution of this approach is limited to about 10-20 Å by the grain size of the staining reagent (Gallagher et al., 2019).

In order to access high-resolution information, we can instead make use of phase contrast in unstained specimens. To generate this type of contrast, the phase difference between elastically scattered electrons and the unscattered electron wave fronts needs to be converted into measurable phase contrast. This is most commonly achieved by recording images not directly in the back focal plane of the objective lens, but slightly before that, at a defined defocus  $\Delta z$  (Kohl & Reimer, 2008). This brings scattered and unscattered electron waves in interference conditions and translates the original difference in their phases into an intensity difference, which can be measured on an electron detector. When generating phase contrast through defocussing, information is not homogenously transferred from the specimen to the detector. The measurable phase contrast is instead transferred at defined spatial frequencies, at which the electron waves experience constructive or destructive interference. This is expressed in the phase contrast transfer function, short CTF (Eq. 3, Kohl & Reimer, 2008).

Eq. 3:

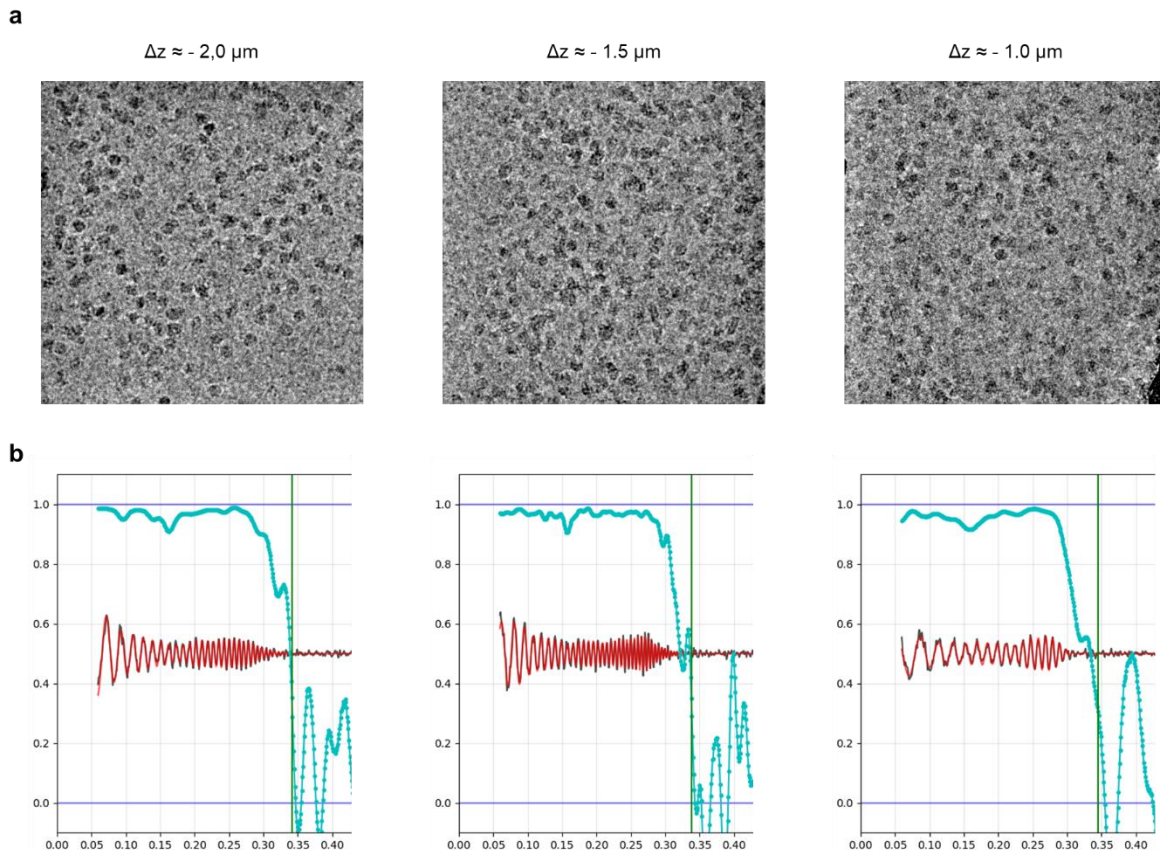
$$CTF(s) = -2 * \sin\left(\frac{\pi}{2} * (C_s * \lambda^3 * s^4 + 2 * \Delta z * \lambda * s^2)\right)$$

$s$  is the spatial frequency on the detector,  $C_s$  the spherical aberration of the objective lens,  $\lambda$  the de Broglie wavelength of the electrons, and  $\Delta z$  the applied defocus.

The CTF is a sinusoidal function, which means that it passes periodically through zero. At spatial frequencies, at which the CTF is zero or close to zero, information from the specimen is effectively lost in the recorded images (Fig. 4). For single particle cryo-EM, which uses the information from thousands to millions of individually images, the easiest way to circumnavigate this problem is to slightly alter the defocus  $\Delta z$  from exposure to exposure. Using this approach, the information-containing maxima and minima of CTF are shifted to different spatial frequencies in every image (Fig. 4). When aggregating

many EM images in 2D or 3D alignments, the entire accessible spatial frequency range is thereby eventually filled with contrast information from the specimen.

An important aspect to be aware of in cryo-EM is radiation damage. When recording images, inelastic electron scattering events deposit energy in the specimen, which can induce secondary chemical reactions. Typical examples of these are the radiolysis of carboxyl moieties, disulphide bonds, and metal centres (Holton, 2009, Garman, 2010, Naydenova et al., 2020). This problem can be treated on computational level, as described in the Section below.



**Fig. 4: Cryo-EM micrographs and their CTFs at different defoci.** **a** Micrographs of hSLC19A3 in complex with Nb3.3 and hydroxychloroquine. Individual exposures were recorded at a defocus of  $-2.0 \mu\text{m}$ ,  $-1.5 \mu\text{m}$ , and  $-1.0 \mu\text{m}$ , respectively. **b** Correlation of the radially averaged power spectrum (black curve) and the fitted CTF function (red curve) of the individual micrographs. The correlation coefficient (y-axis) as a function of spatial frequency (x-axis, in  $\text{\AA}^{-1}$ ) is shown as a turquoise curve.

## Data processing in cryo-EM

Nowadays, cryo-EM data processing is performed using mainly two different software packages: the freely available RELION, developed by Sjors Scheres and colleagues (Scheres, 2012), and the commercially available cryoSPARC (Structura Biotechnology, Punjani et al., 2017). For this thesis I used almost exclusively cryoSPARC as it provides significant advantages in terms of processing speed and project management. The theoretical principles described in the following are thus inspired by workflows in cryoSPARC but are also generally true for RELION.

One of the major technological advances in cryo-EM was the development of direct electron detectors (DED). These high-end cameras are able to detect individual electrons with a high spatial and, very importantly, also high temporal resolution. This enables the recording of movies, rather than static images. The movie format has the great advantage that beam-induced motion of the specimen can be computationally traced and corrected during data processing. As movement by even a few nanometres leads to a loss of high-resolution information in simply averaged images, computational motion correction of movies is a prerequisite for high-resolution structure determination (Scheres, 2014). Another important advantage of the movie format is dose weighting. Dose weighting has been developed to account for radiation damage accumulating over the duration of the exposure of the specimen to the electron beam (Zheng et al., 2017, Zivanov et al., 2019). Radiation damage first affects high-resolution features (e.g. decarboxylation of aspartate and glutamate side chains, reduction of disulphide bonds). Therefore, high-resolution information is only taken into full account from the first frames of the movies and is mathematically quenched in later frames. Thereby, the later frames, in which the target macromolecule has already experienced a high electron dose, are set to contribute only low-resolution information (Naydenova et al., 2020, Yip et al., 2020). Once the movies are motion corrected and dose weighted, they can be treated as images, which are from now on referred to as micrographs. For each micrograph, its experimental CTF can be estimated by fitting the theoretical CTF equation to the radially averaged power spectrum of the micrograph (Eq. 3). Important parameters of an individual micrograph can be inferred from the fitted CTF curve. This includes its defocus, CTF resolution and the relative thickness of the vitreous ice in the imaged area (Punjani et al., 2017). This information can be used to exclude low-quality micrographs from further data processing. Motion-corrected and CTF-estimated micrographs can now be used for single particle analysis. In a first step, target particles need to be identified in the micrographs. This can be done by manual particle picking. Nowadays, this is usually automated using template- or neural network-based picking approaches. For particles that adopt a roughly globular shape, template-based picking using a spherical or elliptical “blob” as template is usually sufficient. Once initial particle coordinates are identified in the micrographs, particle images can be extracted. This entails copying and saving a square image of a defined number of pixels (defined as box size) around the respective particle coordinates for further processing. The box size needs to be adjusted to cover all relevant spatial information of the particle. Importantly, this does not only comprise the clearly visible, high-contrast, low-frequency information that a human observer would identify as a target particle. When movies are recorded in defocus, high-resolution information is delocalised further away from the particle centre (Glaeser & Downing, 2007). Therefore, a box size of 2-3 times the physical particle diameter should be chosen to capture most of the valuable high-resolution information. After extraction, the particle images can be aligned in a process called 2D classification (Frank, 2009). The low signal-to-noise ratio (SNR) of the individual images is compensated when many particles of similar tomographic views can be aligned and averaged with a certain translational and angular precision. If the alignment works, medium-resolution features of the target molecule, like secondary structure

elements, can become visible. This process provides the basis to manually select 2D classes representing projections of the intact target molecule and to discard so-called junk particles. Junk particles can be everything from disintegrated macromolecules, particles located on the carbon film of the grid, empty detergent micelles in the case of membrane proteins, to randomly picked background noise. Usually, several rounds of 2D classification and 2D class selection are performed iteratively to successively clear the particle stacks from junk particles (Scheres, 2010). Once the particle image stacks are cleaned-up to a satisfactory degree in 2D, the remaining “good” particles can be used to reconstruct a three-dimensional density map of the target molecule. It is very important to note, that all of the 3D operations described in the following are performed in Fourier space, which enables the application of the Projection-Slice Theorem (Nogales & Scheres, 2015). To create a first low-resolution volume from the selected particle images, stochastic gradient descent (SGD) algorithms are commonly used. They calculate a 3D volume likely explaining the observed particle images (Punjani et al., 2017). This so-called ab-initio reconstruction also provides a crude tomographic pose assignment for the individual particles. The underlying SGD algorithm is, however, only sensibly interpretable for resolutions  $> 5 \text{ \AA}$  (Punjani et al., 2017). To achieve higher resolution reconstructions, translational and angular refinement of the particle poses needs to be performed in distinct refinement programs. In simple terms, these programs iteratively optimise the particle poses relative to a provided reference volume, which is then itself iteratively refined by the back projection of the pose-refined particles. For macromolecules containing both highly ordered and flexible parts, non-uniform refinement (NU refinement) is able to account for this spatial variability and improve achievable resolutions (Punjani et al., 2020). This is particularly important for integral membrane proteins, which are themselves mostly structured but surrounded by highly disordered lipid membranes or membrane mimetics. Once the global pose refinement has been successfully performed, local refinement allows one to further improve the resolution in selected parts of the density map. For this, the focus volume of interest is enclosed in a corresponding 3D mask in real space. All real space signal of the particles outside this mask is excluded. Pose refinement is, however, performed in Fourier space, where the entire available spatial frequency range is used.

### **The particle size limit of cryo-EM**

The particle size limit of single particle cryo-EM becomes apparent at different stages of particle alignment, i.e. 2D classification, ab-initio reconstruction, and 3D refinement. Many small particles may still align well in 2D and even give seemingly accurate solutions in ab-initio reconstructions. However, the ensuing 3D refinement often fails, as the individual particle images do not contain enough structured mass to allow for an unambiguous pose assignment. For membrane proteins, this is further complicated by the obligate membrane mimetics surrounding their transmembrane domains. These mimetics, which can be detergent micelles, lipid nanodiscs, or even entire liposomes, add a high level of background noise to the particles. For solute carriers in particular, their inherent pseudo-C2-symmetry (Drew et al.,

2021) causes another challenge for pose refinement. In the low SNR situation of cryo-EM, it is close to impossible to unambiguously assign a tomographic viewing direction to a small pseudo-C2-symmetric particle, as each pose is almost identical to its 180° rotation (Wentinck et al., 2022). This leads to an effective averaging of the pseudo-mirror images of the particles. Technical advances on the microscope and software level have allowed to overcome these obstacles for some small targets (Yuan et al., 2022). However, the size range of 40-60 kDa, representing a great part of the proteome, is still difficult to access (Wentinck et al., 2022). Target molecules with less than 40 kDa of structured mass are likely to remain inaccessible for high-resolution structure determination by cryo-EM (Lander & Glaeser, 2021, Kimanius et al., 2024). For these cases, the particle size needs to be increased on a biochemical level. Corresponding approaches are described in the following Section 1.5.

## **1.5 Structural fiducials for small protein targets**

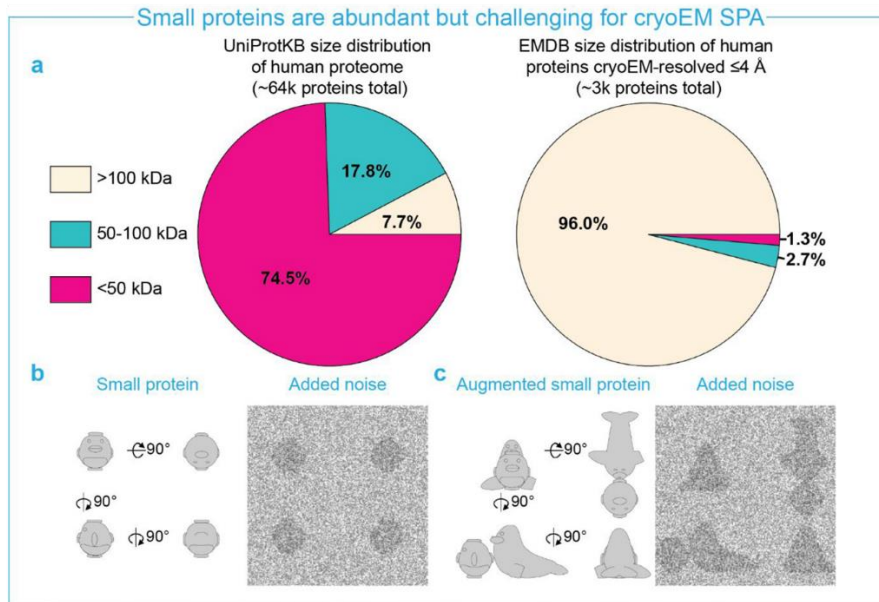
### **Overview of structural fiducials for cryo-EM**

Structure determination of small targets by cryo-EM is difficult due to the low signal they provide for particle alignment. The current lower limit of structured mass that can be resolved by cryo-EM is at  $\approx$  40-50 kDa for proteins (e.g. Yuan et al., 2022, Kimanius et al., 2024) and nucleic acids (e.g. Zhang et al., 2019, Ma et al., 2022). The median length of human proteins is about 375 amino acids in humans, and about 270 amino acids in bacteria (Brocchieri & Karlin, 2005, Nevers et al., 2023). This corresponds to molecular weights of approximately 41 kDa and 30 kDa, respectively (assuming an average molecular weight of amino acids of 110 Da). This renders a large fraction of the proteome inaccessible for cryo-EM (Fig. 5). A solution to this problem was found in the use of structural fiducials. These are well-structured biomolecules, which are rigidly linked to the target of interest. They add structured mass to the target and shift the resulting particles to a cryo-EM-compatible size range (Fig 5). Over the past decade, a plethora of such fiducials has been developed for the use in cryo-EM (Fig. 5). In the following, I will give a brief overview of structural fiducials for small protein targets. One strategy to increase the effective size of a protein of interest is to genetically fuse it with well-structured protein domains. Successful cases of this approach feature the bacterial cytochrome b562, also known as BRIL domain (Mukherjee et al., 2020), the NanoBit strategy (Sun et al., 2020), as well as the combination of maltose binding protein (MBP) and an MBP-binding DARPIn (Pidathala et al., 2023). Another widely used strategy to provide a small target with structural fiducials is to develop well-structured modules that bind rigidly and with high affinity to the protein of interest. The most dominant class of fiducials in this category are antibody fragments. They can be derived from normal immunoglobulins in the form of Fab fragments (Sakuragi et al., 2021, Ciută et al., 2023) or scFv fragments (Zhou et al., 2019, Oishi et al., 2023). Alternatively, the variable fragments of heavy-chain antibodies (HCAbs) from camelids or sharks can be used. These so-called nanobodies have enabled the structure determination of many small

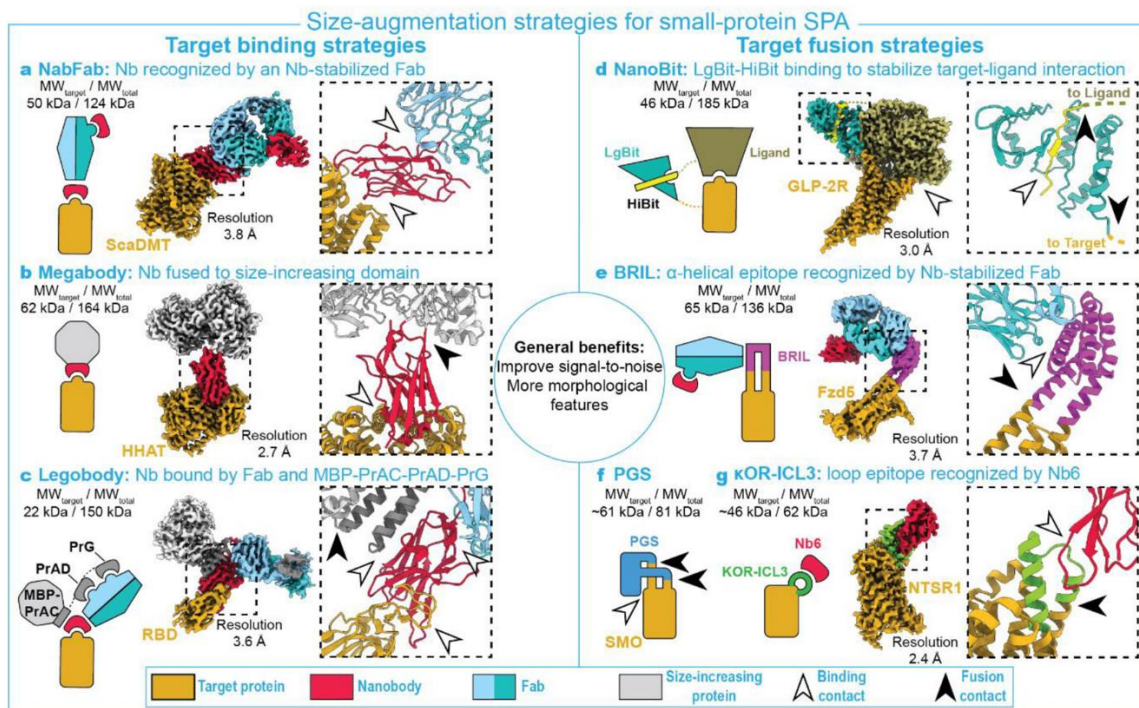


proteins by cryo-EM (Pardon et al., 2014, Hillier et al., 2024). Another type of fiducial are Designed Ankyrin Repeat Proteins (DARPin), which are, however, rarely used on their own in cryo-EM (Rothenberger et al., 2022). There are also nucleic acid-based fiducials in the form of aptamers. So far, their success has been limited in terms of the number of resolved structures and the achieved resolutions (Sampedro Vallina et al., 2023, Rahman et al., 2023).

a



b



**Fig. 5: Structural fiducials in single particle cryo-EM. a** Illustration of the size limitation problem in cryo-EM. **b** Overview of currently used structural fiducials for cryo-EM of small targets. The abbreviations are explained in the main text. This figure was adapted from Wentinck et al., 2022.

These strategies for the generation of structural fiducials are often used in elaborate combinations (Fig. 5), e.g. in the form of BRIL-Fab complexes (Mukherjee et al., 2020), KOR-ICL3:Nb6 complexes (Che et al., 2020), Legobodies (Wu et al., 2021), NabFabs (Bloch et al., 2021), Megabodies (Uchański et al., 2021), and molecular cages (Liu et al., 2019, Castells-Graells et al., 2023). For the work on the human thiamine transporters, I pursued three distinct strategies to provide these small membrane protein targets with useful fiducials for cryo-EM. I made use of the discovery of nanobodies and sybodies, as well as the BRIL fusion approach. In the following, I will elaborate on the origin and application of nanobodies and sybodies. The BRIL fusion approach will be described in more detail in the corresponding results Section 3.2.

## Nanobodies

Nanobodies are the isolated variable fragments (VHH) of heavy chain antibodies (HCAb), which are found in camelids and sharks (Muyldermans, 2013). HCABs were first identified in camels (*Camelus dromedarius*) by Hamers-Casterman and colleagues at the Vrije Universiteit Brussel in 1993. Two years later, in 1995, Greenberg and co-workers described another type of HCAb in nurse sharks (*Ginglymostoma cirratum*), which they termed nurse shark antigen receptor (IgNAR). Camelid and shark HCABs have evolved separately. In both cases, however, the antigen binding VHH domain has an immunoglobulin (Ig) fold, which consists of a 2-layer  $\beta$ -sandwich of 9  $\beta$ -strands in the case of camelid HCABs (Muyldermans, 2013), and 8  $\beta$ -strands in IgNARs (Stanfield et al., 2004). The stable  $\beta$ -sandwich serves as a scaffold for three highly variable loop regions, known as complementarity determining regions (CDR1,2,3). The CDRs mediate antigen binding and undergo affinity maturation by somatic hypermutation upon repeated exposure to the corresponding antigen (Teng & Papavasiliou, 2007). VHH fragments can be easily produced devoid of the HCAb constant domains. As single-domain antibodies, or in this case nanobodies, they are usually highly stable and can be purified in high quantities from overexpression in *E. coli*. Nanobodies have a smaller paratope compared to normal antibodies. This allows them to insert into otherwise hardly accessible cavities of antigens. Due to these features, nanobodies are widely used as research tools and are gaining popularity as biological therapeutics (Muyldermans, 2013, Fridy et al., 2024). In structural biology, nanobodies have become indispensable tools as crystallisation chaperones for X-ray crystallography and as structural fiducials for cryo-EM (Pardon et al., 2014). In these capacities, nanobodies have been particularly impactful for the work with small membrane proteins (Gelová et al., 2024, Hillier et al., 2024). Stably bound nanobodies provide necessary crystal contacts, as well as additional molecular mass for these targets. For my work on the human thiamine transporters, I decided to pursue nanobody discovery in llamas (*Lama glama*, part of the *Camelid* family). In the following, I provide a brief overview of the discovery process, which is described in more technical detail in the methods Section 2.5. To induce the proliferation and maturation of target-specific HCABs in a llama, the animal is immunised with the purified target molecule in conjunction with a suitable adjuvant to evoke a strong immune reaction (Pardon et al., 2014). The

immunisation is usually repeated six times on a weekly basis. This ideally induces a strong B cell response and affinity maturation of the antibody repertoire. Once the immunisation process is complete, B cells can be isolated from a blood sample of the llama. Subsequently, the VHH repertoire of the animal can be cloned from the B cell mRNA into suitable phage display vectors. Through phage display, nanobodies binding to the target antigen can be selectively enriched. After usually 1-3 rounds of phage display selection, and if the enrichment of specific binders is sufficient, individual nanobodies can be expressed and analysed using ELISA and DNA sequencing. Once target binding of the selected nanobodies is confirmed, they can be further characterised using biophysical techniques such as biolayer interferometry (BLI) and thermal stabilisation assays (e.g. nanoDSF).

## **Sybodies**

I also worked on the selection of sybodies as binders of the human thiamine transporters. Sybodies essentially build on existing nanobody scaffolds but have their CDRs synthetically randomised to generate a pool of more than  $10^{12}$  different sybody molecules. This synthetic pool of sybodies can be subjected to phage display selections, analogously to the procedure described above. The sybody platform was originally developed by the research group of Markus Seeger at the University of Zurich and was published in two seminal papers in the years 2018 and 2020 (Zimmermann et al., 2018, Zimmermann et al., 2020). Since then, this technology has been successfully used to select sybodies against a broad spectrum of targets. They have also proven to be useful fiducials for structure determination of membrane proteins by cryo-EM (Shahsavari et al., 2021, Deneka et al., 2021, Trinco et al., 2021, Bärländ et al., 2022, Botte et al., 2022).

## **Aim of the Study**

The main objective of this study was to determine high-resolution structures of the human thiamine transporters SLC19A2 and SLC19A3 by cryo-EM. It was known from the beginning that these transporters on their own would likely be too small for single particle cryo-EM. Therefore, a key element to achieving useful EM reconstructions was to develop suitable structural fiducials for the transport proteins. This should be attempted by pursuing fusion approaches, as well as the selection of suitable antibodies and nanobodies against hSLC19A2 and hSLC19A3. The overall aim of these efforts was to answer the following major questions:

1. What are the three-dimensional atomic structures of the human thiamine transporters?
2. Can different conformational states of the solute carriers be resolved?
3. How is the substrate thiamine recognised by hSLC19A2 and hSLC19A3?
4. How do known thiamine uptake inhibitors interact with the transporters?

To answer these questions comprehensively, the structural work should be complemented with biophysical techniques. Thermal shift and cellular uptake assays should give insights into the structure-function relationships of hSLC19A2 and hSLC19A3.

Based on the cryo-EM structures, medical compounds should be screened for their propensity to block thiamine transport in humans. For this purpose, computational docking should be performed. Hits from the virtual screening should subsequently be validated experimentally on a biophysical, functional, and structural level.

Another objective of the presented work was to attempt structure determination of the human enzyme TPK1 by cryo-EM. This part of the study aimed mainly at seeking a prove of principle. I wanted to test whether, and under which conditions, cryo-EM could be used to determine high-resolution structures of a small soluble protein like hTPK1.

## 2. Methods

### General remarks

- All H<sub>2</sub>O used for this work was de-ionised water purified with a Milli-Q® IQ 7000.
- Buffers for packed column-based purifications (in particular size-exclusion chromatography) were filtered (0.2 µm) and degassed for > 30 minute under a 100 mbar atmosphere prior to their use.
- Room temperature means 22 ± 2 °C.

### 2.1 Cloning of mammalian vitamin transporters

All membrane transporters were cloned, expressed, and purified largely following the procedure described in the following. Deviations from this standard protocol are detailed in the respective result Sections. cDNA encoding the respective proteins was obtained from SourceBioscience or GenScript. The plasmids containing the cDNA of the mammalian vitamin transporters (Table 5, Section 3.1) were initially amplified in and isolated from *E. coli* DH5α. Subsequently, the cDNA was selectively amplified by polymerase chain reaction (PCR) using the oligonucleotide primers listed in Supplementary Table 5, in a reaction mixture detailed in Table 1, and following a thermocycler program described in Table 2. The success of the PCR amplification was assessed by agarose gel electrophoresis. Clearly visible bands in the size range of the target sequence indicated a successful amplification.

**Table 1: PCR reaction mixture.** \*Green Phusion Mix (2 x concentration): 40 mM Tris-HCl, pH8.9, 4 mM MgCl<sub>2</sub>, 120 mM KCl, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.02 mM EDTA, 0.2% TritonX-100 (w/v), 8% glycerol (w/v), 0.005% Xylene Cyanol FF (w/v), 0.05% Orange G (w/v), 0.4 mM dNTPs, 0.04 U/µL Phusion polymerase.

Component	Stock concentration	Final concentration in 50 µL
Green Phusion Mix*	2 x	1 x
Template DNA	~ 300 ng	~ 3 ng
Forward primer	10 µM	0.5 µM
Reverse primer	10 µM	0.5 µM
H <sub>2</sub> O (nuclease-free)	N/A	To 50 µL

**Table 2: PCR thermocycler program.** A cycle comprises 30 seconds dissociation of the dsDNA template at 98 °C, 30 seconds annealing of the oligonucleotide primers at 60 °C, and 90 seconds synthesis of new DNA strands through the Phusion polymerase at 72 °C, assuming an elongation time of 30 seconds per 1,000 base pairs.

Step	Temperature	Duration (min:sec)
Initial denaturation	98 °C	00:30
Cycles (35 x)	98 °C	00:30
	60 °C	00:30
	72 °C	01:30
Final extension	72 °C	05:00
Hold	4 °C	∞

After PCR amplification, PCR products were cloned into pXLG-derived vectors using the SLiCE cloning approach (Zhang et al., 2014). For SLiCE cloning, 50 µL PCR product were initially supplied with 1 µL of the enzyme DpnI (20 U/µL) and incubated for 20 minutes at 37 °C to digest any methylated template DNA from the PCR reaction. Subsequently, the PCR product DNA was purified using the QIAquick PCR Purification Kit (Qiagen). The pXLG expression vectors were digested and linearised using suitable restriction enzymes (NotI-HF and EcoRV-HF for pXLG-C3S) in 1x CutSmart buffer for 2 hours at 37 °C. Afterwards, the linearised vectors were dephosphorylated using Antarctic Phosphatase (5 U) for 30 minutes at 37 °C, followed by an inactivation of the enzymes for 5 minutes at 75 °C. After the dephosphorylation, the vector DNA was separated on an agarose gel (1% agarose in TAE buffer (400 mM Tris-acetate, pH8.3, 10 mM EDTA), run for 1 hour at 120 V). The band corresponding to the linearised vector was excised and extracted using the QIAquick Gel Extraction Kit. For the insertion of the cDNA-derived PCR products into the pXLG expression vectors via SLiCE cloning, 100 ng linearized vector, 50 ng insert DNA (PCR product), and 1 µL 10 x PPY SLiCE extract were mixed in 1x T4 DNA ligase buffer in a total volume of 10 µL. The reaction mixture was subsequently incubated for 20 minutes at 37 °C. 3 µL of the SLiCE reaction mixture were transformed into 50 µL of chemically competent *E. coli* DH5α by a 45 seconds-long 42 °C heat shock. The cells were then plated on ampicillin-containing (100 µg/mL) LB agar. The plates were incubated overnight at 37 °C. Colonies were picked from these plates were analysed by colony PCR and used to inoculate 5 mL LB-ampicillin cultures (100 µg/mL). For colony PCR the picked colonies were briefly dipped into 1x Green Phusion Mix (see description Table 1), containing 0.4 µM of the sequencing primers CMV-for and WHV-5R, respectively. The colony PCR was performed using the thermocycler program detailed in Table 3. The presence of successfully inserted PCR product was subsequently assessed by agarose gel electrophoresis (1% agarose in TAE buffer, run for 1 hour at 120 V).

**Table 3: Colony PCR thermocycler program.**

Step	Temperature	Duration (min:sec)
Initial denaturation	95 °C	01:00
Cycles (35 x)	95 °C	00:30
	52 °C	00:30
	72 °C	01:30
Final extension	72 °C	05:00
Hold	4 °C	∞

The simultaneously inoculated 5 mL LB-ampicillin cultures were grown overnight in a shaking incubator at 37 °C. The grown bacteria were subsequently harvested by centrifugation (10 minutes, at 3,200 x g, at 10 °C). The cell pellets were aspirated and the plasmid DNA of successful clones (as judged by colony PCR) was isolated using the QIAprep Spin Miniprep Kit. The extracted plasmids were then sequenced using the CMV-for and the WHV-5R primers and the Sanger sequencing service provided by the company Microsynth AG. Point mutants and BRIL fusions of hSLC19A3 were created by combining two sets of forward and reverse primers, namely outer and inner primers, in the PCR amplification. For each construct, FG023-h19A3-for and FG066-h19A3-rev were used as outer primers, which anneal to the flanks of the hSLC19A3 insert, as well as the cloning site in the pXLG-C3S base vector. Point mutations were introduced by including inner forward (ifor) and inner reverse (irev) primers encoding the respective mutations along with overlapping sequences of at least 10 base pairs (see Supplementary Table 5). Five BRIL constructs, positioned in extracellular loops (ECL) 2, 3, 4, 5, or 6, were generated through two successive PCR reactions by Lea Spriestersbach, which is described in full detail in her master's thesis (Spriestersbach, 2022).

## 2.2 Expression and purification of mammalian vitamin transporters

Within the respective pXLG vectors, the target proteins were fused to an N- or C-terminal Twin-Strep tag, connected by an HRV 3C or TEV protease cleavage site and a flexible GSSG linker. Plasmid DNA was produced at large scale in *E. coli* (DH5 $\alpha$ ) and extracted and purified using the NucleoBond Xtra Maxi Kit or the NucleoBond PC 10000 EF Giga Kit (both Macherey-Nagel). For overexpression of the target proteins, the corresponding plasmids were transfected in log-phase Expi293F<sup>TM</sup> cells using PEI MAX<sup>®</sup> as transfection reagent. Transfection was performed by incubating Expi293F<sup>TM</sup> cells at a density of 20 million/mL with 30 ng/mL plasmid DNA and 60 ng/mL PEI MAX<sup>®</sup> for 45 min at 37 °C, 8% (v/v) CO<sub>2</sub>, and 270 rpm in FreeStyle<sup>TM</sup> medium. The cells were subsequently diluted to 2 million/mL with fresh FreeStyle<sup>TM</sup> medium and grown for 48-96 hours to allow for sufficient overexpression of the

transfected constructs. Afterwards, cells were harvested for 5 min at 3,000 x g and washed with 1 x PBS (pH7.4). Cell pellets were frozen and stored at -70 °C. The overexpressed membrane proteins were extracted using whole cell solubilisation with detergents. For this, 5 g of frozen cell pellet were thawed in a water bath at room temperature and subsequently resuspended in 25 mL of solubilisation buffer containing LMNG as main solubilising detergent (1x PBS, pH7.4, 200 mM NaCl, 5% glycerol (v/v), 1% LMNG, 0.1% CHS, 1 x EDTA-free protease inhibitors, 0.5 mM TCEP, 20 U/mL DNase I, 2.5 U avidin). The suspension was then incubated for 1 hour at 4 °C on a roller platform. Insoluble debris was subsequently pelleted by centrifugation for 30 min at 35,000xg at 4 °C. The cleared supernatant was loaded on 1 mL StrepTactin resin, equilibrated in StrepA buffer (20 mM Tris-HCl, pH 7.4, 350 mM NaCl, 5% glycerol (v/v), 0.5 mM TCEP, 0.02% LMNG (w/v), 0.002% CHS (w/v)). The resulting slurry was incubated for 1 hour at 4 °C to allow for quantitative binding of the Strep-tagged target proteins to the streptavidin immobilised on the resin. Afterwards, the slurry was loaded on a gravity column. The flowthrough was discarded, and the resin was washed 4x with 20 column volumes of StrepA buffer. The target protein was then eluted with 5 column volumes StrepB buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.002% LMNG, 0.0002% CHS, 10 mM desthiobiotin, pH adjusted to 7 with 1 M NaOH). The eluted protein was concentrated to 0.5 mL using a molecular weight cut-off of 50 kDa and was subsequently subjected to size-exclusion chromatography (SEC). SEC was performed in TBS-D buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.002% LMNG, 0.0002% CHS) on a Superdex S200 10/300 Increase column (Cytiva) using an ÄKTA pure system.

If a protein should be site-specifically biotinylated (primarily for sybody and nanobody discovery), an Avi-tag was inserted between the protein of interest and the respective protease cleavage site. For the biotinylation, StrepB-eluted protein was desalted using a PD Minitrap™ G-25 column with TBS-D, and supplied with 120 µg/mL of BirA, 5 mM ATP, 50 µM biotin, 0.5 mM TCEP and 40 µg/mL 3C protease. The reaction mixture was incubated overnight at 4 °C on a roller platform. Afterwards, the solution was concentrated using a 50 kDa molecular weight cut-off and subjected to SEC, as described above, to separate the added enzymes and small molecules from the biotinylated target protein. Protein samples were analysed by SDS-PAGE using SERVAGel™ TG PRiME™ 4-12 % precast gels in 1x Tris-glycine buffer at 250 V for 40 min. Protein bands were stained using Quick Coomassie Stain (SERVA).

### **2.3 Cloning, expression, and purification of the human TPK1**

Human thiamine pyrophosphokinase 1 (hTPK1) cDNA was obtained from GenScript and cloned into the pSb\_init vector for periplasmic expression in *E. coli*. In this vector, the wildtype hTPK1 was fused to an N-terminal PelB signal sequence for periplasmic secretion (cleaved post-translationally), and a C-terminal Myc-His<sub>6</sub>-tag, cleavable by HRV 3C protease. The cDNA was amplified by PCR using primers creating suitable overhangs for fragment exchange (FX) cloning (Supplementary Table 5, Geertsma,



2013). The PCR was performed with 1 ng of template cDNA and 0.5  $\mu$ M of the respective forward and reverse primers in 1 x Green Phusion Mix (see description Table 1), using the thermocycler program detailed in Table 4. Remaining methylated template DNA was digested with DpnI and the PCR product was purified using the QIAquick PCR Purification Kit.

**Table 4: PCR thermocycler program for the amplification of hTPK1.**

Step	Temperature	Duration (min:sec)
Initial denaturation	98 °C	00:30
Cycles (35 x)	98 °C	00:30
	65 °C	00:30
	72 °C	01:30
Final extension	72 °C	05:00
Hold	4 °C	$\infty$

For FX cloning, both the pSb\_init vector and the PCR product were digested and ligated in a single-tube, two step reaction with the type IIS restriction enzyme BspQ1. This enzyme cleaves DNA asymmetrically outside its non-palindromic recognition sequence (GCTCTTCN), to create 3-nucleotide 5' sticky end. These ends were designed to enable the complementarity-guided insertion of the hTPK1 DNA into the pSb\_init vector. For this DNA recombination, 300 ng of hTPK1 PCR product were mixed with 100 ng pSb\_init in 1 x NEBuffer 3.1 and supplied with 1  $\mu$ L BspQ1 in a total volume of 10  $\mu$ L. The reaction mixture was incubated for 1 hour at 37 °C. Subsequently, BspQ1 was inactivated for 20 minutes at 80 °C. For the ensuing ligation of the digested DNA fragments, the sample was provided with 1.25  $\mu$ L of 10 mM ATP and 1.25  $\mu$ L of T4 DNA ligase and incubated for 1 hour at room temperature. The DNA ligase was then heat-inactivated for 20 minutes at 65 °C. 5  $\mu$ L of the reaction mixture were subsequently transformed into 50  $\mu$ L of chemically competent *E. coli* (DH5 $\alpha$ ), which were then plated on LB-chloramphenicol (100  $\mu$ g/mL) agar plates. Colony PCR was performed using the pBAD-for and pBAD-rev primers (Supplementary Table 4) and analysed by agarose gel electrophoresis. Successful cloning was confirmed by Sanger sequencing. For expression, 0.5  $\mu$ L of hTPK1-pSb\_init plasmid were transformed in 150  $\mu$ L of chemically competent *E. coli* (MC1061). These cells were grown for 30 min at 37 °C in 20 mL antibiotic-free LB medium, which was subsequently supplied with 100  $\mu$ g/mL chloramphenicol. The cells were grown further overnight. This preculture was subsequently used to inoculate 6 L of TB-chloramphenicol (100  $\mu$ g/mL) medium (6 x respectively 1 mL preculture in 1 L medium in 2 L flask). The cells were grown at 37 °C, at 190 rpm to an OD<sub>600</sub> of 0.8. Overexpression of hTPK1 was induced by adding a final 0.02% of L-(+)-arabinose (w/v). To improve the yields of the recombinantly expressed thiamine kinase, 100  $\mu$ M of thiamine were simultaneously added to the growth medium. The cultures were grown for another 3 hours and harvested by centrifugation for 20 minutes at

4,500 rpm. The cell pellets were aspirated and frozen at -70 °C. For periplasmic extraction of the overexpressed hTPK1, 10-20 g of cell pellet were thawed in a 35 °C water bath and resuspended in 200 mL periplasmic extraction buffer (20% sucrose (w/v), 50 mM Tris-HCl, pH7.4, 0.5 mM EDTA, pH8.0, 1 x EDTA-free protease inhibitors, 0.5 µg/mL lysozyme, 0.5 mM TCEP). The suspension was stirred for 30 minutes at 4 °C. Afterwards, 400 mL of Mg buffer (20 mM Tris-HCl, pH7.4, 350 mM NaCl, 1 mM MgCl<sub>2</sub>, 0.5 mM TCEP) were added. Cell debris was subsequently pelleted by centrifugation. The soluble supernatant containing the protein of interest was transferred to a fresh bottle and supplied with 1 mL Ni-NTA resin, equilibrated in HisA buffer (20 mM Tris-HCl, pH7.4, 350 mM NaCl, 20 mM imidazole, pH7.4). After 1 hour of incubation on a stirring platform, the suspension was loaded on a gravity column. The flowthrough was discarded, and the resin washed with 4 x 20 mL of HisA buffer. The immobilised hTPK1 was eventually eluted with 5 mL of HisB buffer (20 mM Tris-HCl, pH7.4, 350 mM NaCl, 400 mM imidazole, pH7.4, 0.5 mM TCEP). The elution was then concentrated to 0.5 mL using a 10 kDa molecular weight cut-off concentrator and subjected to size-exclusion chromatography on a Superdex S200 10/300 Increase column on an AKTA Pure system in TBS buffer (20 mM Tris-HCl, pH7.4, 150 mM NaCl). Protein samples were analysed by SDS-PAGE as described above.

## 2.4 Sybody discovery

### Antigen preparation

The hSLC19A3 antigens for sybody discovery were designed to include a C-terminal Avi-tag, cloned in-frame with the transporter sequence and directly upstream of the HRV 3C protease cleavage site (cf. Supplementary Table 6). The posttranslational N-glycosylation of hSLC19A3 should be reduced as far as possible, in order to prevent masking of potential sybody-binding epitopes through the flexible oligosaccharides. For the first discovery campaign, hSLC19A3-wt-CA3S was therefore expressed in Expi293F<sup>TM</sup> GnTI- cells, lacking the N-acetylglucosaminyltransferase I. This enzyme is required for the build-up of complex N-linked glycan trees on secretory proteins. The transporter expressed in GnTI- was purified and subsequently treated with the enzyme EndoH, which cleaves the N-glycan trees down to their last N-acetylglucosamine unit (Freeze & Kranz, 2010). For later discovery campaigns, the fully glycosylation-free mutant hSLC19A3-gf-CA3S (Supplementary Table 6) was used, as it could be purified at > 10 x higher yields. In both cases, the antigens were biotinylated as described in Section 2.2, and treated with HRV 3C protease to remove their TwinStrep-tags.

### Sybody selections

A sybody discovery campaign consisted of several rounds of display-based selections of antigen-binding sybodies against the immobilised antigen. The sybody-encoding mRNA libraries (S, M, and L, Zimmermann et al., 2020), were obtained from the laboratory of Markus Seeger at the University of

Zurich. These libraries were used separately throughout the procedure. The mRNA libraries were used for the in vitro translation of the sybodies by ribosomes using the PUREfrex 2.1 kit and its DS supplement. Through the design of the mRNA sequences, translation was automatically stalled after the sybody polypeptide chain, whereby the phenotype (folded sybody) was physically connected to the respectively encoding genotype (mRNA). The resulting ribosome-mRNA-sybody complexes were then used to perform an initial ribosome display against the hSLC19A3-antigen in solution (solution panning). The antigens were then immobilised through their biotinylation on magnetic Dynabeads MyOne Streptavidin T1. The bound complexes were washed, subsequently destabilised by the chelation of  $Mg^{2+}$  with EDTA, and the sybody mRNA eluted with yeast RNA. Afterwards, the retrieved mRNAs were reversely transcribed using the Affinity Script reverse transcription kit, and the resulting cDNA amplified by PCR. The PCR product was then cloned by FX cloning (Geertsma, 2013, Zimmermann et al., 2020) into the pDX\_init vector, which fused the sybody sequences in-frame with the M13 phage tip protein pIII. The resulting phagemid libraries were then used to perform two to three successive rounds of phage displays against the hSLC19A3 antigens. For the production of sybody-carrying phages, the phagemid libraries were transformed into electrocompetent *E. coli* (SS320) through electroporation (1 mm-gap cuvettes, 1.8 kV). The transformed cells were then grown in 2YT medium, containing 200  $\mu\text{g/mL}$  ampicillin and 2% glucose (w/v) overnight at 37 °C. For the phage production, respectively 10 mL of these overnight cultures were infected with  $2.7 \times 10^{10}$  plaque-forming units of the M13K7 helper phage. The infected cells were centrifuged, aspirated, resuspended in 2YT medium, containing 200  $\mu\text{g/mL}$  ampicillin and 25  $\mu\text{g/mL}$  kanamycin, and grown overnight. The cells were then pelleted by centrifugation and the phage-containing supernatant transferred into a fresh tube. The phages were precipitated by mixing the supernatant with 0.2 volumes of a PEG6000/NaCl solution (20% (w/v), 2.5 M). The phages were then resuspended in 1 x PBS (pH7.4), cleared from remaining cell debris by centrifugation, and then used for the actual phage display-based selection. In the first round of phage display, the hSLC19A3 antigens were incubated with the phages (solution panning) and then immobilised on a neutravidin-coated Maxisorp plate. After extensive washing, bound phages were eluted by limited proteolysis using trypsin. Ensuing rounds of phage display were prepared in a similar fashion. The immobilisation of the antigen-phage complexes was, however, done on Dynabeads MyOne Streptavidin C1 to reduce the carry-over of substrate-binding phages. Antigen-specific enrichment of phages relative to a membrane protein negative control (chicken PHT1) was analysed by qPCR. After two to three rounds of phage display, the sybody libraries were subcloned into the pSb\_init vector using FX cloning. Individual sybody clones were then expressed (48 per library) in the periplasm of *E. coli* (MC1061). The periplasmic extracts from these small-scale expressions (cf. Section 2.3) were then analysed for the presence of antigen-specific sybodies in an ELISA format. For this, the sybodies contained in the periplasmic extracts were immobilised by their C-terminal Myc-tag on a protein A/anti-Myc antibody-coated Maxisorp plate. Subsequently, the plate was incubated with biotinylated antigen. After extensive washing, the presence of bound biotinylated antigen could be detected through the

enzymatic activity of a streptavidin-conjugated horse radish peroxidase (HRP). Promising sybody clones could then be expressed and purified at larger scale, analogously to the procedures described in Section 2.5. The sybody discovery protocol can be found in more detail in Zimmermann et al., 2020.

## 2.5 Nanobody discovery and production

Important note: For the procedures described in this Section, it was essential to ensure the production of high-fidelity VHH libraries and, of course, the well-being of the llamas. Therefore, the immunisations and ensuing steps up until the generation of VHH libraries were performed by experienced staff members of the Vrije Universiteit Brussel.

### Antigen preparation and immunisation of llamas

For the immunisation of llamas (*Lama glama*), hSLC19A3-gf-C3S and hSLC19A2-wt-C3S were expressed in Expi293F™ cells and purified as described in Section 2.2, including the cleavage of their TwinStrep-tags using HRV 3C protease. After size-exclusion chromatography, the proteins were transferred to storage buffer (20 mM Tris-HCl, pH7.4, 350 mM NaCl, 10% glycerol (w/v), 0.002% LMNG (w/v), 0.0002% CHS (w/v), 2 mM thiamine). Subsequently, the proteins were concentrated to 2 mg/mL ( $A_{280}$  of 3.8 and 4.4, respectively), aliquoted (8 x 50  $\mu$ L), and flash-frozen in liquid nitrogen. The antigen samples were subsequently stored at -70 °C. The two antigens, hSLC19A3-gf and hSLC19A2-wt, were respectively injected six times into a llama on a weekly basis. For the first, second, third and fifth injections, the antigens were mixed at a 1:1 ratio with the GERBU adjuvant (Pardon et al., 2014). For the fourth and sixth injection, the antigens and the adjuvant were administered spatially separated to avoid potential denaturation of native epitopes of the transporter by the adjuvant.

### Nanobody selections

After the completion of the immunisation (three to five days after the last injection), 100 mL blood was drawn from the two llamas, respectively. Subsequently, peripheral blood mononuclear cells were isolated from the blood samples. This cell fraction contains the HCAb-producing B cells. Their mRNA was extracted and reverse-transcribed into cDNA. The cDNA then served as a template to selectively amplify the VHH locus and thereby create a first nanobody DNA library. This library was subsequently cloned into the phage display vector pMESy4. In this vector, each nanobody is fused in-frame to an N-terminal PelB sequence for recombinant periplasmic expression in *E. coli*, and a C-terminal His<sub>6</sub>-EPEA-tag, followed by an amber stop codon (TAG) and the pIII protein of the M13 bacteriophage. The *E. coli* (TG1) used for the nanobody phage displays are amber-suppressing. In this strain, the TAG codon is translated into a glutamine residue, wherefore the polypeptide chains of the nanobody and the pIII proteins are fused (Suzuki et al., 2005). The pIII protein leads to the integration of the nanobody-His<sub>6</sub>-

EPEA-pIII fusion at the tip of assembling M13 bacteriophages. Once the VHH phagemid libraries were generated, phage displays were performed. The proceedings for the nanobody phage displays were in principle similar the ones for the sybody selections (Section 2.4). Important differences are described in the following. Antigens were immobilised prior to their incubation with the nanobody-containing phages. This is referred to as solid panning. The immobilisation of antigen was exclusively performed in a plate-format. For the panning, antigens were directly adsorbed to the solid support in their unmodified form (solid phase immobilisation). Alternatively, biotinylated antigens were immobilised on a neutravidin-coated surface (neutravidin-capture immobilisation). The latter approach was preferred, as it is more likely to preserve native epitopes of the antigens. Antigen-specific enrichment of phages was monitored by the number of infectious particles, rather than qPCR. Once satisfying enrichments were achieved (10-1,000 x relative to an antigen-free control), individual nanobody clones were expressed in *E. coli* (TG1) in a 96 to 192 well format. Periplasmic extracts were analysed by ELISA analogously to Section 2.4. Alternatively, the clones were directly sequenced by Sanger sequencing. Sequence alignment, using the QIAGEN CLC Sequence Viewer, allowed to quickly identify enriched nanobody families. Selected representatives of these families were tested for their interaction with antigen through an ELISA. Antigen-binding nanobodies were then further assessed for their physical interaction with hSLC19A3 and hSLC19A2 by biolayer interferometry (BLI, see Section 2.6). The protocol for the nanobody discovery was developed by Jan Steyaert and Els Pardon and their co-workers. Throughout the discovery campaigns I had great support from their team, particularly from Els Pardon, Alison Pirro Lundqvist and Eva Beke. A more detailed version of the nanobody discovery protocol can be found in Pardon et al., 2014.

## **Nanobody production**

For the larger-scale expression of a selected nanobody, the respective pMESy4-nanobody plasmid DNA was transformed into *E. coli* (WK6). Switching to this *E. coli* strain is essential for the production of pIII-free nanobodies, as WK6 cells are not amber-suppressing. By setting the amber stop codon back into full effect, the proteinogenic connection between the nanobody and the encoded pIII protein is never established. Another important notion for work with the pMESy4 plasmid is that the nanobody expression is controlled by both the IPTG-sensitive LacI repressor and the indirectly glucose-inhibited CAP activator protein. For the propagation of the vector and the proper growth of the bacterial cells, high levels of glucose of 2% (w/v) should be added to the medium to avoid a leaky transcription of the nanobody genes. This can otherwise be severely bacteriostatic - I speak from experience. For the overexpression of nanobodies, glucose levels should be reduced and IPTG should be added to release the LacI repressor once an adequate cell density is reached.

Precultures of pMESy4-nanobody transformed *E. coli* (WK6) were set up in LB medium, containing 100 µg/mL ampicillin and 2% glucose (w/v) and grown overnight at 37 °C in a shaking incubator. For large-scale nanobody production, these precultures were inoculated at a ratio of 1:200 (v/v) in 2 x 500

mL TB medium, containing 100 mg/mL ampicillin and 0.1% glucose, in 1 L flasks, and grown at 37 °C, shaking at 165 rpm. At an  $OD_{600}$  of 0.8-1.0, overexpression of the nanobodies was induced by adding 1 mM IPTG. The cultures were then grown overnight at 22 °C, shaking at 165 rpm. Cells were subsequently harvested for 15 min at 7,000 rpm, yielding typically about 15 g of cell mass. Cell pellets were frozen and stored at -70 °C. Periplasmic extraction and purification of the nanobodies was carried out as described for hTPK1 (Section 2.3). The only nanobody-specific deviation was the use of a size-exclusion column with a smaller pore size (Superdex S75 10/300 Increase), which has a better size-resolution for the 12-15 kDa nanobodies. At the end of the purification, the nanobodies were in TBS buffer (20 mM Tris-HCl, pH7.4, 150 mM NaCl). The purified nanobodies were concentrated using 10 kDa molecular weight cut-off concentrators, flash-frozen in liquid nitrogen, and stored at -70 °C.

## **2.6 Biolayer interferometry (BLI)**

The interaction between membrane transporters and nanobodies was analysed using biolayer interferometry (BLI) on an Octet RED96 system (FortéBio). The nanobodies were immobilized at a concentration of 300 nM through their C-terminal His<sub>6</sub>-tag on Octet® HIS1K Biosensors (pre-equilibrated in BLI buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.02% LMNG (w/v), 0.002% CHS (w/v), 0.1% BSA (w/v), 0.5 mM thiamine)). Following a 60-second baseline measurement, the biosensors were transferred into BLI buffer containing the membrane transporters at concentrations between 25 and 800 nM and incubated for 240 seconds, allowing the transporters to engage the immobilized nanobodies. The sensors were then moved to fresh BLI buffer to monitor the dissociation of bound transport proteins over 600 seconds. All BLI experiments were conducted at 22 °C. Data analysis was performed using the Octet Data Analysis software v.10.0 (FortéBio), with reference subtraction and alignment applied. Binding affinities were calculated based on the maximum response values, assuming a steady-state binding equilibrium and a 1:1 binding stoichiometry.

## **2.7 Thermal stability assays (nanoDSF)**

The thermal stability of proteins and their binding affinities for small molecules were assessed using nanoDSF on a Prometheus NT.48 (NanoTemper Technologies). For sample preparation, the respective purified protein was diluted to a final concentration of 4 µM in TBS-D buffer (20 mM Tris-HCl, pH7.4, 150 mM NaCl; for membrane proteins supplied with 0.002% LMNG (w/v) and 0.0002% CHS (w/v)). When the effect of a small molecule compound on the thermal stability of a protein should be analysed, that particular compound was added at concentrations ranging from 2 µM to 2 mM to the protein sample. Prior to nanoDSF measurements, samples were incubated for 30 minutes at room temperature to allow for a steady-state binding equilibrium to be established. Melting scans were performed over a

temperature range from 20 to 95 °C, with a temperature increase of 1 °C per minute. Thermal unfolding was monitored by measuring the fluorescence of the protein's tryptophanes at 350 nm (F350) and 330 nm (F330). Thermal unfolding curves were commonly based on the ratio of F350/F330. Some compounds, however, masked the melting transitions of the proteins in the fluorescence ratio. In these cases, the F350 alone was used as an alternative readout. The protein melting temperature ( $T_m$ ) was identified as the inflection point of the thermal unfolding curve, which corresponds the local maximum or minimum of its first derivative. The melting scan data were analysed using the manufacturer's software (PR.ThermControl, v 2.1.2, NanoTemper Technologies). Ligand-induced thermal shifts ( $\Delta T_m$ ) were calculated as the difference between the melting temperatures of the ligand-bound state ( $T_m$ ) and the apo state ( $T_m^{apo}$ ) of the analysed protein. To determine apparent binding affinities at 25 °C,  $T_m$  values from dilution series of the respective small molecule ligands were used to fit the parameters of the following equation (Eq. 4).

$$\text{Eq. 4} \quad \frac{\Delta T_m}{T_m^{apo}}(L) = \frac{-RT_{std}}{E_{a1}} * \ln\left(\frac{K_{d,app}}{K_{d,app}+L}\right)$$

with L being the ligand concentration, R the universal gas constant,  $T_{std}$  the chosen standard temperature of 298.15 K (25 °C), and  $E_{a1}$  the activation energy for the unfolding of the apo state. As R,  $T_{std}$  and L are known and the  $\Delta T_m$  and  $T_m^{apo}$  are experimentally determined, it is possible to use non-linear curve fitting to estimate  $E_{a1}$  and  $K_{d,app}$ . The fitting was done in symfit (v. 0.5.3), as described in Kotov et al., 2023, based on theoretical considerations formulated in Hall, 2019.

## 2.8 Mass spectrometry-based cellular thiamine uptake assays

The uptake of deuterated thiamine (thiamine- $d_3$ ) by hSLC19A3-wt and mock-transfected Expi293F<sup>TM</sup> cells was measured using LC-MS/MS. Expi293F<sup>TM</sup> cells were transfected as described before and grown for 48 hours at 37 °C, under 8% CO<sub>2</sub> (v/v), and shaking at 270 rpm. Thereafter, the cells were harvested by centrifugation for 5 minutes at 250 x g. For the uptake assays, the cells were resuspended at a cell density of 2 million cells per mL in 1 x PBS (pH 7.4). The resulting cell suspension was subsequently aliquoted into a 96-deep well plate. The assay was set up with three technical replicates per condition ( $n = 3$ , 1 mL per well). Each well was supplied with 200  $\mu$ M of the respective compound to be tested. The 96-well plate was then incubated for 1 minute on a shaking platform. Next, 2  $\mu$ M of thiamine- $d_3$  were added, and the plate was incubated for an additional 5 minutes. The cells were then pelleted for 5 minutes at 500 x g and washed three times with 1 mL of 1 x PBS (pH 7.4). Finally, the cell suspensions were transferred to a fresh 96-deep well plate. The cells were pelleted, aspirated, and stored at -70 °C.

From here on, the mass spectrometry workflow was performed by Bernhard Drotleff at the EMBL Metabolomics Core Facility in Heidelberg. The description of this method is therefore adopted directly from Gabriel et al., 2024: "For the LC-MS/MS analysis, the samples were extracted by adding 500  $\mu$ L

of a mixture of H<sub>2</sub>O:MeOH:ACN (1:1:1, v/v), containing 5% formic acid. After vortexing and ultrasonication in a water bath for 5 min at 4°C, samples were incubated for 20 min at -20°C. Ultimately, samples were centrifuged at 15,000 × g and 4 °C for 10 minutes using a 5415R microcentrifuge (Eppendorf, Hamburg, Germany). The supernatants were then transferred for LC-MS analysis, which was initiated within one hour after sample preparation. LC-MS/MS analysis was performed on a Vanquish UHPLC system coupled to an Orbitrap Exploris 240 high-resolution mass spectrometer (Thermo Fisher Scientific, MA, USA) in positive ESI (electrospray ionisation) mode. Chromatographic separation was carried out on an Atlantis Premier BEH Z-HILIC column (Waters, MA, USA; 2.1 mm x 100 mm, 1.7 μm) at a flow rate of 0.25 mL/min. The mobile phase consisted of water:acetonitrile (9:1, v/v; mobile phase A) and acetonitrile:water (9:1, v/v; mobile phase B), which were modified with a total buffer concentration of 10 mM ammonium formate. The aqueous portion of each mobile phase was adjusted to pH 3.0 via addition of formic acid. The following gradient (8 min total run time including re-equilibration) was applied (time [min]/%B): 0/90, 3/85, 3.5/60, 4/60, 5/90, 8/90. Column temperature was maintained at 40°C, the autosampler was set to 4°C and sample injection volume was set to 3 μL. Analytes were recorded via a full scan with a mass resolving power of 120,000 over a mass range from 60 - 900 *m/z* (scan time: 100 ms, RF lens: 70%). MS/MS fragment spectra were recorded via targeted product ion scans for Thiamine ([M]<sup>+</sup>, *m/z* = 265.1118) and Thiamine-d<sub>3</sub> ([M]<sup>+</sup>, *m/z* = 268.1306) at a resolving power of 15,000, stepped collision energies [%]: 20/35/50, and total cycle time of 3 sec. Ion source parameters were set to the following values: spray voltage: 3500 V, sheath gas: 30 psi, auxiliary gas: 5 psi, sweep gas: 0 psi, ion transfer tube temperature: 350°C, vaporizer temperature: 300°C. All experimental samples were measured in a randomized manner. Pooled quality control (QC) samples were prepared by mixing equal aliquots from each processed sample. Multiple QCs were injected at the beginning of the analysis in order to equilibrate the analytical system. A QC sample was analysed after every 6<sup>th</sup> experimental sample to monitor instrument performance throughout the sequence. For determination of background signals and subsequent background subtraction, an additional processed blank sample was recorded. Data was processed using TraceFinder 5.1 and raw peak area data was exported for relative metabolite quantification.”

## 2.9 Cryo-EM sample preparation and data collection

For cryo-EM, all proteins were prepared in TBS-D buffer (20 mM Tris-HCl, pH7.4, 150 mM NaCl, 0.002% LMNG (w/v), 0.0002% CHS (w/v)). hSLC19A3-wt was prepared in complex with the nanobodies Nb3.3 and Nb3.4, while hSLC19A3-gf was combined with Nb3.7. hSLC19A2-wt and hSLC19A3-ALFA were prepared in complex with Nb2.1 and Nb2.2, and Nb-ALFA, respectively. The membrane transporters were concentrated to 60 μM and mixed and diluted with a 1.5-fold molar excess of nanobody in TBS buffer (20 mM Tris-HCl, pH7.4, 150 mM NaCl), resulting in final concentrations of 30 μM membrane transporter and about 45 μM nanobody. The following compounds were added



alongside the nanobody at the indicated final concentrations: 500  $\mu\text{M}$  thiamine, 250  $\mu\text{M}$  fedratinib, 300  $\mu\text{M}$  amprolium, 200  $\mu\text{M}$  hydroxychloroquine, 250  $\mu\text{M}$  domperidone. The samples were subsequently incubated on ice for 4 to 18 hours to allow for the full saturation of the transporters with nanobodies. Cryo-EM samples of hTPK1 were prepared by diluting the concentrated protein prepared in TBS buffer with 10 volumes TBS-D buffer, containing 0.002% LMNG (w/v) and 0.0002% CHS (w/v). hTPK1 was then concentrated again to a final concentration of 30  $\mu\text{M}$ . This exchange to detergent-containing buffer was performed to improve the quality of the vitreous ice (more homogeneous ice gradient, better spreading of sample on the grid squares/foil holes) and to interfere with preferred orientation of the kinase on the air-water interface on the EM grids. For the preparation of EM grids, 3.6  $\mu\text{L}$  of the respective samples were applied to glow-discharged holey carbon film grids (Quantifoil 300 mesh, Au, R1.2/1.3 or R2/1) for 15 seconds at 10 °C and 100% humidity using a Vitrobot Mark IV (Thermo Fisher Scientific). Blotting was performed for 3.5 seconds with a blot force of -5 and a drain time of 0.5 seconds. The grids were then plunged into a liquid ethane/propane mixture and stored in liquid nitrogen.

Cryo-EM data were generally collected on the two Titan Krios G3 (FEI) electron microscopes at the CSSB Hamburg, both equipped with an X-FEG operating at 300 kV, a K3 direct electron detector camera (Gatan), and a BioQuantum imaging filter (Gatan) with a slit width set to 20 eV. These data were collected at a nominal magnification of 105,000 x, corresponding to a pixel size of 0.83 or 0.85 Å, and at total doses of 45-60  $e^-/\text{Å}^2$ . The data of hTPK1 and the hSLC19A3-Nb3.3-hydroxychloroquine complex were collected on a Titan Krios G4 (FEI) electron microscope at the EMBL Heidelberg Imaging Centre, equipped with a cold-FEG operating at 300 kV, a Falcon 4i electron detector camera, and a SelectrisX imaging filter set to a slit width of 10 eV. These data were collected at a nominal magnification of 215,000 x, corresponding to a pixel size of 0.572 Å, and at total doses of 60-70  $e^-/\text{Å}^2$ . All data were collected in electron counting mode. More details on the data collections can be found in Supplementary Table 1.

## 2.10 Cryo-EM data processing

The processing of acquired EM movies was carried out using cryoSPARC (v.4.4-4.6). Initially, the collected movies were subjected to patch motion correction. Motion corrected micrographs meeting high-quality criteria were selectively retrieved for further processing (CTF resolution < 3.5 Å, relative ice thickness < 1.05, total frame motion distance < 25 Å). Particle picking was performed using a template picker. The 2D references for this approach were generated from a 3D reconstruction, obtained through the processing of a small subset of the data (200-600 micrographs, subjected to blob-picking (100 Å blob), 2D classification, ab-initio reconstruction, and non-uniform refinement (NU-refinement)).

After template-based particle picking across the entire dataset, protein-containing and well aligning particles were iteratively enriched through several rounds of 2D classification. The selected particles

were then subjected to a series of multi-class *ab-initio* reconstructions and NU-refinements to produce a first high-resolution reconstruction, typically in the nominal resolution range of 3.5-4.0 Å. This density map was used, along with random-generated volumes, for performing multiple rounds of heterogeneous refinement on the initial template-picked particle stacks to retrieve good particles through 3D alignment. The selected particles were then subjected to NU-refinement, followed by local motion correction and re-extraction at larger box sizes of ~326 Å (384 px without Fourier cropping for datasets of transporter-nanobody complexes, collected on K3 cameras), ~292 Å (1024 px, Fourier cropping to 512 px for the hSLC19A3-Nb3.3-hydroxychloroquine dataset, collected on Falcon 4i camera), and ~225 Å (784 px for the hTPK1 dataset, collected on Falcon 4i camera). After another NU-refinement, the resulting density map was locally refined with a mask encompassing the transporter-nanobody complex, or the hTPK1 dimer. Map sharpening was performed using the global B-factor estimates from the respective Guinier plots. For further details on EM data processing, see Supplementary Fig. 1. Q-scores were calculated the MapQ plugin in UCSF Chimera, developed by Grigore Pintilie (Pintilie et al., 2020).

## **2.11 Model building and refinement**

AlphaFold2 predictions were used as initial structure models for hSLC19A3, the nanobodies, and hTPK1 (Jumper et al., 2021). These protein models were refined in a first round in the corresponding EM density maps using ISOLDE in ChimeraX (Croll, 2018, Pettersen et al., 2021). Small molecule ligands were built with eLBOW package of Phenix (Afonine et al., 2018) and integrated into the protein structure models in Coot. The resulting models underwent iterative refinement with Phenix (auto-refinement) and Coot (manual refinement). Visualisations were created using UCSF ChimeraX (Pettersen et al., 2021).

### 3. Results

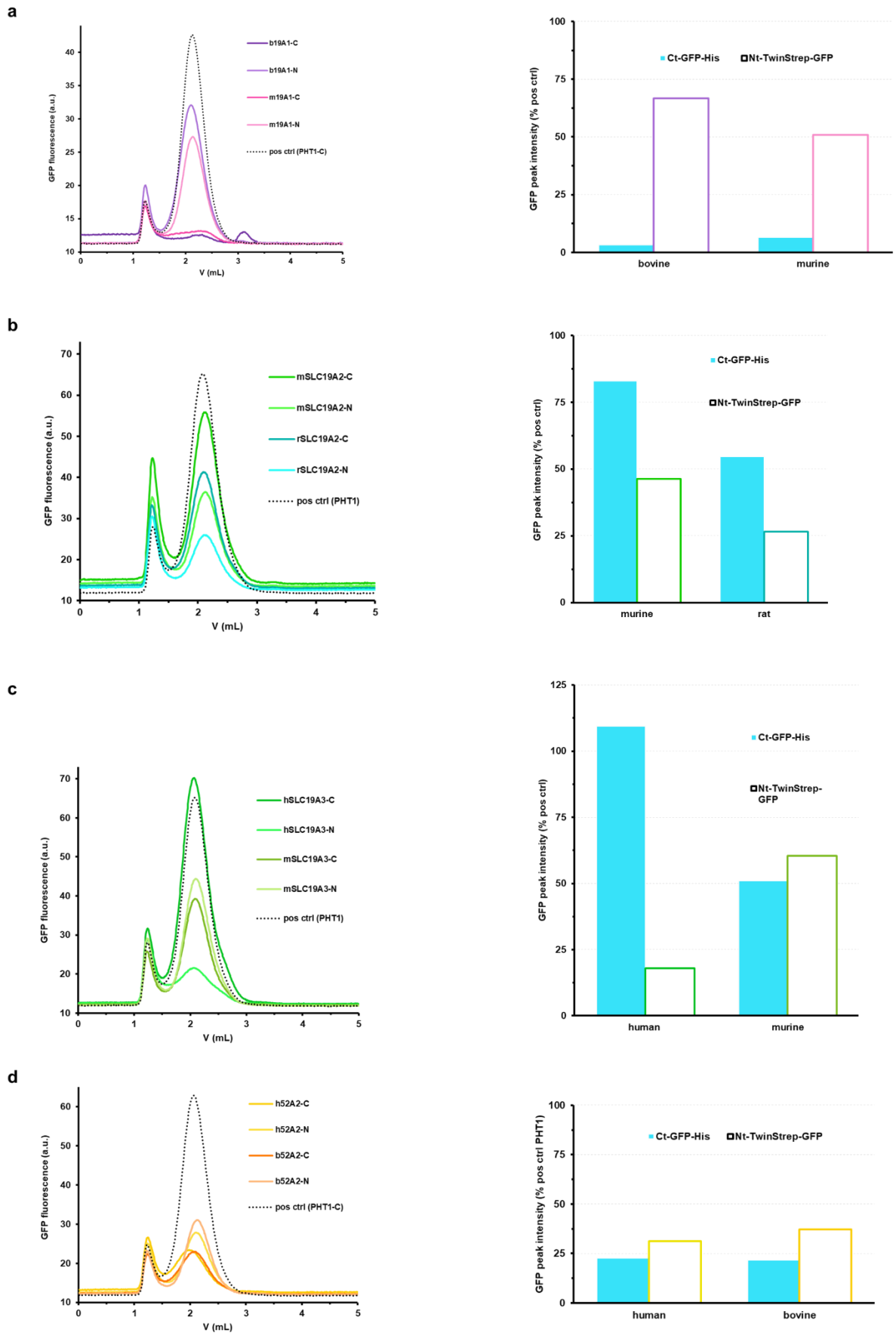
#### 3.1 Expression, purification, and biophysical characterisation of hSLC19A3

##### Fluorescence-based screening of mammalian vitamin transporters

In the beginning of this study, I screened twelve mammalian vitamin transporters (MVTs) for their expression and solubilisation behaviour (Table 5). For this purpose, I cloned their respective cDNA into pXLG expression vectors, providing the transporters with either an N-terminal Twin-Strep-GFP-cTEV or a C-terminal cTEV-GFP-His<sub>6</sub>-tag (cTEV: cleavage site for TEV protease). The GFP-tags allowed to measure the expression and solubilisation levels of the transporters by fluorescence microscopy (data not shown) and fluorescence-detection size exclusion chromatography (FSEC, Fig. 6). For FSEC, the MVT-GFP fusion proteins were expressed in HEK293F cells and solubilised in DDM/CHS-containing buffer. The screening of the different MVT-GFP constructs revealed that the mammalian transporters for folate (vitamin B<sub>9</sub>), thiamine (vitamin B<sub>1</sub>), and riboflavin (vitamin B<sub>2</sub>) could be produced at yields sufficient for biophysical and structural studies. This was assessed by comparing the FSEC traces of the MVTs to the well-purifiable transporter PHT1 (Fig. 6). Among the studied MVTs, the human SLC19A3 (hSLC19A3) showed particularly high yields (Fig. 6). When comparing the N- and C-terminally tagged hSLC19A3, the latter one showed a higher peak in the FSEC (Fig. 6). I therefore decided to continue my work with the C-terminally tagged hSLC19A3.

**Table 5: Overview of the screened mammalian vitamin transporters (MVTs).**

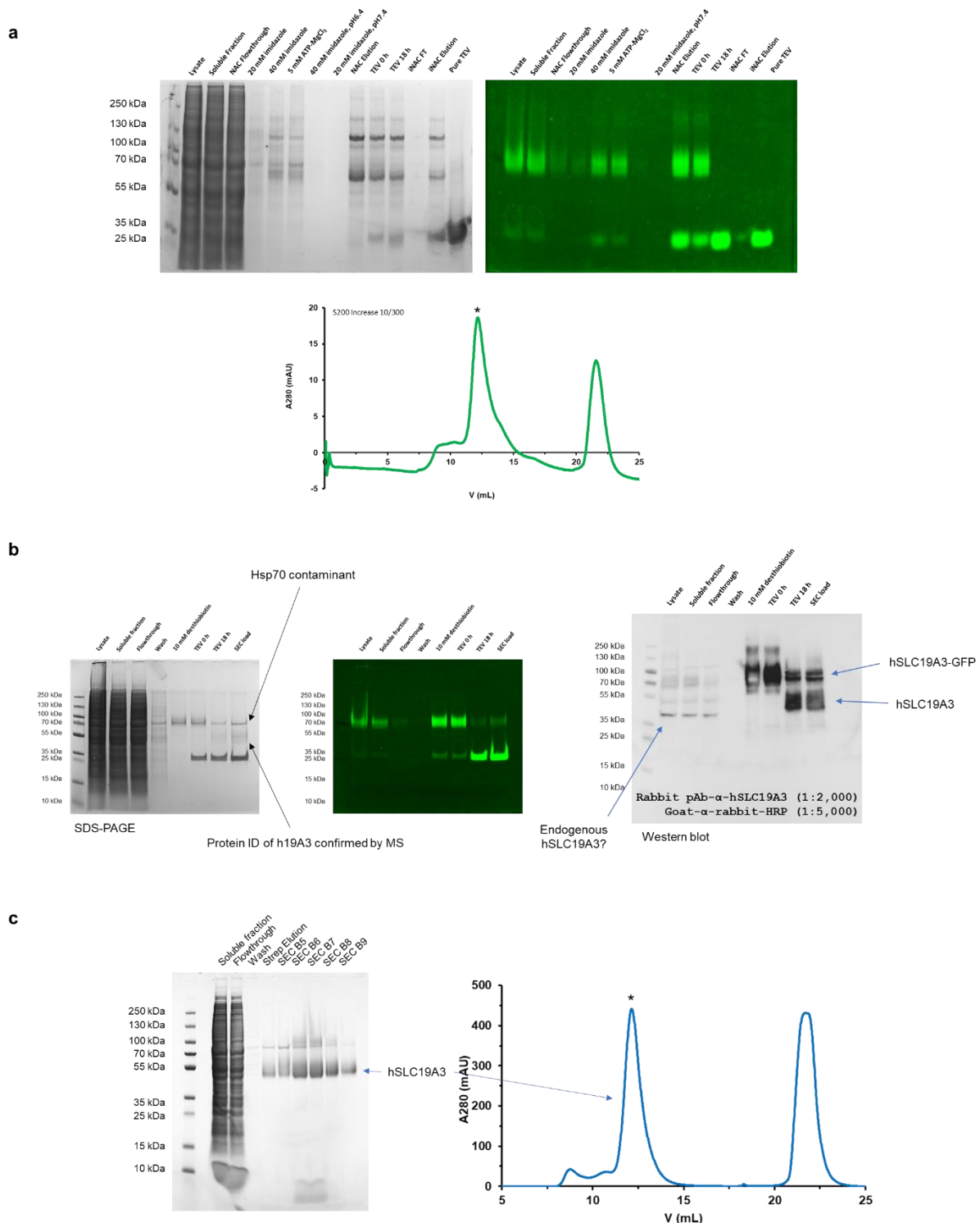
Transporter	Alternative name	Main substrate	Species
SLC19A1	Reduced folate carrier	Folate	Homo sapiens, (human), Bos taurus (cattle)
SLC19A2	Thiamine transporter 1	Thiamine	Mus musculus (mouse), Rattus norvegicus (rat)
SLC19A3	Thiamine transporter 2	Thiamine	Homo sapiens (human), Mus musculus (mouse)
SLC52A1	Riboflavin transporter 1	Riboflavin	Homo sapiens (human), Bos taurus (cattle)
SLC52A2	Riboflavin transporter 2	Riboflavin	Homo sapiens (human), Bos taurus (cattle)
SLC52A3	Riboflavin transporter 3	Riboflavin	Homo sapiens (human), Bos taurus (cattle)



**Figure 6: FSEC of mammalian vitamin transporters.** 48 h post transfection, mammalian vitamin transporter-GFP fusion protein (MVT-GFP) overexpressing HEK293F cells were harvested and solubilised in 1 x PBS pH7.4, containing 1% DDM (w/v) and 0.1% CHS (w/v). FSEC of the soluble fractions of the respective cell lysates was performed using a Superose 6 5/150 Increase column on an Agilent system. The FSEC chromatograms were used to assess the quantities of solubilised and properly folded vitamin transporters. The left panels show the individual chromatograms. The dotted line indicates the FSEC trace of a positive control transporter, PHT1-GFP. The right panels show the peak height (here = peak intensity) of the respective constructs. N- and C-terminally tagged constructs are directly compared. **a** Folate transporters (SLC19A1). **b** Thiamine transporter 1 (SLC19A2). **c** Thiamine transporter 2. **d** Riboflavin transporter (SLC52A2).

### Optimising the purification of hSLC19A3

Early on, I decided to optimise the purification of hSLC19A3 in detergent solution. Due to its favourable biochemical characteristics, the detergent combination of lauryl maltose neopentyl glycol (LMNG) and cholesterol hemisuccinate (CHS), mixed at a ratio of 1:10 (w/w) was chosen as the default solubilisation agent for the membrane transporter. LMNG has a low critical micelle concentration (CMC) of around 0.001% (w/v). This allows one to work with close to detergent-free buffers at later stages of the purification (Cho et al., 2015). Furthermore, the combination of LMNG with the negatively charged cholesterol derivative CHS is known to produce compact micelles, which are well-suited for determining high-resolution cryo-EM structures of integral membrane proteins (Yuan et al., 2022). Using LMNG/CHS, the hSLC19A3-cTEV-GFP-His<sub>6</sub> construct was successfully solubilised and purified (Fig. 7). However, both the purity and the yield of the transporter came short of what would be required for biophysical and structural studies. To improve the purification of hSLC19A3, the transporter was subcloned to switch its C-terminal tag first to a cTEV-GFP-TwinStrep-tag. The Strep-tag-based clean-up of the transporter showed a strong improvement in terms of purity (Fig. 7). By subsequently removing the GFP from the tag, we could see a strong boost in yields of the recombinantly expressed hSLC19A3-c3C-TwinStrep (c3C: cleavage site for HRV 3C protease). A change from the HEK293F to the Expi293F<sup>TM</sup> cell line for the overexpression of hSLC19A3 further improved the yields (Fig. 7). With these adaptations, recombinant hSLC19A3 could be purified at yields of 1-2 mg from 250 mL expression culture (Fig. 7).

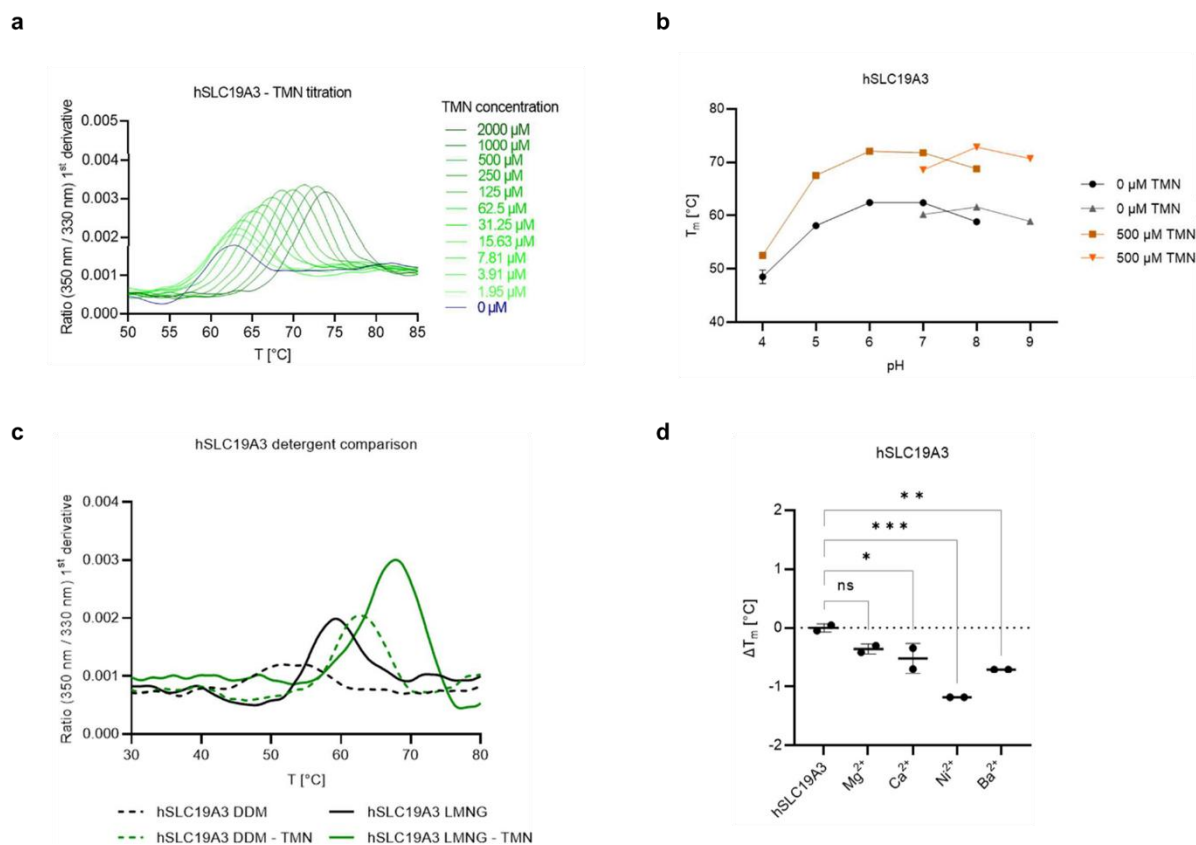


**Figure 7: Purification of hSLC19A3 in LMNG/CHS. a** Purification of hSLC19A3-cTEV-GFP-His<sub>6</sub>, including TEV protease cleavage. The left panel shows SDS-PAGE gel of the purification stained with Coomassie. The panel to the right shows in-gel GFP fluorescence. The panel below shows the SEC trace of the transporter after Nickel-affinity chromatography (NAC), TEV cleavage, and inverse Nickel-affinity chromatography (iNAC). **b** Purification of hSLC19A3-cTEV-GFP-TwinStrep. The left panel shows a corresponding SDS-PAGE gel stained with Coomassie. The panel in the middle shows in-gel GFP fluorescence. The panel to the right shows a Western blot of the same SDS-PAGE gel using a polyclonal rabbit anti-hSLC19A3 primary antibody (ab237630, Abcam), and a goat anti-rabbit horse radish peroxidase (HRP)-conjugated secondary antibody at the indicated dilutions (v/v). **c** Purification of hSLC19A3-c3C-TwinStrep. The panel on the left shows a Coomassie-stained SDS-PAGE gel of the purification. The panel to the right shows the SEC trace of the Strep-tag affinity-purified transporter. The \* indicate the SEC peaks corresponding to hSLC19A3. Purifications were done from 200 mL cell culture (HEK293FTM for (a)

and (b), Expi293FTM for (c)), respectively. Protein mass spectrometry confirmed the identity of the purified hSLC19A3. The faint band at 70 kDa could be assigned to the chaperone Hsp70 (MS experiments performed by Proteomics Core Facility, EMBL Heidelberg; data not shown).

## First biophysical insights

Once pure hSLC19A3 could be produced, I tested its interaction with thiamine using a thermal stabilisation assay in the form of nanoDSF. This assay showed a melting transition for the transporter at around 63 °C in its apo state (Fig. 8). For a human membrane protein this is already a remarkably high thermal stability. Thiamine further stabilised hSLC19A3 in a concentration dependent manner, leading to a melting temperature of about 74 °C in the presence of 2 mM thiamine (Fig. 8, Gabriel et al., 2024). The melting transitions and the responsiveness to its endogenous substrate were indications that the purified hSLC19A3 was functionally folded in LMNG/CHS. We then went on to study the behaviour of the transporter under different chemical conditions, including its response to different pH, different detergents used for solubilisation, as well as the presence of divalent ions. These findings are published in Lea Spriestersbach's master thesis (Spriestersbach, 2022), and I will just briefly summarise them here. Changing the pH over a range from pH6 to pH9 has no detectable impact on the thermostability or the response of hSLC19A3 to thiamine. However, at more acidic conditions, especially at pH4, the transporter is strongly destabilised and thermostabilisation by thiamine is not as pronounced (Fig. 8). With regard to the chosen detergent, we found that hSLC19A3 is more stable in LMNG/CHS, compared to DDM/CHS. The transporter responds to thiamine in both its LMNG- and DDM-solubilised form (Fig. 8). When assessing the impact of divalent ions on hSLC19A3, we find that the transporter is mildly destabilised in the presence of 0.5 mM  $Mg^{2+}$ ,  $Ca^{2+}$ ,  $Ni^{2+}$ , and  $Ba^{2+}$ . The most pronounced destabilisation is induced by  $Ni^{2+}$ , with a  $\Delta T_m$  of about -1 °C (Fig. 8). Thiamine binding seems to be unperturbed by these ions (Spriestersbach, 2022).



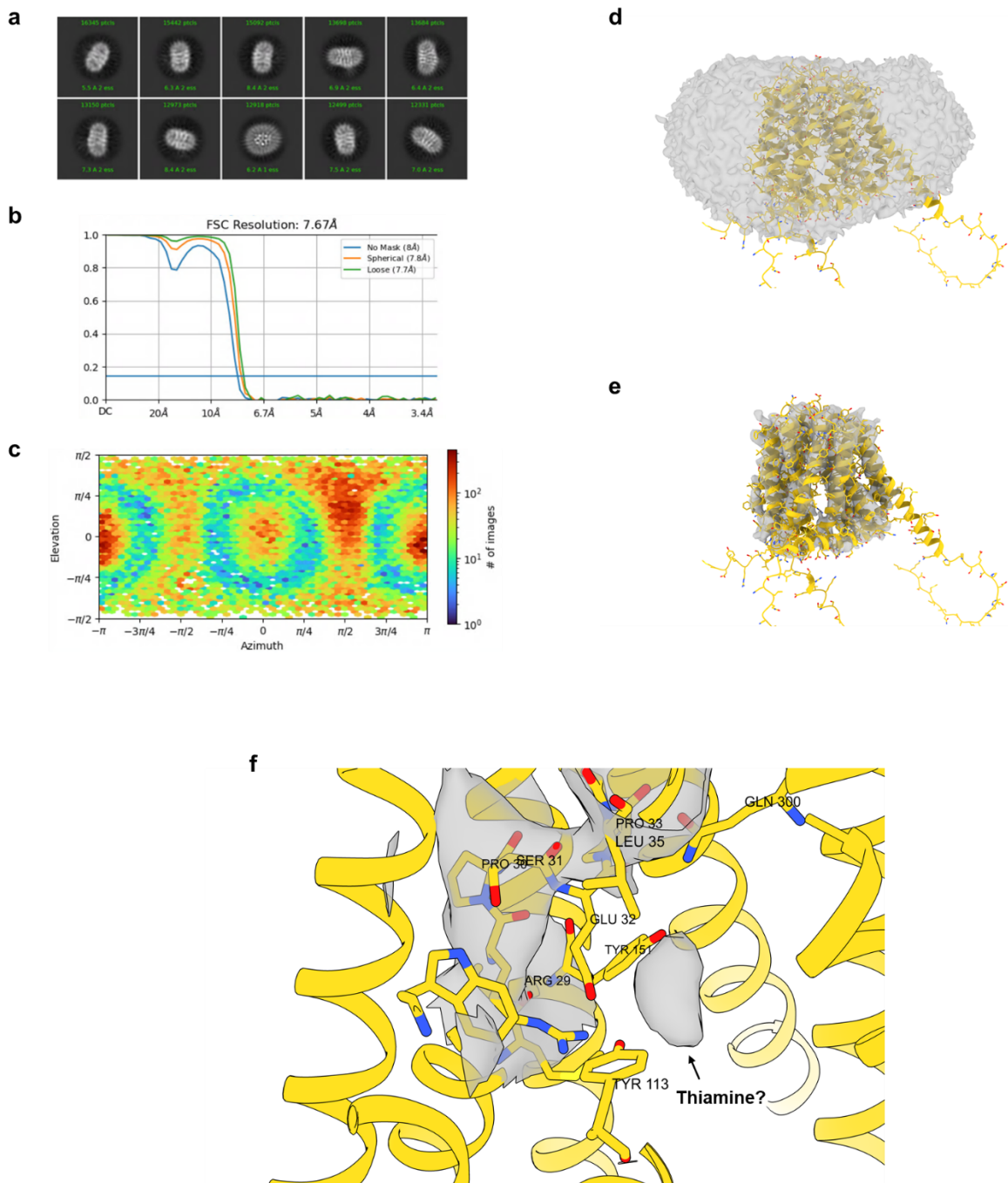
**Figure 8: Biophysical characterisation of hSLC19A3.** **a** Thermostabilisation of hSLC19A3-wt by thiamine. The graph shows the melting curves of the transporter in the presence of different thiamine concentrations, read out as the first derivative of the ratio F350/F330 of the tryptophan fluorescence (nanoDSF). **b** Melting temperatures of apo and thiamine-bound hSLC19A3-wt over a pH range from pH4 to pH9. **c** Melting curves of hSLC19A3-wt in different detergents. LMNG/CHS shows a stabilising effect when compared to DDM/CHS. **d** Effect of different divalent ions (0.5 mM) on the thermostability of hSLC19A3-wt (here in its apo state). TMN: thiamine. The panels of this figure were adapted from Priestersbach, 2022.

### 3.2 Development of structural fiducials for hSLC19A3

hSLC19A3, like most proteins of the solute carrier family, presents a small target for structure determination by cryo-EM. Its twelve transmembrane helices (about 42 kDa) were the only part of the protein likely to adopt a defined three-dimensional structure, whereas its N- and C-terminus, as well as the intracellular loop 3 (ICL3) were predicted to be flexible. Nevertheless, we attempted to refine a density map of wildtype hSLC19A3 on its own using cryo-EM. For a few other solute carriers, this approach had successfully yielded high-resolution reconstructions (e.g. Yuan et al., 2022). In our case, we could determine a low-resolution, and in retrospect accurate reconstruction of the inward-open state of the transporter (Fig. 9). However, angular refinement on a per-particle level failed to yield higher resolutions, irrespective of the algorithm used (homogeneous refinement and non-uniform refinement). A probable explanation for this problem is the low signal-to-noise ratio of the individual particle images, paired with the high degree of pseudosymmetry of the protein, which impairs proper pose refinement (Drew et al., 2021, Wentinck et al., 2022). This led us to pursue a biochemical solution of the problem



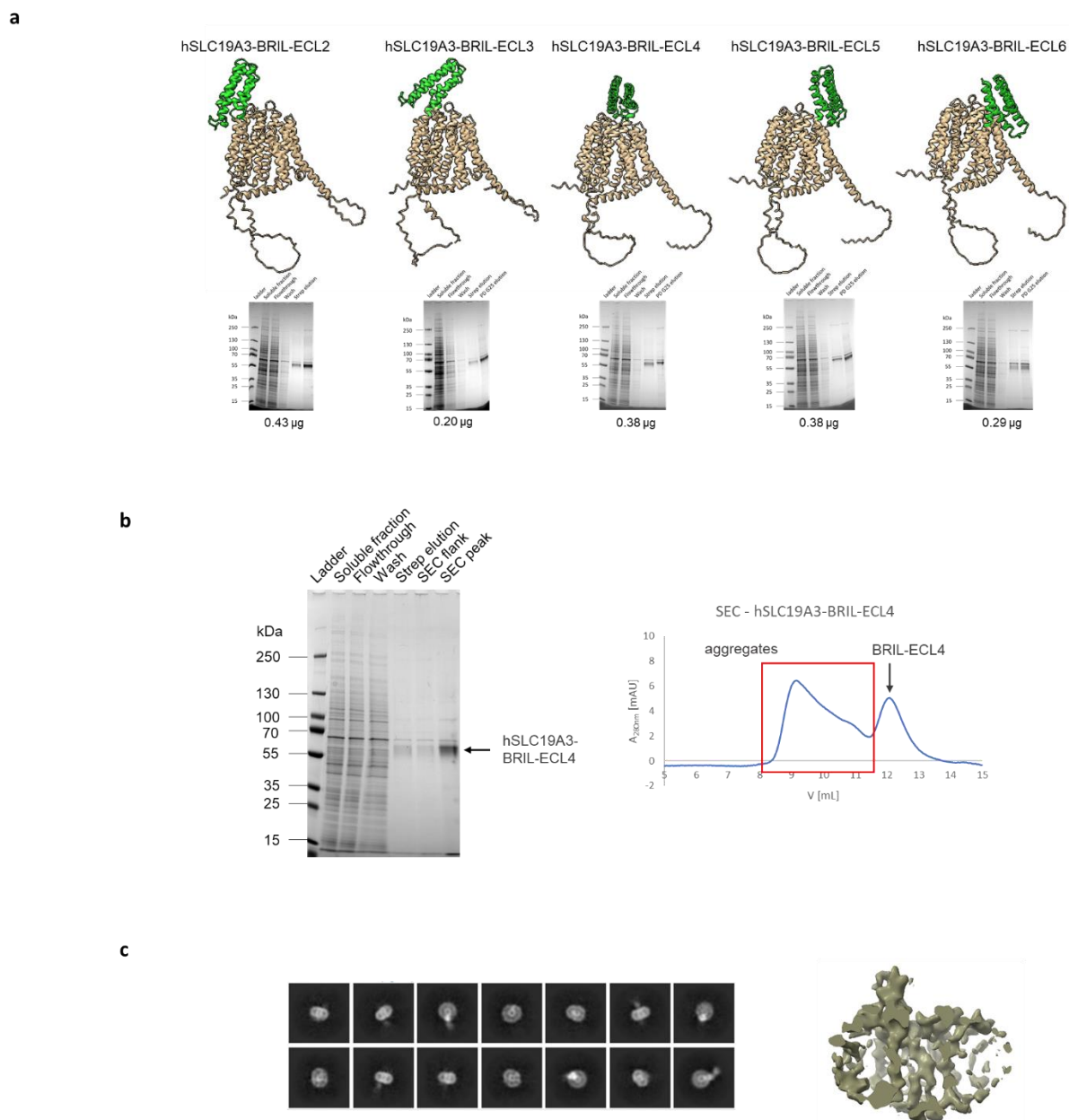
and to trying to increase the size of the structured part of the particles. The corresponding approaches and their outcome are described in the following.



**Figure 9: Low resolution reconstruction of the fiducial-free thiamine-bound wildtype hSLC19A3.** a 2D class averages of the fiducial-free transporter. b Fourier shell correlation (FSC) curve of the 3D reconstruction (non-uniform refinement) from 111,198 particles. c Pose distribution of the particles shows slightly preferred orientation. This might, however, be an artifact from missing certain views in the 2D class average-based selection of particles or the 3D alignment process on this small membrane protein target. d hSLC19A3-wt:thiamine cryo-EM density map (grey) superimposed with the AlphaFold2 model of the transporter (yellow), seen from the membrane plane. The LMNG/CHS micelle is asymmetrically deformed by the protein. e The same setting as in (d), but at a higher density map threshold. Individual  $\alpha$ -helices can be seen. f Close-up on the putative substrate binding site of hSLC19A3. An observed extra density in the binding site might be the bound substrate thiamine, which was supplied in the cryo-EM sample at a concentration of 2 mM. The location of this extra density suggests a coordination of the vitamin by Glu32, Trp59, Tyr113, and Gln300. The resolution of the 3D reconstruction is, however, too low to draw any definitive conclusions.

## BRIL constructs

In a first approach, we attempted to increase the particle size of hSLC19A3 by replacing its extracellular loops (ECLs) with a well-structured and stable bacterial apocytochrome b562 (BRIL domain, Mukherjee et al., 2020). The BRIL fusions were designed so the cytochrome would form continuous  $\alpha$ -helices with the transporter  $\alpha$ -helices adjacent to the respectively replaced ECL (structure modelling by AlfaFold2, Fig. 10). The idea behind this design was that continuous helices would provide a structurally rigid linker between the transporter and the BRIL domain. Five constructs were cloned in total (Fig. 10). They were expressed and purified following the standard protocol (methods Section 2.2). All constructs could be purified. Their yields were, however, significantly lower compared to the wildtype hSLC19A3 (Fig. 10). We prepared cryo samples of the best expressing construct (hSLC19A3-BRIL-ECL4) by adding a 1.2x excess of the high-affinity anti-BRIL Fab fragment BAK5 (Mukherjee et al., 2020) and 0.5 mM thiamine, and collected datasets on a Titan Krios G3i, at a magnification of 105,000x. Particles were picked and classified in 2D. The 2D class averages revealed the presence of the BRIL domain and the binding of the Fab fragment to that domain, as well as a good random orientation of the particles (both top views and side views were clearly present, Fig. 10). These particle averages also showed, however, that the BRIL-Fab complex does not aid the resolution of the pseudosymmetry and the detection of any higher-resolution features (e.g. helical pitch) of the transporter. This was eventually confirmed by the attempt to refine the structure of the hSLC19A3-BRIL-ECL4:Fab complex in 3D, which only reached nominal resolutions of 7-8 Å, and showed inaccurate features, like discontinuous helices crossing at the centre of the membrane plane. A probable reason for this failure of high-resolution structure determination was the flexibility between the transporter and the structural fiducial. Looking at the 2D class averages, we assume that the BRIL domain retained a high degree of conformational freedom, as its density fans out relative to the transporter (Fig. 10). The same flexibility-problem appeared to be true for the elbow region of the Fab fragment, which is a known problem for high-resolution structure determination (Kung et al., 2024). We therefore concluded that BRIL fusion was not a promising strategy for our aim to determine the structure of hSLC19A3 by cryo-EM. Most of the BRIL-related work was published in Spriestersbach, 2022.

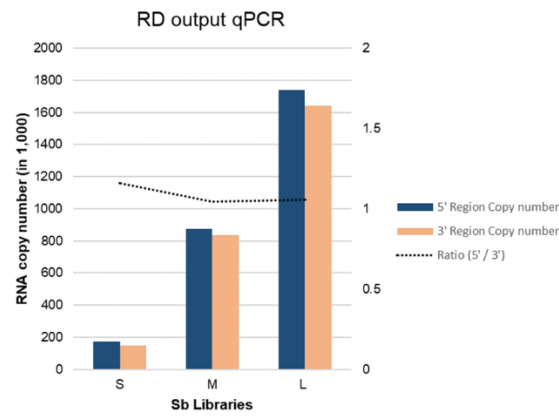


**Figure 10: Design, purification, and cryo-EM analysis of hSLC19A3-BRIL fusions.** **a** The depicted structure models show the design of the five tested BRIL fusion constructs of hSLC19A3. The transporter is coloured in light brown, the BRIL domains are shown in bright green. Below, the purification of the individual BRIL constructs is shown in the form of Coomassie-stained SDS-PAGE gels, with the respective protein yields from 250 mL expression culture indicated below. **b** Purification of hSLC19A3-BRIL-ECL4 for the preparation of a cryo-EM sample from 250 mL expression culture. The left panel shows a Coomassie-stained SDS-PAGE gel of the purification, the right panel the SEC trace of the Strep-tag affinity-purified sample. **c** The left panel shows 2D class averages from a cryo-EM dataset collected of hSLC19A3-BRIL-ECL4 in complex with the anti-BRIL Fab fragment BAK5. The right panel shows a cross-section of a 3D reconstruction from this dataset at a nominal resolution of about 7 Å. The panels in this figure were adapted from Spriestersbach, 2022.

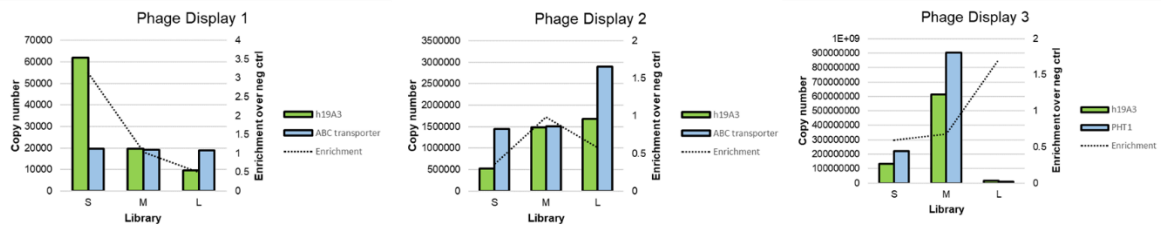
## Sybody discovery

Another strategy to provide hSLC19A3 with a structural fiducial was the selection of sybodies that would stably bind to the transporter. The VHH libraries and the protocol used for sybody discovery were initially developed in the laboratory of Markus Seeger at the University of Zurich (Zimmermann et al., 2018, Zimmermann et al., 2020). For the presented work, four sybody discovery campaigns were performed, differing in the used hSLC19A3-antigen and hSLC19A3-specific ligands used over the course of the discovery process. Fig. 11 shows exemplarily the results of the first sybody discovery campaign, performed strictly adhering to the protocol by Zimmermann et al., 2020, using wildtype hSLC19A3 expressed in Expi293F-Gnt1- cells and deglycosylated with EndoH, and 2 mM thiamine as ligand in all buffers. The outcome of each selection round was monitored by qPCR. The respective analysis after the initial round of ribosome display showed that the mRNA retrieved from the display showed no significant signs of degradation (Fig. 11). While the first round of the subsequent phage displays showed an about 3.5x enrichment for hSLC19A3, this could be neither recapitulated nor enhanced in the ensuing rounds of phage display. In phage display round two and latest in round three, we would expect an enrichment of at least 10-100x, if hSLC19A3-specific binders were captured in the discovery process. Despite this outcome, we expressed 138 individual sybodies from the three different libraries after phage display round three and tested them for specific hSLC19A3 binding using ELISA. Twelve sybodies showed mildly specific binding (enrichment of ELISA signal 1.1-4.2x over negative control). After sequencing of these clones, eight of these binders were confirmed to be unique. When tested for hSLC19A3 binding in size-exclusion chromatography (SEC) and biolayer interferometry (BLI), however, none of these sybodies showed complex formation with the transporter. For the ensuing sybody discovery campaigns, the following changes were implemented to improve the chances of a positive outcome. The antigen was prepared using a glycosylation-free mutant of the transporter, hSLC19A3-gf, in which its two predicted N-glycosylation sites were mutated (Asn45Gln, Asn166Gln). For each campaign, a different inhibitor of the transporter was used with the objective to stabilise different conformations of hSLC19A3. The employed inhibitors were amprolium, fedratinib, and amitriptyline (based on Vora et al., 2020), used at concentrations of 0.5 mM, 50  $\mu$ M, and 200  $\mu$ M, respectively. Unfortunately, these discovery campaigns were also not successful (data not shown). Eventually we concluded that sybody discovery was not likely to yield any useful binders for hSLC19A3. Over the lengthy procedure of each discovery campaign, many of the chosen parameters could explain the failure of the selections. But the four-time repetition under different conditions suggests the lack of hSLC19A3-specific sybodies in the input mRNA libraries, that would be sufficiently expressed and soluble enough to be enriched through the ribosome and phage display rounds.

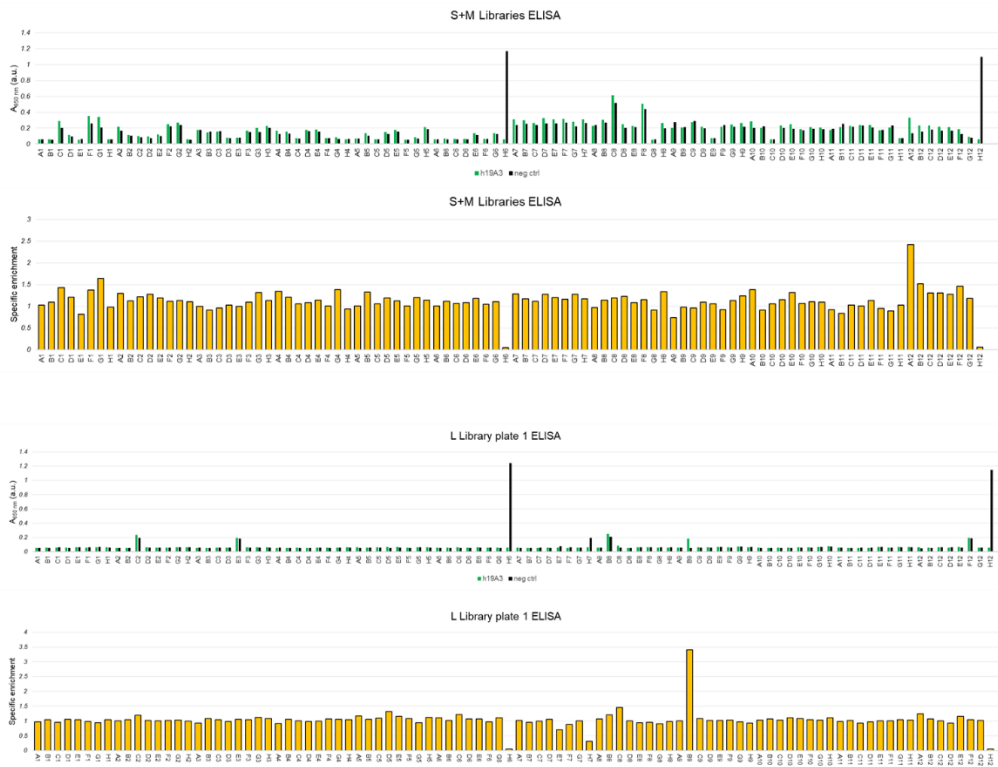
a



b



c



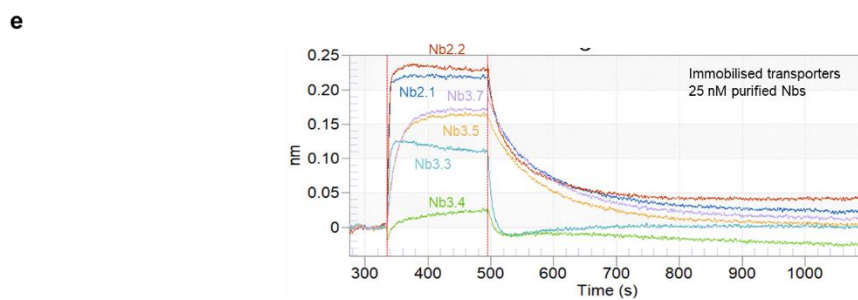
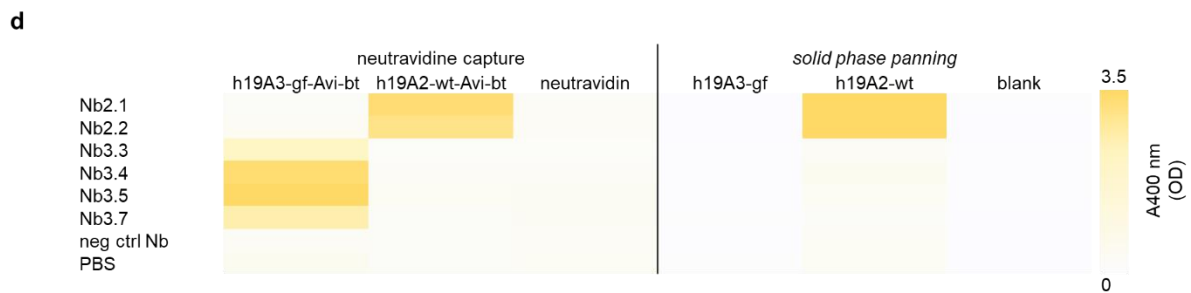
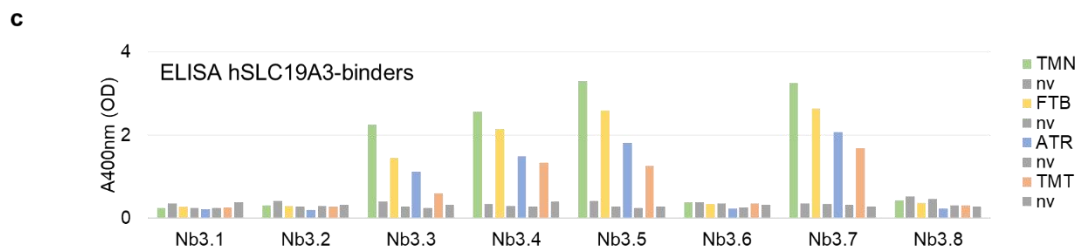
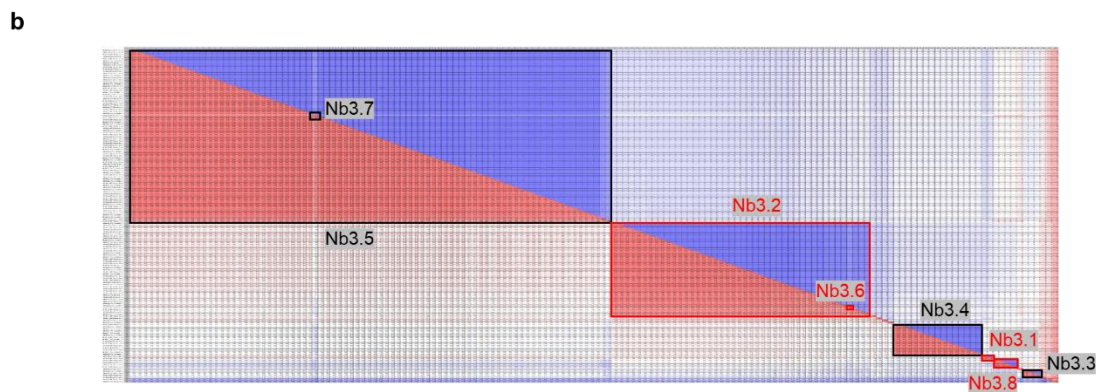
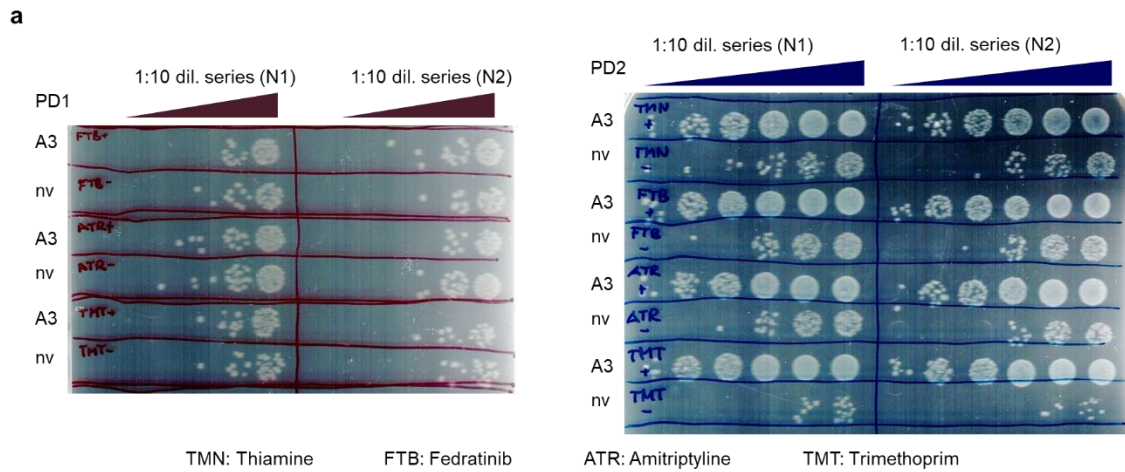
**Fig. 11: Sybody discovery campaign for hSLC19A3.** Results of one of the sybody discovery campaigns performed for hSLC19A3. For antigen preparation for this campaign, hSLC19A3-wt was expressed in Gnt1- cells and treated with EndoH in the purification process. **a** qPCR quantification of 5' and 3' mRNA sequences of the individual mRNA libraries (S, M, and L) after ribosome display. The near equal amounts of 5' and 3' mRNA indicate no major RNA degradation. **b** qPCR quantification of phages from the different rounds of phage display selection. The comparison of the phages retrieved from the target (hSLC19A3) and a negative control (ABC transporter or PHT1, as indicated) allows one to calculate the target-specific enrichment. **c** ELISAs of individual sybody clones. The upper panels show the ELISA readout (green: hSLC19A3, black:

negative control, no target added). The lower panels (yellow bars) show the hSLC19A3-specific enrichment over the negative control.

## Nanobody discovery

After both, the BRIL fusion and sybody discovery failed to provide useful fiducials for the structure determination of hSLC19A3, we decided to pursue nanobody development through the immunisation of llamas. This work was performed in collaboration with the research group of Jan Steyaert at the Vrije Universiteit Brussel. For the immunisation, the glycosylation-free mutant hSLC19A3-gf (Asn45Gln, Asn166Gln) was purified in LMNG/CHS in the presence of 2 mM thiamine. The Twin-Strep-tag of this construct was cleaved using 3C protease to avoid the generation of antibodies against this peptide sequence. The SEC-purified protein was then used as antigen for the immunisations of a llama, as described in the methods Section 2.5. The presence of hSLC19A3-specific heavy-chain-only antibodies (HCAbs) was confirmed by ELISA of the serum of the immunised llama against solid phase adsorbed hSLC19A3 (Fig. 12). After reverse transcription of mRNA extracted from peripheral blood mononuclear cells of the llama, the VHH-fragments were amplified by PCR and cloned into phage display vectors. Building on this immune library, two nanobody discovery campaigns were performed, using different panning strategies and ligands. In a first attempt, hSLC19A3 was solid phase adsorbed to MaxiSorp™ plates and two consecutive rounds of phage display were performed in the presence of 0.5 mM thiamine. This approach did not yield any significant enrichment of hSLC19A3-specific phages, most likely due to the denaturation of native epitopes of the target transporter in the solid phase adsorption. Subsequently, another discovery campaign was started using neutravidin capture of biotinylated hSLC19A3 to immobilise the antigen. In this campaign, four different ligands (thiamine, amprolium, fedratinib, and amitriptyline) of the transporter were used in separate selections to promote the discovery of conformation-specific binders. Strong enrichments were achieved after two rounds of phage display (Fig. 12). From the resulting phage libraries, 192 individual clones were sequenced. This allowed for the identification of major families of binders in the libraries. Representatives of these families were then tested for hSLC19A3-specific binding in ELISA. Eventually, three unique binders, Nb3.3, Nb3.4 and Nb3.7, were isolated and validated using biolayer interferometry (BLI, Supplementary Fig. 2). They show binding affinities in the nanomolar range and induce thermostabilisation of hSLC19A3, suggesting that they bind against structured, native epitopes of the transporter (Gabriel et al., 2024, Supplementary Fig. 3).



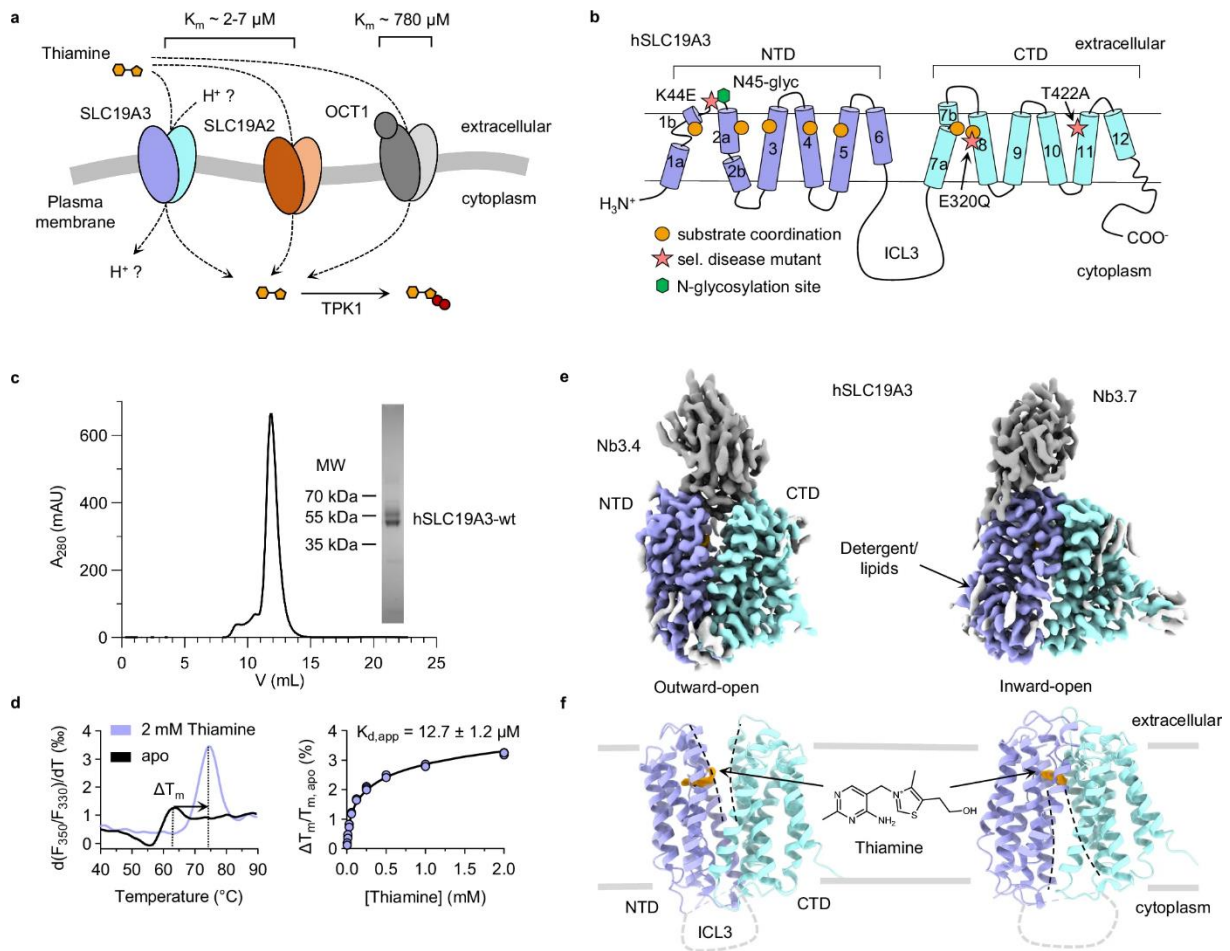


**Fig. 12: Nanobody discovery campaign for hSLC19A3.** The antigen used for this campaign was hSLC19A3-gf. **a** Quantification of target-specific phage enrichment by infection assay. The left panel shows the outcome of the first round of phage display selection, the right panel that of the ensuing second round. **b** Clustering of 192 individual clones from the phagemid libraries after two rounds of phage display. The boxes indicate families of nanobodies that were later probed for their binding behaviour. **c** ELISA of representatives of the identified nanobody families. **d** ELISA testing the specificity of binders selected against hSLC19A2 and hSLC19A3. **e** Biolayer interferometry (BLI) analysis of binding of the selected nanobodies to the immobilised transporters. Nb2.1 and Nb2.2 bind to hSLC19A2. Nb3.3, Nb3.4, Nb3.5, and Nb3.7 bind to hSLC19A3. Nanobody concentration at 25 nM.

## Cryo-EM structures of hSLC19A3 in complex with nanobodies

The nanobodies identified in the discovery process were subsequently used for structure determination of hSLC19A3 by cryo-EM (Supplementary Fig. 6-10, Supplementary Table 1). All three nanobodies bind to the extracellular solvent-exposed surface of the transporter (Gabriel et al., 2024). While the epitopes bound by the nanobodies overlap, they show differences that are important for the recognition of conformationally and chemically distinct species of hSLC19A3 (Fig. 13, Supplementary Fig. 2 & 3). Nb3.3 binds to the inward-open state of hSLC19A3 and engages both the NTD and the CTD of the transporter. A major part of the interaction between the nanobody and hSLC19A3 consists of a salt bridge network formed by the positively charged arginine residues Arg65<sup>Nb3.3</sup> and Arg108<sup>Nb3.3</sup> of CDR2 and CDR3 of Nb3.3 and Glu50<sup>hSLC19A3</sup> and Glu54<sup>hSLC19A3</sup> of the transporter. The association of the two proteins is further supported by a polar interaction of Tyr59<sup>Nb3.3</sup> with Gln428<sup>hSLC19A3</sup> and Asn432<sup>hSLC19A3</sup>. Furthermore, Trp103<sup>Nb3.3</sup> inserts between the transporter at Leu431<sup>hSLC19A3</sup> and the detergent micelle. Nb3.7 also binds to the inward-open state of hSLC19A3 and interacts only with the NTD of the transporter. It furthermore binds on top of residue 45, which allows the nanobody to stably interact with the glycosylation-free mutant Asn45Gln of hSLC19A3 but prevents binding to the wildtype transporter that is N-glycosylated in this position. The major polar contact formed between Nb3.7 and hSLC19A3 is the salt bridge formed between Arg109<sup>Nb3.7</sup> and Glu50<sup>hSLC19A3</sup>. The nanobody further inserts the apolar residues Ile103 and Leu105 of its extended CDR3 in between the transporter at around Gly100 and the detergent micelle. Nb3.4 binds to the outward-open conformation of hSLC19A3 but is, like Nb3.3, compatible with the N-glycosylation of the transporter on Asn45. The nanobody inserts its CDR3 in between the two six-helix bundles of the transporter, disrupting contacts of the extracellular gate. It inserts Lys106<sup>Nb3.4</sup> to break the hydrogen bond between the backbone carbonyl oxygen of Pro33<sup>hSLC19A3</sup> and Gln300<sup>hSLC19A3</sup>. Nb3.4 further disrupts the apolar sealing of the extracellular gate by engaging the hydrophobic residues Ile36<sup>hSLC19A3</sup> and Ile301<sup>hSLC19A3</sup> with Phe109<sup>Nb3.4</sup> and Tyr103<sup>Nb3.4</sup>, respectively. The binding of Nb3.4 to the transporter is stabilised by additional polar contacts of CDR3 with the NTD (Lys108<sup>Nb3.4</sup>-Pro37<sup>hSLC19A3</sup>, Asn110<sup>Nb3.4</sup>-Ser40<sup>hSLC19A3</sup>, Gln100<sup>Nb3.4</sup>-Ser48<sup>hSLC19A3</sup>,) and the CTD (Thr106<sup>Nb3.4</sup>-Asp304<sup>hSLC19A3</sup>). CDR1 and CDR2 of Nb3.4 engage the CTD of hSLC19A3 (e.g. Arg31<sup>Nb3.4</sup>-Asn432<sup>hSLC19A3</sup>). Nb3.4 forms no contact with the detergent micelle, in contrast to the other two nanobodies described here.





**Fig. 13: “Structure and function of hSLC19A3. a** Humans and other mammals have at least three distinct membrane transporters that can mediate thiamine uptake: SLC19A3, SLC19A2 and OCT1 (SLC22A1). SLC19A3 is essential for the uptake of thiamine across the intestinal and blood-brain barrier under physiological conditions. In the cytoplasm, thiamine is phosphorylated by the enzyme TPK1 to form the biologically active coenzyme thiamine pyrophosphate. For a comparison of the thiamine binding modes of hSLC19A3 and hOCT1, see Supplementary Fig. 16. **b** Cartoon representation of hSLC19A3, based on the determined cryo-EM structures. The transporter follows the canonical MFS fold, with twelve transmembrane helices (TMs) folding in two symmetrically related six-helix bundle domains (NTD and CTD). Substrate coordination is mediated mostly by residues of TM1-5 of the NTD and by TM7 and TM8 of the CTD (orange circles). The positions of three disease mutants which were studied in this work are indicated by pink stars. **c** Representative size-exclusion chromatography (SEC) trace of purified hSLC19A3-wt. The protein elutes as a monomer of about 50 kDa. A fraction of the protein is glycosylated and appears as an extra band at ~60 kDa in SDS-PAGE (inset). The purification of hSLC19A3-wt has been performed >10 times with close to identical outcomes. **d** Thermal shift assay of hSLC19A3 using nanoDSF. The left panel shows the first derivative of the melting curve, measured as ratio of the fluorescence recorded at 350 nm (F350) and at 330 nm (F330), in the absence and presence of thiamine. Thiamine induces a strong stabilisation of the transporter ( $\Delta T_m = 10.9 \pm 0.3$  °C by 2 mM thiamine). The right panel illustrates the concentration dependent thermal shifts ( $n = 3$ ) with the resulting apparent dissociation constant  $K_{d,app}$  for thiamine (mean  $\pm$  s.d.). **e** Cryo-EM maps of thiamine-bound hSLC19A3 in its outward-open (left) and inward-open (right) conformation. The NTD (purple) and CTD (cyan) of the transporter, as well as the fiducial nanobodies (grey), are resolved. ChimeraX contour levels: outward-open: 0.876; inward-open: 0.148. **f** Structure models of the respective conformational states. The density of thiamine is shown in orange. ChimeraX contour levels: outward-open: 0.547; inward-open: 0.148. The N-terminus, C-terminus and ICL3 are expected to be structurally disordered and could consequently not be resolved.” Figure adapted from Gabriel et al., 2024.

### **3.3 Structural insights into the transport cycle of hSLC19A3**

#### **Structurally resolvable states of hSLC19A3**

Using the three hSLC19A3-specific nanobodies Nb3.3, Nb3.4, and Nb3.7, we could resolve two conformational states of the transporter (Gabriel et al., 2024, Supplementary Fig. 6-10, Supplementary Table 1). While Nb3.3 and Nb3.7 bind to the inward-open state of hSLC19A3, Nb3.4 stabilises an outward-open conformation (Fig. 13, Supplementary Fig. 6-10). This allowed us to investigate the structural basis of the transport cycle of hSLC19A3, including its molecular gates and the recognition of its substrate thiamine.

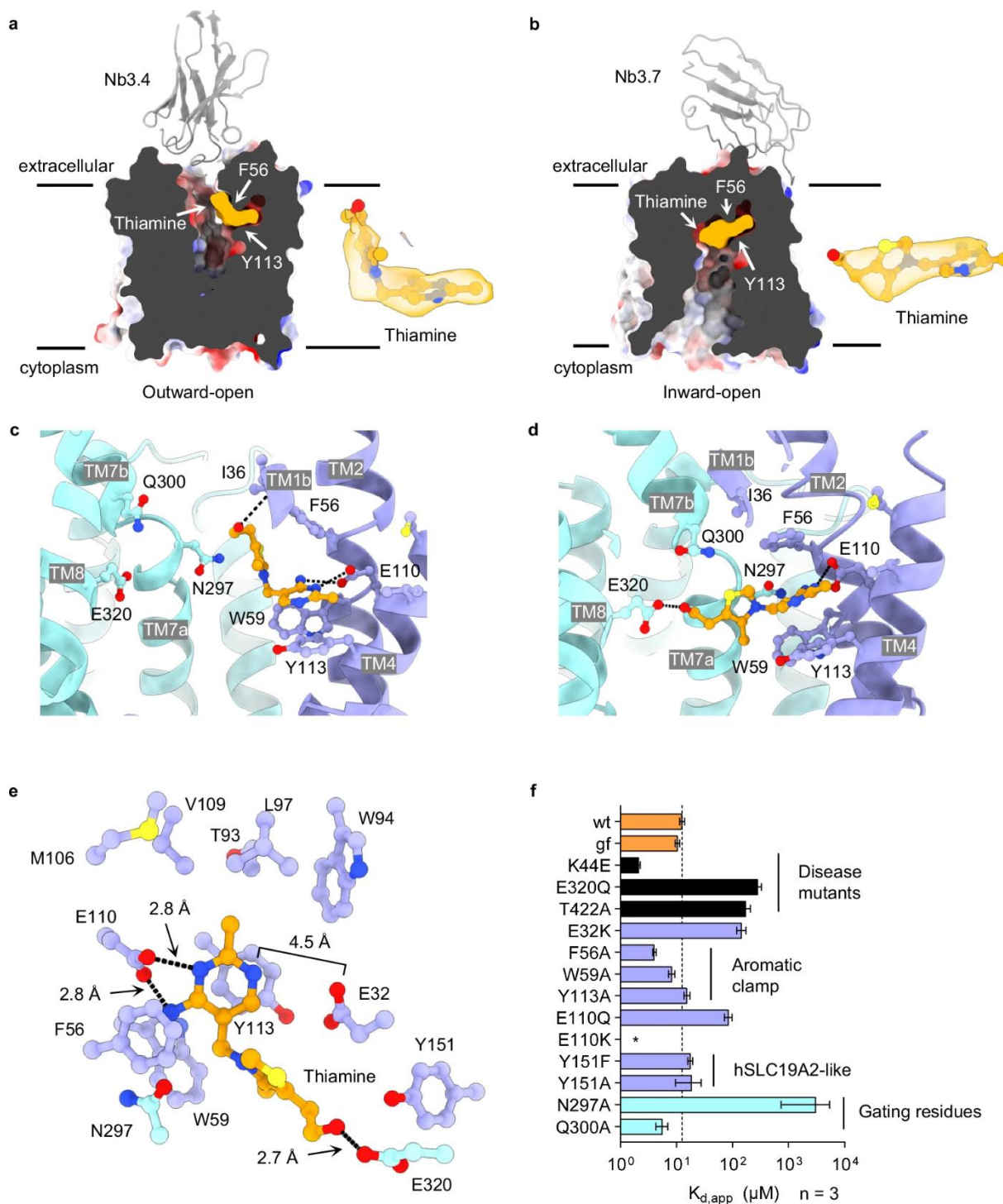
#### **Molecular gates of hSLC19A3**

The cryo-EM structures of hSLC19A3 in complex with Nb3.3 and Nb3.7 allowed the detailed structural analysis of the inward-open state of the transporter. The first observation was that the structure of hSLC19A3 was identical whether bound to Nb3.3 or Nb3.7 (Supplementary Fig. 11). Thus, we assume that the nanobodies do not perturb the native structure of the transporter. In the inward-open state, the extracellular gate of the carrier protein is closed, sealing the substrate binding site towards the extracellular space (Supplementary Fig. 17). This gate comprises at its core the hydrophobic interaction between Ile36 and Ile301, as well as the polar interactions between Gln300 and the backbone of Pro33, as well as Asn297 and the backbone of Phe56 (Gabriel et al., 2024). It is these gate interactions, that are specifically disrupted by Nb3.4, which inserts its CDR3 to engage and separate both the hydrophobic leucine residues, as well as Gln300 and Pro33 of the transporter. By breaking the extracellular gate, Nb3.4 stabilises the outward-open state of hSLC19A3. Upon transitioning from its inward- to its outward open conformation, the two six-helix bundles of the transporter rearrange relative to each other in a rocker-switch-like manner (Drew et al., 2021). I.e. there are hardly intra-domain structural changes, but the domains as entire entities perform a rocking motion to open the extracellular and close the cytoplasmic gate (Supplementary Fig. 11). This cytoplasmic gate has, like its extracellular counterpart, a hydrophobic interaction at its core (between Tyr128, Ala395, and Tyr403), which is supported by polar contacts between Asp75 and the backbone of Ala404 and Ala405, and Gln137 and the backbone of Lys338, respectively (Supplementary Fig. 17).

#### **Binding of thiamine to the outward- and inward open states of hSLC19A3**

In the outward-open state, thiamine inserts its aminopyrimidine ring into a binding pocket of the NTD of hSLC19A3 (Fig. 14, Gabriel et al., 2024). This pocket is lined by three aromatic residues, Phe56, Trp59, and Tyr113, which form a modality we termed aromatic clamp. Deeper into this pocket, Glu110 provides an essential polar contact point for the amine and the ring nitrogen of thiamine (Supplementary Fig. 1 & 5). The rear end of the pocket is a hydrophobic cavity, formed by the side chains of Trp94,

Thr93, Leu97, Met106, and Val109. They accommodate the methyl group of the aminopyrimidine ring of thiamine (Fig. 14). Furthermore, Glu32 is in spatial proximity to the aminopyrimidine ring and likely contributes to the coordination of thiamine in hSLC19A3. In the outward-open state of hSLC19A3, the thiazole ring and the hydroxyethyl tail of thiamine adapt a kinked conformation relative to the aminopyrimidine ring, reaching towards the extracellular space. In this, the thiazole ring is in spatial proximity to and likely engaging in  $\pi$ - $\pi$  stacking interactions with Phe56, while the hydroxyethyl tail of the vitamin forms a polar contact with the backbone of Ile36.



**Fig. 14: “Thiamine binding site of hSLC19A3. a** Localisation of the thiamine binding site of hSLC19A3 in the outward- and **(b)**, inward-open conformation of the transporter. The transporter is shown in electrostatic surface representation. Density for thiamine (orange) is shown separately (ChimeraX contour levels: 0.547 and 0.148, Q-scores: 0.45 and 0.57). **c** Residues involved in thiamine coordination in the outward- and **(d)**, inward-open conformation of hSLC19A3, shown in the membrane plane. **e** Extracellular view of hSLC19A3 highlighting the thiamine-coordinating residues in the substrate binding site of the transporter. Residues of the NTD are coloured in purple, residues of the CTD in cyan, and thiamine in orange. Black dashes indicate hydrogen bonds (cut-off at 3.0 Å) **f** Thiamine binding affinities (apparent dissociation constants  $K_{d,app}$ ) of single point mutants of hSLC19A3. The corresponding titration curves were measured in three independent experiments ( $n = 3$ , see Supplementary Fig. 4 & 5). The bar plot shows the mean  $K_{d,app}$  and its standard deviation derived through a global fit of the measured data to Eq. (4), as described in the methods. For the purification of the respective mutants, see Supplementary Fig. 20.” Figure adapted from Gabriel et al., 2024.

When hSLC19A3 transitions to the inward-open state, the aminopyrimidine ring stays coordinated in the NTD of the transporter, whereas the tail of thiamine rotates about 90°, detaching from its interaction with the NTD and reaching over to the CTD, where the hydroxy moiety engages in a strong hydrogen bond with Glu320 (Fig. 14). Thiamine forms two additional polar contacts in the inward-open state of hSLC19A3, namely through its amine group with Asn297, and through its hydroxy moiety with Glu320 (Supplementary Fig. 1). I therefore speculate that the thiamine-affinity of hSLC19A3 could be higher in its inward-, compared to its outward-open state.

### Genetic diseases caused by mutations in hSLC19A3

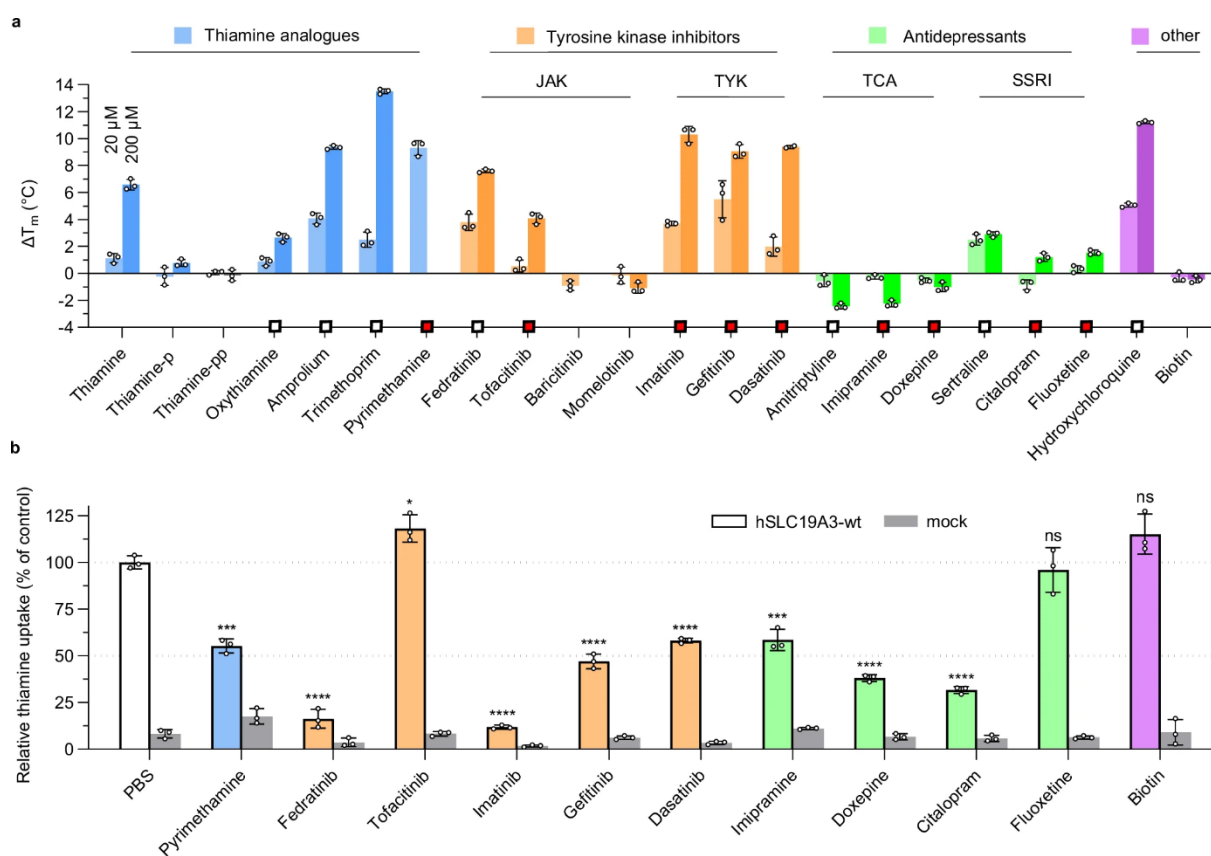
Loss-of-function mutations of hSLC19A3 lead to biotin- and thiamine responsive basal ganglia disease (BTBGD) and Wernicke’s-like encephalopathy (WLE) in humans (Kono et al., 2009, Aburezq et al., 2023). To date, more than thirty single point mutations, along with an increasing number of deletions and frameshift mutations, have been identified as the genetic cause of these neurometabolic disorders (Gabriel et al., 2024, Supplementary Fig. 18). Both BTBGD and WLE present with severe neurological symptoms and are inherited in an autosomal recessive manner. The median age of onset for symptoms ranges from three months to four years after birth, depending on the respective mutations in the SLC19A3 gene. Current treatment recommendations involve a combination of high-dose thiamine and biotin; however, over 30% of patients continue to experience moderate to severe disease effects, and approximately 5% may die despite treatment (Wang et al., 2021). To further elucidate the molecular mechanisms underlying these conditions, we investigated the effects of one of the most prevalent point mutations associated with BTBGD, SLC19A3-Thr422Ala (Gabriel et al., 2024). This mutation is predominantly reported in Saudi Arabia, where the heterozygous carrier frequency is about 1 in 500 (Alfadhel et al., 2019). Patients commonly exhibit symptoms such as subacute encephalopathy, confusion, convulsions, dysarthria, and dystonia. At the cellular level, the SLC19A3-Thr422Ala mutant localises normally to the apical plasma membrane in Caco-2 and MDCK cells and can be successfully expressed and purified *in vitro*. However, its thiamine uptake activity is significantly compromised. Our data indicate that the Thr422Ala mutation results in a greater than 10-fold reduction in thiamine binding affinity, decreasing from  $12.7 \pm 1.2 \mu\text{M}$  to  $172 \pm 35 \mu\text{M}$  (Fig. 14, Supplementary Fig. 19). The Thr422 residue, located on transmembrane helix 11 (TM11), is spatially positioned in the outer region of the

transporter but is connected to the core of hSLC19A3 through hydrogen bonding with Gln294 on TM7. This segment of TM7 plays a critical role in forming the substrate binding site and the extracellular gate with Asn297, Gln300, and Ile301. Structural alterations in this region may account for the diminished affinity and transport activity associated with the Thr422Ala mutation. The cryo-EM structures obtained in this study can also provide insights into the molecular mechanisms of other BTBGD mutations. Mutations such as Trp59Arg, Trp94Arg, Tyr113His, and Glu320Lys are likely to interfere with thiamine binding by introducing a positive charge into the substrate binding pocket of hSLC19A3. Binding studies of the Glu32Lys mutant suggest that the introduction of a positive charge in this area can lead to a significant loss of thiamine affinity, exceeding 10-fold. This reduction in binding affinity likely renders the transporter ineffective at physiological thiamine concentrations, which may explain the pathogenicity of these mutations.

### **3.4 Exploration of drug interactions of hSLC19A3**

Previous research has highlighted a wide range of FDA-approved drugs that inhibit thiamine uptake through hSLC19A3 (Giacomini et al., 2017, Vora et al., 2020). This finding has clinical relevance, especially considering that many of these medications are prescribed as long-term treatments. Among the known thiamine uptake inhibitors (TUIs) are antidepressants like sertraline (Zoloft® and others) and amitriptyline, the antiparasitic hydroxychloroquine, and the Janus kinase (JAK) inhibitor fedratinib (Inrebic®). Inhibition of SLC19A3 can result in systemic and tissue-specific thiamine deficiencies, which may not be detectable through standard serum thiamine level diagnostics. Evidence from BTBGD-causing mutations and studies involving SLC19A3 knockout mice indicate that such deficiencies can lead to the loss of critical cell populations in vital organs, particularly in the brain (Schänzer et al., 2014, Wen et al., 2023). We conducted thermal shift assays with recombinantly expressed and purified hSLC19A3 to investigate its interactions with six known TUIs and eleven pharmacologically similar drugs (Fig. 15, Gabriel et al., 2024). The results confirmed the physical binding of the known inhibitors to hSLC19A3 and revealed nine novel drug interactions, seven of which also showed SLC19A3-mediated thiamine uptake inhibition in cells (Fig. 15). Initially, we tested several drugs that are structurally related to thiamine. The known inhibitors amprolium and trimethoprim (Vora et al., 2020) stabilised the transporter in a concentration-dependent manner, with the newly assessed antiparasitic pyrimethamine showing an even stronger effect (Fig. 15). This compound induced a thermal stabilisation of  $9.3 \pm 0.5$  °C and reduced thiamine uptake through hSLC19A3 by approximately 50% in cellular assays. We also examined representative JAK inhibitors. Fedratinib, a potent hSLC19A3 inhibitor, demonstrated high-affinity binding in thermal shift assays ( $K_{d,app} = 1.0 \pm 0.2$  μM, Supplementary Fig. 21) and almost completely shut down cellular thiamine uptake (Fig. 15). We assessed three additional JAK inhibitors: tofacitinib (Xeljanz®), baricitinib (Olumiant®), and momelotinib (Ojjaara®). Tofacitinib stabilised hSLC19A3 but did not inhibit thiamine transport.

Meanwhile, baricitinib and momelotinib did not stabilise the hSLC19A3 at all, suggesting that these compounds do not interact with the transporter (Gabriel et al., 2024). These findings are consistent with previous studies on JAK inhibitors (Giacomini et al., 2017).



**Fig. 15: “Drug interactions of hSLC19A3. a** Thermal shift screen of known and potential thiamine uptake inhibitors (TUIs). Thermal shift assays were performed in three independent experiments ( $n = 3$ ) in the presence of 20  $\mu$ M and 200  $\mu$ M of the respective compound and the stabilisation effect compared to the apo stability of hSLC19A3. The melting temperature was determined as the inflection point of the thermal unfolding curve of the protein (Fig. 13d). Wherever possible, the ratio of the fluorescence at 350 nm and 330 nm (F350/F330) was selected as a readout for the thermal unfolding curve. The bar plot shows the mean  $\pm$  s.d. of the measured thermal shifts. Due to intrinsic fluorescence of the following compounds, the fluorescence trace at 350 nm (F350) only was used as a readout: oxythiamine, trimethoprim, pyrimethamine, tofacitinib, imatinib, gefitinib, dasatinib, sertraline, fluoxetine and hydroxychloroquine. The two compounds, pyrimethamine and baricitinib, interfere at 200  $\mu$ M too strongly with the fluorescence detection to provide an unambiguous readout and were omitted in this dataset. The squares indicate identified TUIs. White squares indicate already known TUIs, and red squares highlight hSLC19A3-inderactions identified in this study. For affinity measurements of individual compounds, see Supplementary Fig. 21. **b** Cellular thiamine-uptake and its inhibition by 200  $\mu$ M of selected drugs. Shown is the relative uptake of deuterated thiamine (thiamine-d3, normalised to PBS-only control) by Expi293FTM cells overexpressing hSLC19A3-wt (indicated in colour), compared with mock transfected cells (indicated in grey). Three independent measurements were performed ( $n = 3$ ). The data are shown as mean  $\pm$  s.d. (ns:  $p \geq 0.05$ , \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ , \*\*\*\* $p \leq 0.0001$ , Student’s t test, two-tailed).” Figure adapted from Gabriel et al., 2024.

We also tested tyrosine kinase inhibitors, including imatinib (Gleevec®), gefitinib (Iressa®), and dasatinib (Sprycel®). They strongly stabilised hSLC19A3 with apparent binding affinities of  $4.4 \pm 0.5$   $\mu$ M,  $116 \pm 36$  nM, and  $7.5 \pm 1.5$   $\mu$ M, respectively (Fig. 15). All these non-JAK tyrosine kinase inhibitors disrupted thiamine uptake, with imatinib showing an inhibitory effect comparable to that of fedratinib. Chemically, fedratinib and the tyrosine kinase inhibitors share a central aminopyrimidine moiety,

potentially significant for their interactions with hSLC19A3, as suggested by prior research (Giacomini et al., 2017).

All tested tricyclic antidepressants (TCAs), including the known TUI amitriptyline and the newly evaluated imipramine and doxepin, interacted with hSLC19A3 and inhibited its thiamine transport activity (Fig. 15). Notably, TCAs destabilised the transporter in a concentration-dependent manner, unlike compounds from other classes, which typically had a stabilising effect. Among selective serotonin reuptake inhibitors (SSRIs), sertraline stabilised hSLC19A3 at low micromolar concentrations, aligning with previous findings on thiamine uptake inhibition. The newly assessed SSRIs citalopram (Celexa® and others) and fluoxetine (Prozac® and others) also interacted with hSLC19A3, but only citalopram inhibited thiamine uptake. Hydroxychloroquine demonstrated strong stabilisation of hSLC19A3, binding with an apparent affinity in the nanomolar range ( $K_{d,app} = 170 \pm 55$  nM), consistent with previous reports of its potent inhibition of hSLC19A3. In contrast, biotin did not interact with the transporter, which aligns with our and prior cellular uptake data (Fig. 15, Subramanian et al., 2006).

### 3.5 Structural basis of drug interactions of hSLC19A3

The thiamine uptake inhibitors (TUIs), both known and newly identified, exhibit considerable structural diversity (Supplementary Fig. 21). This diversity poses challenges for developing a universal ligand-based pharmacophore model. To better understand how these compounds interact with hSLC19A3, I determined cryo-EM structures of the transporter in complex with three known inhibitors: fedratinib, amprolium, and hydroxychloroquine (Fig. 16, Gabriel et al., 2024). These inhibitors belong to structurally and functionally distinct drug classes. The cryo-EM structures were solved using either Nb3.7 or Nb3.3 as structural fiducials. These nanobodies bind to the transporter in the inward-open conformation. The cryo-EM maps of the transporter-drug complexes were resolved to between 3.1 and 3.7 Å, with core resolutions reaching 2.5-3.2 Å (Supplementary Fig. 22-26, Supplementary Table 1). All analysed inhibitors bind orthosterically within the substrate binding site of the transporter (Fig. 16).

From a structural perspective, the binding of these compounds to hSLC19A3 is influenced by at least three key factors:

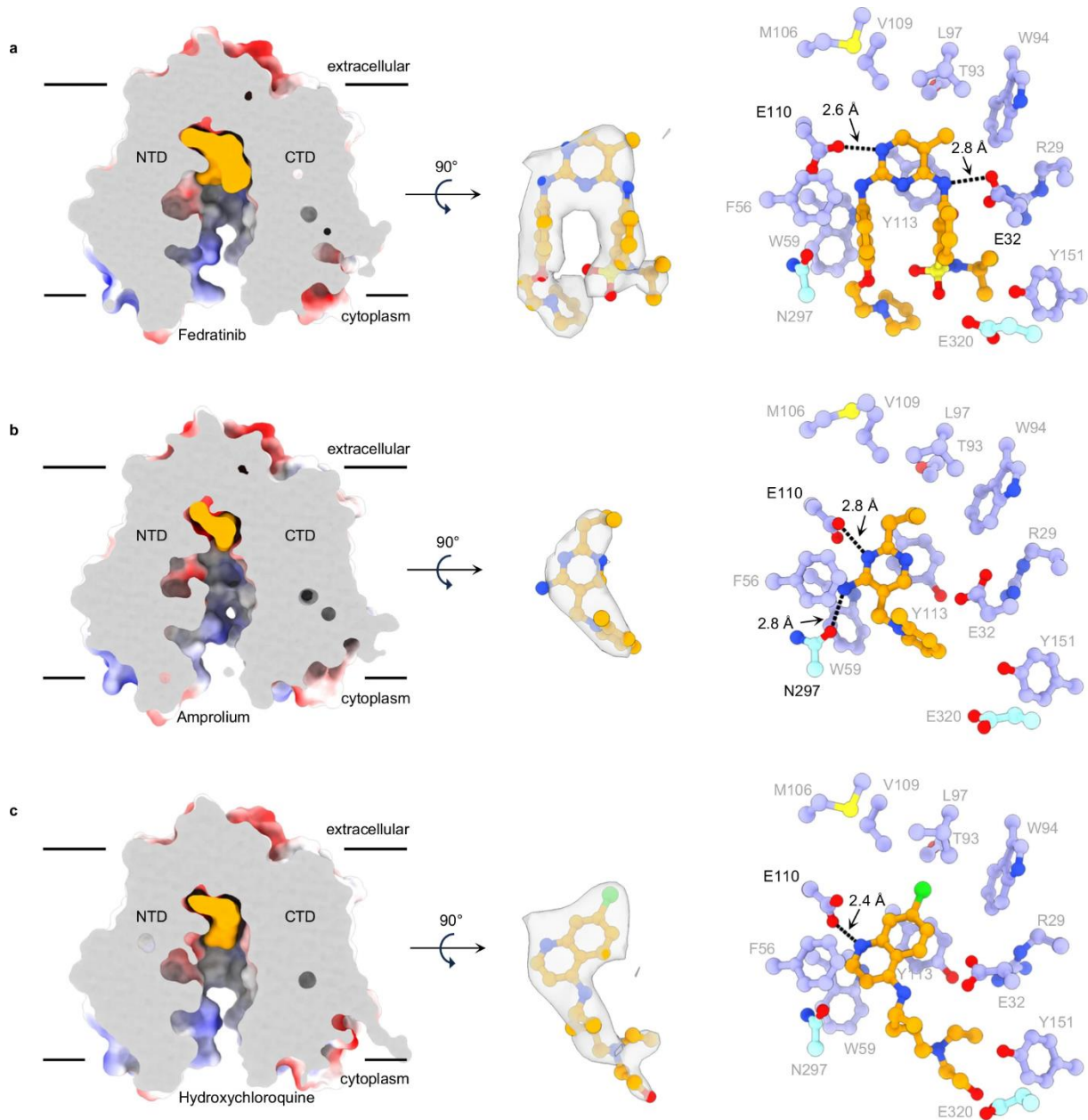
- 1. Intercalation of aromatic ring systems in the aromatic clamp:** Cryo-EM data indicate that the aromatic ring systems of all three inhibitors engage in direct  $\pi$ - $\pi$  stacking interactions with Tyr113, and to a lesser extent with Phe56 and Trp59 (Fig. 16). This appears to be a critical feature for the interaction of TUIs with hSLC19A3, as most known and newly identified ligands contain at least one aromatic ring. This interaction may be particularly vital for the inhibitory



effect of compounds lacking compatible polar contacts, such as the tested antidepressants amitriptyline, imipramine, doxepin, sertraline, and citalopram (Fig. 15).

- 2. Electrostatic compatibility with polar contact points provided by Glu32, Glu110, and Asn297:** Fedratinib, amprolium, and hydroxychloroquine feature hydrogen bond donors and acceptors that align well with the polar residues in the substrate binding site (Fig. 16). All three inhibitors are within hydrogen bonding distance to Glu110, which may be protonated, allowing it to function as both a hydrogen bond donor and acceptor. A protonation of Glu110 is supported by the interactions of fedratinib and amprolium with this residue (Fig. 16). Asn297 also plays a significant role, with its carbonyl oxygen contributing to the electrostatically negative surface of the transporter and being within hydrogen bonding distance of amprolium's primary amine (2.8 Å; Fig. 16). While Glu320 is crucial for high-affinity thiamine binding (Fig. 14), it does not appear to interact with any of the inhibitors (Fig. 16). In contrast, Glu32 is within hydrogen bond distance of a secondary amine in fedratinib (2.8 Å; Fig. 16).
- 3. Insertion of a lipophilic moiety into the hydrophobic pocket formed by Thr93, Trp94, Leu97, Met106, and Val109:** The structurally resolved inhibitors all engage in this interaction, extending an apolar substituent into the hydrophobic pocket (Fig. 16). Specifically, fedratinib protrudes a methyl group, amprolium a propyl group, and hydroxychloroquine a chlorine atom into the lipophilic cavity. This interaction likely enhances the binding affinity of the respective ligands for hSLC19A3. A similar methyl group is also present in the structures of the newly identified inhibitors imatinib and dasatinib (Supplementary Fig. 21).





**Fig. 16:** “Cryo-EM structures of hSLC19A3 in complex with high-affinity inhibitors. a–c Bound inhibitors (illustrated in orange) were observed to occupy the same substrate binding site and engage with similar residues as thiamine. Shown on the left are cross sections of the electrostatic surface representations of the ligand-bound structures, with the ligand coloured in orange. The densities for the different compounds are shown in the centre of the respective panels (ChimeraX contour levels/Q-scores for fedratinib: 0.7/0.47, amprolium: 0.15/0.53, hydroxychloroquine: 0.47/0.75); density within 2.5 Å of the atom coordinates is shown. On the right, the coordination of the compounds in the substrate binding site is depicted. Residues of the NTD are shown in purple, the ones of the CTD in cyan and the compound in orange. Black dashes indicate hydrogen bonds (cut-off at 3.0 Å).” Figure adapted from Gabriel et al., 2024.

## 3.6 Structure-based discovery of thiamine uptake inhibitors

### Computational docking of small molecules in the substrate binding site of hSLC19A3

I resolved the cryo-EM structures of hSLC19A3 binding to its substrate thiamine and different thiamine uptake inhibitors (TUIs). With these structures, we now retrospectively understand, how certain inhibitors of hSLC19A3 interact with the transporter. In a next step, I wanted to ask the question, whether these experimental structural data could be used to identify previously unknown thiamine uptake inhibitors. In order to address this question, I teamed up with an expert in computational docking, Björn Windshügel from the Fraunhofer Institute for Translational Medicine and Pharmacology (ITMP). He performed the computational steps described in the following.

We decided to make use of the cryo-EM structures of the inward-open hSLC19A3 in complex thiamine, fedratinib, amprolium, and hydroxychloroquine to perform docking of a selection of medical compounds. The protein structures were aligned and prepared for docking using the software Molecular Operating Environment (MOE v.2022.02, Chemical Computing Group Inc., Montreal, Canada). For this preparation, nanobodies and small molecule ligands were removed. Subsequently, the protein structures were protonated, and their hydrogen bond networks optimised using Protonate3D (MOE). The docking was performed using the software GOLD (v.2023.3.0, Cambridge Crystallographic Data Centre, Cambridge, UK) to screen the binding of 583 compounds of an in-house curated Enzo Life Sciences drug repurposing library. The compounds were docked into the substrate binding site of hSLC19A3, by confining them to a 12 Å sphere around the C4 atom of thiamine (PDB ID: 8S61). For each compound, 50 docking rounds were carried out. The docking poses were scored using the ChemPLP scoring function of GOLD and analysed visually in MOE. After the docking was complete, 34 hits were selected for experimental validation, based on the ChemPLP score and visual inspection (Table 6).

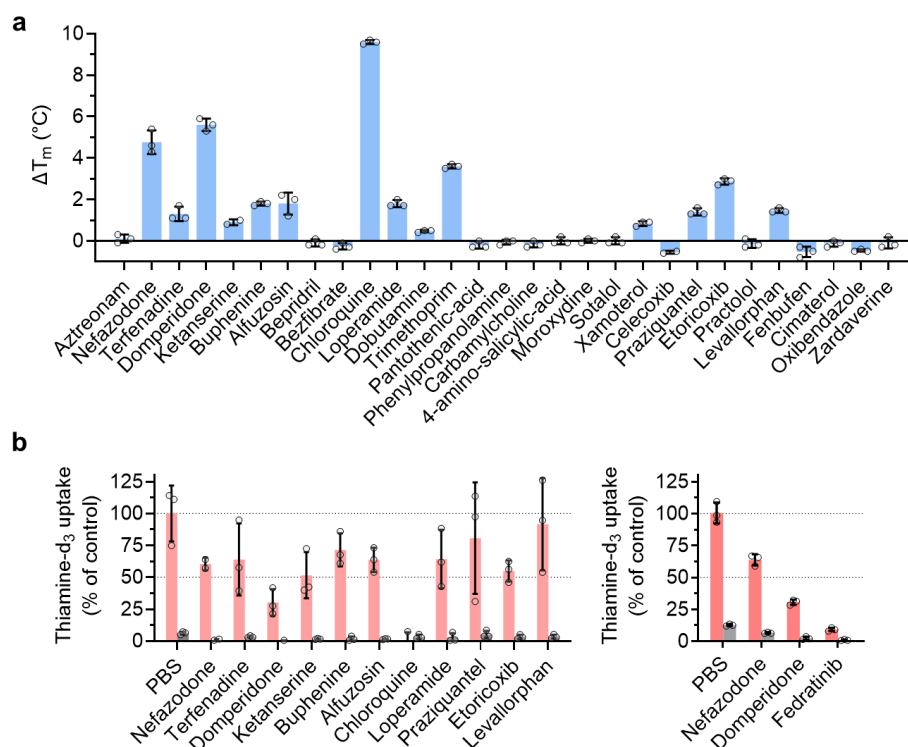
**Table 6: Overview of the selected hit compounds from the virtual screen.** (-) not determined. R and S indicate the respective stereoisomer. The compounds in italics were not compatible with nanoDSF.

Compound	ChemPLP score	$\approx \Delta T_m$ (°C, mean)	$\approx$ Thiamine-d <sub>3</sub> uptake inhibition (% control, mean)
Nefazodone	98.58	4.8	39.5
Terfenadine (R)	94.26	1.3	36
Domperidone	88.23	5.6	69.7
Alfuzosin (S)	84.65	1.8	36.3
Bepridril (R)	84.58	-0.1	-
Ketanserine	83.6	0.9	48.4
Bezafibrate	82.38	-0.3	-
Chloroquine (S)	82.13	9.6	100
Loperamide	81.58	1.8	35.7
Buphenine (R,R,S)	81.1	1.8	28.5
Dobutamine (S)	79.5	0.5	-
Xamoterol (R)	76.76	0.8	-
Celecoxib	74.35	-0.6	-
Aztreonam	73.93	0.1	-
Praziquantel (R)	72.24	1.4	19.3
Etoricoxib	67.59	2.9	45
Practolol	65.71	-0.1	-
Levallorphan	64.41	1.5	8.4
Fenbufen	63.18	-0.5	-
Cimaterol (R)	60.16	-0.1	-
Sotalol (S)	59.67	0	-
Trimethoprim	59.25	3.6	-
Oxibendazole	56.24	-0.5	-
Zardaverine	55.89	-0.1	-
Pantothenic-acid	53.87	-0.2	-
Moroxydine	51.93	0	-
Phenylpropanolamine	50.31	-0.1	-
Carbamylcholine	44.72	-0.2	-
4-amino-salicylic-acid	41.28	0	-
<i>Fluphenazine</i>	<i>91.15</i>	-	-
<i>Naftopidil (S)</i>	<i>86.94</i>	-	-
<i>Sulfasalazine</i>	<i>70.29</i>	-	-
<i>Famciclovir</i>	<i>67.14</i>	-	-
<i>Fluvastatin (S, R)</i>	<i>94.41</i>	-	-

### Experimental validation of hSLC19A3 ligands

I subsequently tested the 34 selected hit compounds for their interaction with purified hSLC19A3 using thermal shift assays (nanoDSF). Of the 34 *in silico* hits, 16 showed a significant effect on the thermostability of hSLC19A3 (Fig. 17). The strongest effects ( $\Delta T_m > +2^\circ\text{C}$ , at 200  $\mu\text{M}$ ) were seen for nefazodone, domperidone, chloroquine, trimethoprim, and etoricoxib (Fig. 17). Chloroquine and

trimethoprim were already known to be TUIs (Giacomini et al., 2017). The fact that these compounds were picked up again through the docking procedure and the thermostabilisation screens confirms our approach. Five compounds, namely fluphenazine, naftopidil, sulfasalazine, famciclovir, and fluvastatin, interfered with the nanoDSF measurement at 200  $\mu$ M compound concentration. I therefore repeated the measurements for these drugs at 20  $\mu$ M. At this concentration, I could not observe any effect of these five compounds on the thermostability of hSLC19A3 (data not shown).



**Fig. 17: Experimental validation of hSLC19A3 docking screen hits. a** Thermal shift assay (nanoDSF) of 29 nanoDSF-compatible hit compounds from the structure-based computational docking screen. **b** Mass spectrometry-based cellular thiamine-d<sub>3</sub> uptake assays. Left panel: initial screen using 200,000 cells and 20  $\mu$ M of the respective compounds. Right panel: Repetition of the assay for nefazodone and domperidone with 2,000,000 cells and at 200  $\mu$ M compound concentration.

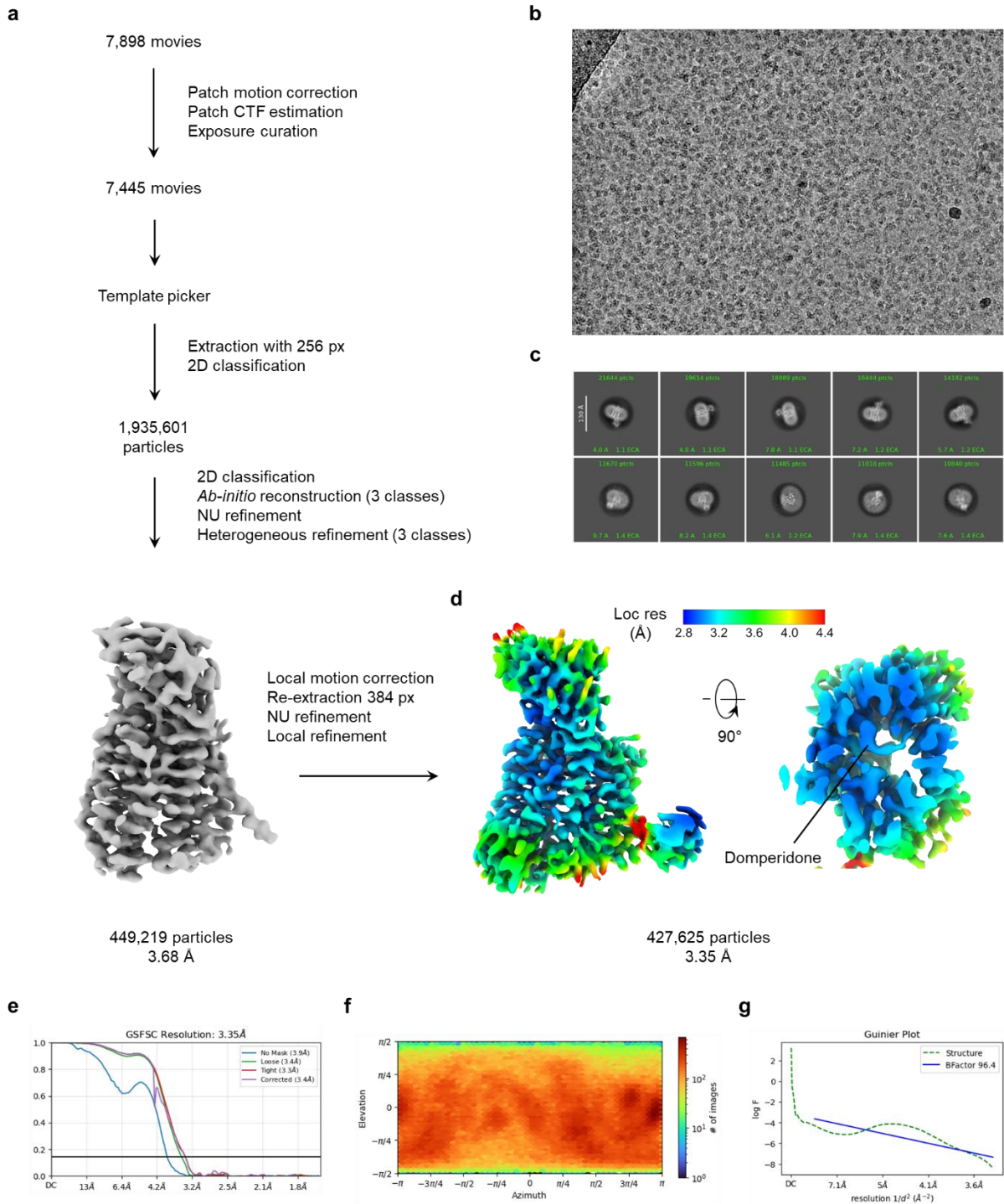
### Testing of the inhibitory potential of new hSLC19A3 ligands

I postulate that compounds inducing a thermostabilisation of more than +1  $^{\circ}$ C in hSLC19A3 can be considered hSLC19A3 ligands. I selected the drugs fulfilling this criterion (N = 11) to perform cell-based thiamine uptake inhibition assays (analogous to the inhibition assays described in Section 3.4). Due to the limited availability of these compounds, I could perform this assay only with a lower cell number ( $0.2 \times 10^6$ ) and at 20  $\mu$ M inhibitor concentration. The corresponding results show consequently relatively large standard deviations (Fig. 17). However, this assay indicated that nine out of the eleven tested hSLC19A3 ligands interfere with thiamine uptake (Fig. 17). The strongest effects (> 40% inhibition) were induced by nefazodone, domperidone, ketanserin, and etoricoxib. Chloroquine was included as a positive control in this assay and completely suppressed thiamine uptake. For validation

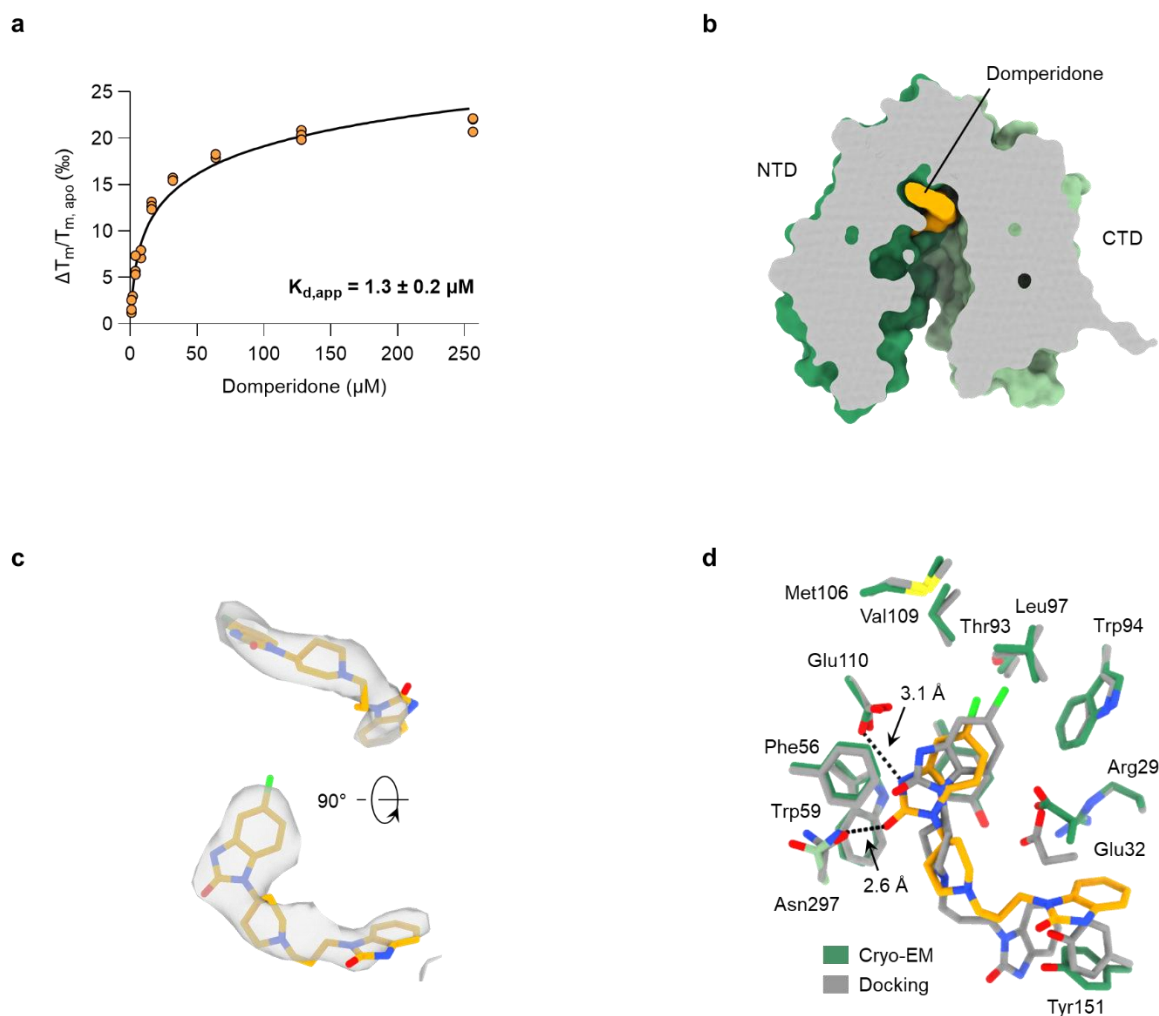
purposes, I decided to repeat this assay for nefazodone and domperidone. I used compounds from a different vendor and performed the assay under optimised conditions ( $2 \times 10^6$  cells,  $200 \mu\text{M}$  inhibitor concentration). As shown in Fig. 17, this assay could confirm the inhibition of hSLC19A3-mediated thiamine uptake by nefazodone ( $36 \pm 4\%$ ) and domperidone ( $69 \pm 2\%$ ), and the positive control fedratinib ( $91 \pm 1\%$ ).

### **Cryo-EM structure of hSLC19A3 in complex with domperidone**

In order to understand the structural basis of the interaction of these newly identified inhibitors, I decided to determine the cryo-EM structure of hSLC19A3 in its inward-open state in complex with domperidone, one of our best hits from the inhibitor discovery. The binding analysis by nanoDSF yielded an apparent  $K_d$  of  $1.3 \pm 0.2 \mu\text{M}$  of domperidone for hSLC19A3. The cryo-EM sample was subsequently prepared with a saturating concentration of  $250 \mu\text{M}$  of the compound. Cryo-EM data of the ternary hSLC19A3-wt:Nb3.3:domperidone complex were collected on a Titan Krios G3 (K3 camera) and the data processed as described in the corresponding methods Sections 2.9 & 2.10 (Fig. 18). The final nominal map resolution reached  $3.35 \text{ \AA}$  and the structure model of the transporter and the nanobody could be fitted well into the 3D reconstruction. The substrate binding site of hSLC19A3 showed clear and separate density for a bound ligand (local resolution  $\approx 2.8 \text{ \AA}$ ). The molecular structure of domperidone showed good agreement with this density and also matched well with coordinating residues of the transporter (Fig. 19). The chloro-benzimidazole moiety of domperidone is inserting into the NTD of hSLC19A3. The chlorine atom is accommodated by the hydrophobic pocket created by Thr93, Trp94 and Leu97 on TM3, and Met106 and Val109 on TM4. As an aromatic ring system, the chloro-benzimidazole is engaging in  $\pi$ - $\pi$  stacking with the aromatic clamp, mainly with Tyr113 and to a lesser extent with Phe56 and Trp59. The secondary ring amine of the chloro-benzimidazole is in hydrogen bond distance to the side chain of Glu110 ( $3.1 \text{ \AA}$ ). At the same time, the carbonyl oxygen of the chloro-benzimidazole is in hydrogen bond distance with Asn297 ( $2.6 \text{ \AA}$ ). The other major contact of domperidone with hSLC19A3 is formed through its second aromatic moiety, the chlorine-free benzimidazole. The density for this part of the molecule is less well-defined. It can, however, be clearly located as inserting into a pocket between TM1b, TM5, and TM8. The benzimidazole is here primarily coordinated by a  $\pi$ - $\pi$  stacking interaction with Tyr151 (Fig. 19). Overall, the domperidone molecule is wrapping around TM1b of the NTD of hSLC19A3. The two benzimidazole groups of domperidone are anchored to the protein mainly by  $\pi$ - $\pi$  stacking interactions. The linker connecting the two aromatic moieties provides the necessary flexibility for the drug molecule to fit in the substrate binding site of hSLC19A3. Comparing the coordination of domperidone to other thiamine uptake inhibitors (TUI) reveals interesting similarities and one important difference. The chloro-benzimidazole moiety of domperidone engages the transporter analogous fedratinib, amprolium, and particularly hydroxychloroquine. Its second benzimidazole, however, forms an interaction network, which has not been observed for any other ligand of hSLC19A3 so far (Fig. 19).



**Fig.18: Cryo-EM data processing for hSLC19A3-wt in complex with Nb3.3 and domperidone.** **a** Data processing workflow of the hSLC19A3-wt:Nb3.3:domperidone cryo-EM dataset. Template picking was performed using 2D reference generated from a 3D refined reconstruction from a subset of 459 blob-picked micrographs. ChimeraX contour level of the non-uniform refinement density map: 0.16. **b** Representative micrograph acquired at a defocus of 1.2  $\mu\text{m}$ . **c** 2D class averages. **d** Final cryo-EM density map of the hSLC19A3 complex coloured by local resolution (ChimeraX contour level 0.74). **e** Gold-standard Fourier shell correlation (GSFSC) curves for two half maps. FSC threshold = 0.143. **f** Angular distribution of particles in the final 3D local refinement. **g** Guinier plot for B factor estimation.



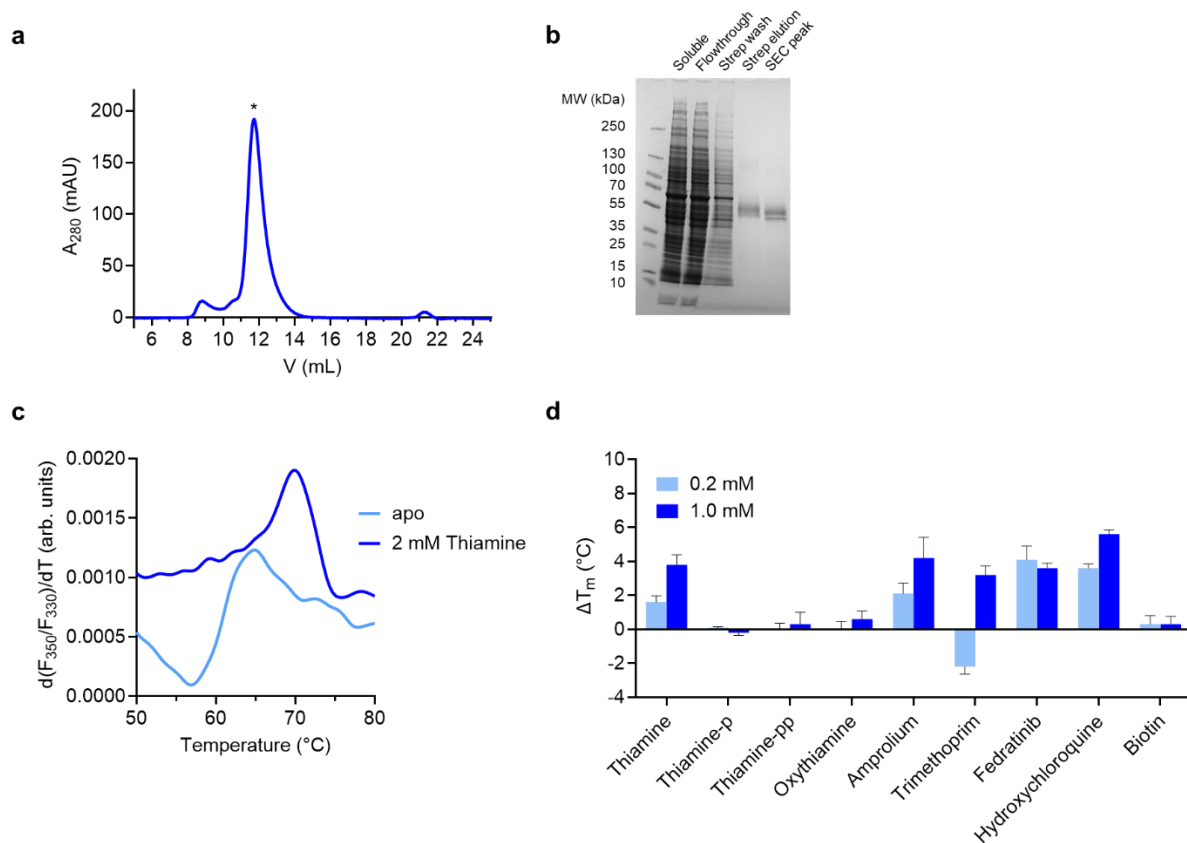
**Fig. 19: Cryo-EM structure of domperidone-bound hSLC19A3.** **a** Thermal shift assay (nanoDSF) for the determination of the apparent binding affinity of domperidone for hSLC19A3-wt. **b** Slice through a surface representation of the hSLC19A3-dimperidone complex. The NTD and CTD of the transporter are coloured in dark and light green, respectively. Domperidone is highlighted in orange. **c** Cryo-EM density (grey surface) and modelled domperidone structure (ChimeraX contour level: 0.65. **d** Substrate binding site of hSLC19A3. Comparison of the domperidone binding pose between the experimental cryo-EM structure and the computational docking in the thiamine-bound hSLC19A3 (PDB ID: 8S61). The docking predicted the localisation of the chloro-benzimidazole moiety largely correctly. The second benzimidazole, however, engages the transporter in an unpredicted manner, by inserting into a pocket behind Tyr151.

### 3.7 Biophysical and structural work on SLC19A2

#### Expression, purification and biophysical characterisation of the human SLC19A2

In the initial expression and solubilisation screens performed for this work, the murine and rat SLC19A2 showed comparably good yields (Fig. 6, Table 5). Later on, we cloned and purified the human SLC19A2. Analogously to hSLC19A3, hSLC19A2 was fused with a C-terminal c3C-Twin-Strep-tag and could be purified at equally high yields and purity when solubilised in LMNG/CHS (Fig. 20). hSLC19A2 was thermostabilised in the presence of thiamine by up to about 4 °C (Fig. 20).





**Fig. 20: Purification and drug interactions of hSLC19A2.** **a** Size-exclusion chromatography of Strep-tag purified wildtype hSLC19A2 (Superdex S200 10/300 Increase column). **b** SDS-PAGE analysis of the different purification steps of hSLC19A2. **c** nanoDSF measurement of hSLC19A2 in its apo state and in the presence of 2 mM thiamine. **d** Thermal shift assay (nanoDSF) of hSLC19A2 in the presence of different compounds at 200  $\mu$ M and 1 mM, respectively.

## Drug interactions of hSLC19A2

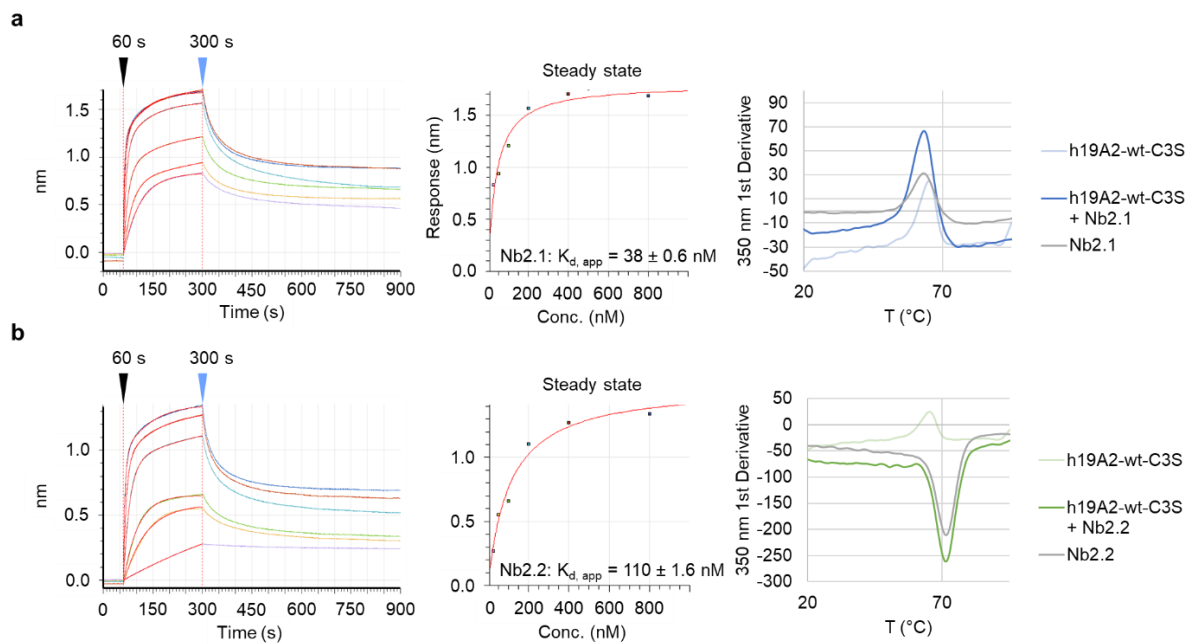
In a next step, we probed the drug interaction profile of hSLC19A2. Thermal shift screens showed that hSLC19A2 shares many drug interactions with hSLC19A3 (Fig. 20). However, the extent of thermal stabilisation by individual compounds respectively differed. This implies differences in the binding affinities of the compounds for the two thiamine transporters. This might have important consequences for the drug interactions of hSLC19A2 and hSLC19A3 *in vivo*.

## Nanobody discovery for hSLC19A2

Analogously to the work on hSLC19A3, I performed a nanobody discovery campaign for hSLC19A2. The antigen for immunisation was in this case the wildtype hSLC19A2 (fully glycosylated, purification tag cleaved with HRV 3C protease). After the immunisation process, the llama serum showed good hSLC19A2-specific HCAb titres. Through phage display selection, I could subsequently isolate two unique anti-hSLC19A2 nanobodies, Nb2.1 and Nb2.2 (Supplementary Table 6). They showed good



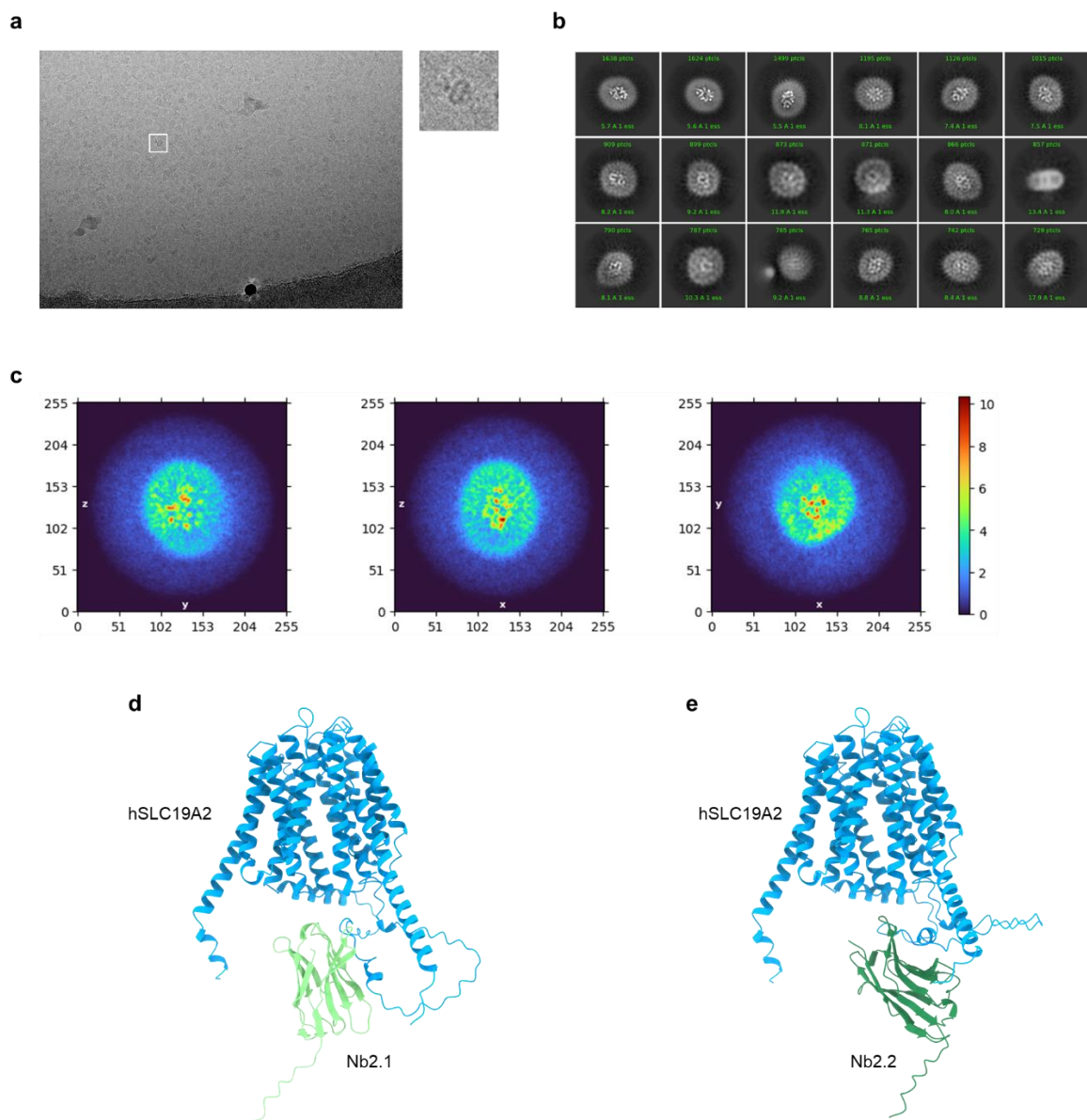
signals and hSLC19A2-specificity in ELISAs and high-affinity binding in BLI. However, we could not observe any thermostabilisation of hSLC19A2 by Nb2.1 or Nb2.2 (Fig. 21).



**Fig. 21: Biophysical characterisation of hSLC19A2-specific nanobodies.** Left panels: Bi-layer interferometry (BLI) measurements of the binding of hSLC19A2 to immobilised nanobodies. hSLC19A2 concentrations: 25-800 nM. Middle panel: Fitting of the measured BLI responses to a steady state binding model for the determination of binding affinities. Right panel: Effect of the nanobodies on the thermal stability of hSLC19A2, measured by nanoDSF. **a** Nb2.1. **b** Nb2.2.

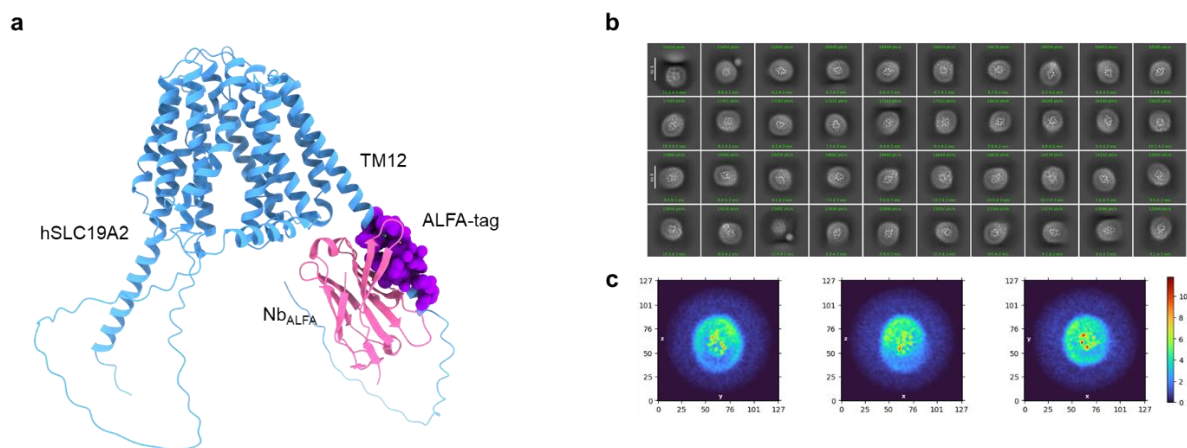
## Cryo-EM of hSLC19A2

Both nanobodies were used to prepare cryo-EM samples with thiamine-bound hSLC19A2. At a first glance, the particle density and distribution on the grids looked good and comparable to the hSLC19A3 preparations. The particles showed preferred orientation. Most of the transporter-nanobody complexes appeared to orient with the micelle parallel to the grid-plane. Furthermore, neither the 2D classes nor the 3D reconstructions reached high resolution (exemplified for Nb2.2 in Fig. 22). Especially on the 2D level, it became apparent that the nanobodies were likely binding to a flexible epitope of the transporter, most likely the disordered N- or C-terminus, or the large ICL3. The nanobodies were consequently not useful for high-resolution structure determination of the transporter. I correlate this with the absence of any thermostabilisation effect of the nanobodies on their target (Fig. 21). Subsequent structure prediction of the hSLC19A2-nanobody complexes using the newly released AlphaFold3 showed that the nanobodies likely bind to sequences in the ICL3 of the transporter (Fig. 22). These predictions would be in line with our observations in BLI, nanoDSF, and cryo-EM.



**Fig. 22: Cryo-EM analysis and AlphaFold3 models of hSLC19A2-nanobody complexes.** **a** Micrograph of hSLC19A2 in complex with Nb2.2 and thiamine. One particle is enlarged to show the visible complex formation between hSLC19A2 and Nb2.2. **b** 2D class averages from the same dataset. **c** Real space slices through a 3D reconstruction from the same cryo-EM dataset. **d** AlphaFold3 model of the hSLC19A2-Nb2.1 complex. **e** AlphaFold3 model of the hSLC19A2-Nb2.2 complex.

In the aftermath of this endeavour, I tried another strategy for providing hSLC19A2 with a structural fiducial. I fused a C-terminal ALFA peptide-tag (Götzke et al., 2019). This short peptide forms a stable  $\alpha$ -helix that is recognised with picomolar affinity by a specifically developed nanobody, Nb<sub>ALFA</sub>. I replaced the most flexible part of TM12 of hSLC19A2 with this tag (Fig. 23), expressed and purified it. This ALFA-tagged construct showed excellent behaviour in terms of yield, purity, stabilisation by thiamine, and binding to Nb<sub>ALFA</sub> (data not shown). In cryo-EM, however, the ternary complex hSLC19A2-ALFA-Nb<sub>ALFA</sub>-thiamine showed preferred orientation (mostly top views), and 2D and 3D alignment suggested a high degree of flexibility between the transporter and the nanobody (Fig. 23). Therefore, this construct was also deemed unfit for high-resolution structure determination.



**Fig. 23: AlphaFold2 model and cryo-EM analysis of the hSLC19A2-NbALFA complex.** **a** AlphaFold2 model of the hSLC19A2-ALFA construct in complex with NbALFA. **b** 2D class averages of a cryo-EM dataset collected on this protein complex. **c** Real space slices through a 3D reconstruction from the same dataset.

## Structure prediction of hSLC19A2

As I was not able to resolve meaningful density maps of hSLC19A2 by cryo-EM, I use AlphaFold predictions of this transporter to discuss its structural features (Supplementary Fig. 14 & 15). In June 2024, a preprint on the cryo-EM structures of hSLC19A2 was published by Qu and colleagues (Qu et al., 2024), which provides a first experimental validation for what I describe in the following. hSLC19A2 and hSLC19A3 share a high degree of sequence identity (53%), with a particular conservation of the transmembrane domains. Consequently, the AlphaFold (AF2) prediction of hSLC19A2 is in good agreement with the experimental cryo-EM structures of the inward-open state of hSLC19A3 (Supplementary Fig. 14). The only notable differences are confined to the poorly conserved and anyways disordered N- and C-terminal regions and ICL3 (Supplementary Fig. 12 & 14). The substrate binding site exhibits a high degree of structural conservation, with only minor variations observed in the form of two specific substitutions, in the following (A) and (B). (A) Phe56 in hSLC19A3 is replaced by Tyr74 in hSLC19A2. (B) Tyr151 in hSLC19A3 is substituted by Phe169 in hSLC19A2. Substitution (A) results in a slight alteration of the aromatic clamp; however, as shown in Supplementary Fig. 14, the replacement of Phe56 with alanine in hSLC19A3 does not impair thiamine binding. Thus, it is reasonable to extrapolate that the substitution of Phe56 with tyrosine has probably no substantial effect on thiamine binding either. Experimental analysis of the second substitution (B) by nanoDSF revealed that replacing Tyr151 in hSLC19A3 with phenylalanine does not significantly affect thiamine affinity of the transporter (Fig. 14). These findings agree with previous studies, which determined very similar transport affinities of hSLC19A2 and hSLC19A3 for thiamine (Yamashiro et al., 2020). As the substrate binding sites are so highly identical on a structural level, it stands to reason that the two transporters have a similar drug interaction profile. This is in line with our thermal stabilisation screens (Fig. 15 & 20) and transport inhibition data on hSLC19A2 reported by Qu and colleagues. There are,

however, differences in the affinities of drugs for hSLC19A2 or hSLC19A3 (Fig. 15 & 20, Qu et al., 2024). This could translate into tissue- or cell type-specific thiamine uptake inhibition, due to the differential expression of the two transporters throughout the human body.

### **3.8 Cryo-EM structure of hTPK1**

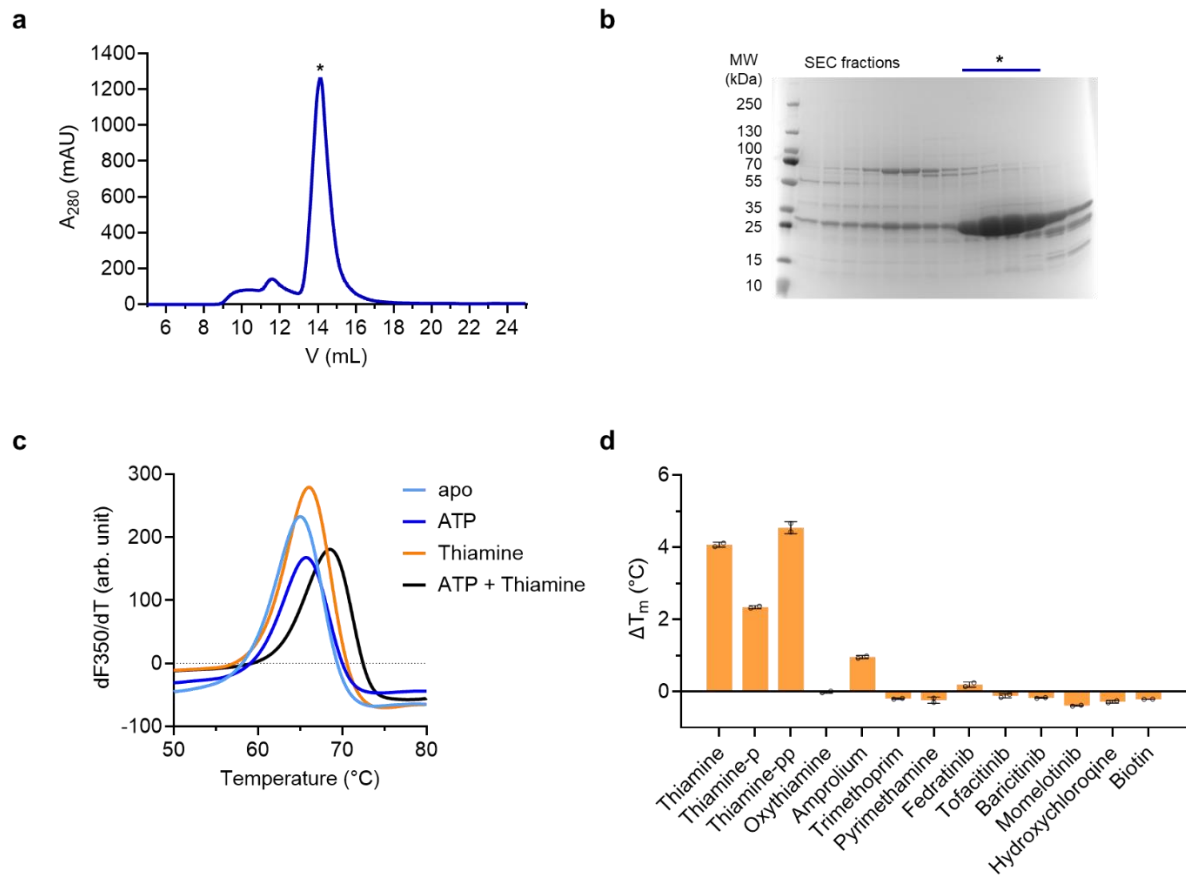
#### **The human thiamine pyrophosphokinase TPK1**

The human TPK1 (hTPK1) is a 27 kDa protein that forms a stable homodimer with a combined molecular mass of 54 kDa. The kinase is expressed and resides in the cytoplasm (Sambon et al., 2022). hTPK1 catalyses the transfer of a pyrophosphate moiety from ATP to thiamine, generating thiamine-pp and AMP (Fig. 3). To perform this phosphoryl transfer, hTPK1 coordinates a catalytically essential  $Mg^{2+}$  ion in its active site (Timm et al., 2001, Sambon et al., 2022). Crystal structures of the enzyme have been solved over the past two decades (see Table 7). I wanted to test, whether the structure of this small human kinase could also be determined using state-of-the-art single particle cryo-EM. Furthermore, we wanted to compare structure models of the same protein determined by cryo-EM and macromolecular crystallography (MX).

#### **Expression, purification, and initial biophysical characterisation of TPK1**

I overexpressed hTPK1 in *E. coli*. For this purpose hTPK1 cDNA (GenScript) was cloned into pSb\_init, a vector originally developed for the periplasmic expression of sybodies in the lab of Markus Seeger at the University of Zurich. In this vector, hTPK1 is fused to an N-terminal PelB sequence to target it to the periplasm of the expression host. This signal peptide is cleaved off by signal peptidases in the bacterial membrane (Lei et al., 1987; Steiner et al., 2006; Singh et al., 2013). Furthermore, pSb\_init provides hTPK1 with a C-terminal Myc-tag, followed by a His<sub>6</sub>-tag for Ni<sup>2+</sup> affinity purification. Interestingly, arabinose-induced overexpression of hTPK1 led to the development of a noticeable ethanol odour of the bacterial cultures over the course of the four-hour expression at 37 °C. This phenomenon was absent in analogous expressions of sybodies from this vector. I assume that the presence of excess amounts of hTPK1 in the periplasm might scavenge free thiamine, thus depriving the bacterial cells of this micronutrient and forcing them to switch from oxidative phosphorylation to fermentation. While I have no proof for this idea, the consequently rationalised addition of 0.1 mM thiamine to the TB growth medium increased the yield of recombinant hTPK1 by about 30-fold (Fig. 24). hTPK1 could be obtained at high purity from periplasmic extraction, followed by Ni<sup>2+</sup> affinity chromatography and size-exclusion chromatography (Fig. 24). Using thermal shift assays (nanoDSF), I could show that the purified hTPK1 is stabilised by thiamine, thiamine-p, thiamine-pp, and the thiamine analogue antibiotic amprolium, as expected (Fig. 24). Other tested compounds, including trimethoprim, pyrimethamine, hydroxychloroquine, biotin, and the JAK inhibitors fedratinib, tofacitinib, baricitinib,

and momelotinib showed no interaction with the kinase (Fig. 24). An important notion at this point is, that the entire purification of hTPK1 from the periplasm of *E. coli* was performed in the absence of thiamine or any thiamine analogues. Theoretically, that should have yielded the enzyme in its apo form. As I will show later in this Section that was, however, not the case.



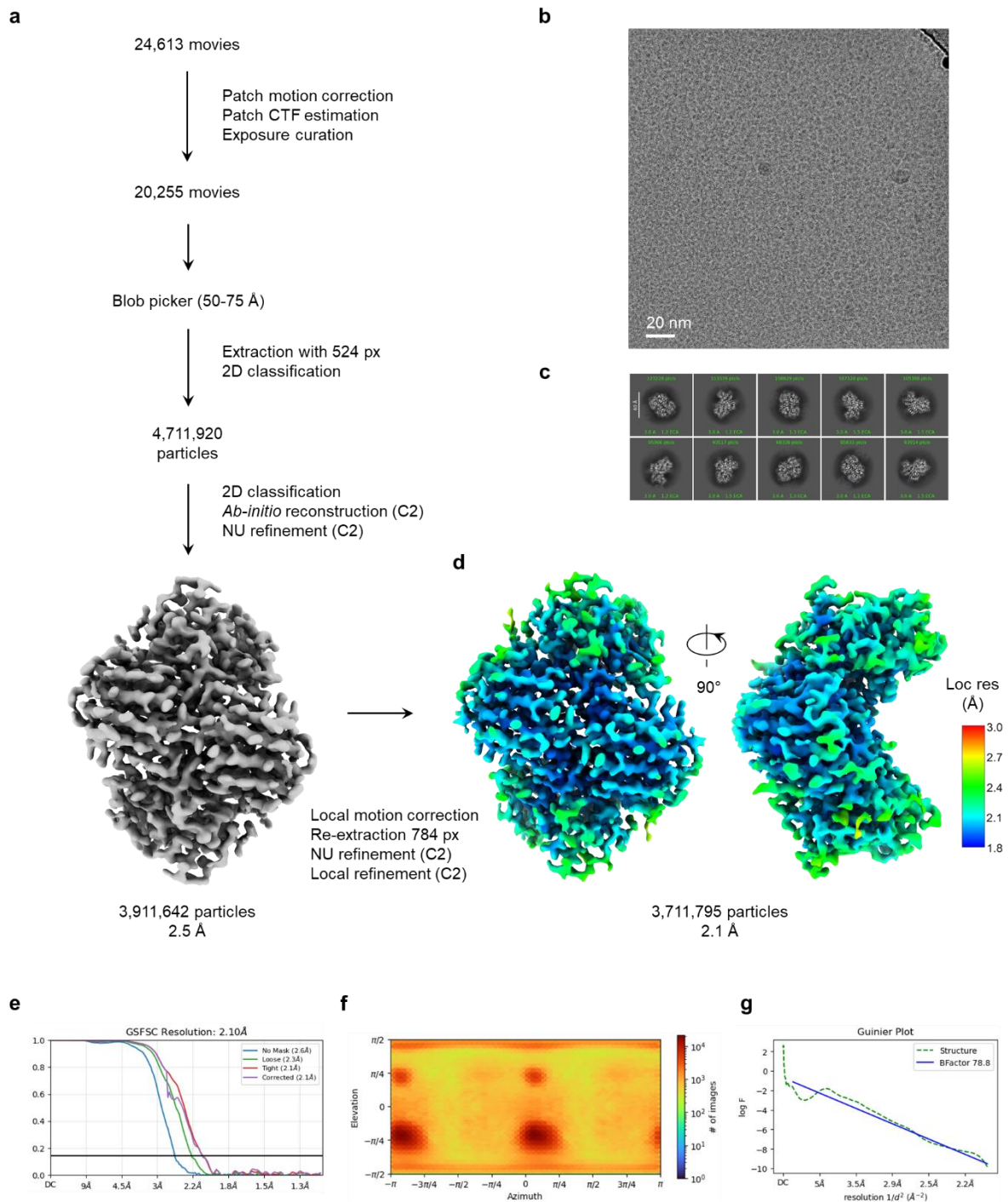
**Figure 24: Purification and thermal shift assay of hTPK1.** **a** Size-exclusion chromatography profile of His-tag purified hTPK1 from 6 L bacterial expression culture (Superdex S200 10/300 Increase column). **b** SDS-PAGE analysis of the SEC fraction. The \* marks the fractions from the main SEC peak containing the highest concentration of hTPK1. In the purified hTPK1-construct, the hTPK1 monomer has a molecular mass of 32 kDa, including tags. **c** nanoDSF analysis of hTPK1 in the presence of its endogenous substrate thiamine (20  $\mu$ M) and the co-substrate ATP (2 mM ATP-Mg<sup>2+</sup>). **d** Thermal shift assay (nanoDSF) of hTPK1 in the presence of 200  $\mu$ M of different compounds.

### Cryo-EM structure of the human TPK1

hTPK1 is, with a combined structured mass of 54 kDa, on the small end of what cryo-EM can resolve. In a first attempt, I collected data at a nominal magnification of 105,000 x (pixel size of 0.83 Å) on a Titan Krios G3, equipped with an X-FEG, a Gatan imaging filter (GIF, 20 eV slit), and a K3 direct electron detector (Gatan). The resulting dataset yielded first refined 3D reconstructions of hTPK1, which was overall in good agreement with the crystal structures of this enzyme. The nominal map resolutions went to about 2.7 Å (data not shown). However, the corresponding 3D maps showed non-protein streak-like densities, which are likely reconstruction artefacts. Such artefacts might ultimately arise from the low signal-to-noise ratio in the hTPK1 particle images, which impedes both the proper cleaning of

particle stacks, as well as 3D pose refinement. Trying to mitigate this problem, I decided to collect data on a higher-end Titan Krios G4 microscope with the help of Joseph Bartho at the Imaging Centre at EMBL Heidelberg. The microscope was equipped with a cold-FEG, a Selectris X imaging filter (Thermo Fisher Scientific, 10 eV slit), and a Falcon 4i camera. The imaging parameters were specifically tailored to favour high-resolution reconstructions of a small target. We decided to collect the data at a significantly higher magnification of 215,00 x (pixel size of 0.572 Å) to optimally use the detector quantum efficiency (DQE) of the camera for high resolution information, as well as a high electron dose of 70 e-/Å<sup>2</sup> to improve the contrast for the small hTPK1 target. Using a standard data processing workflow in cryoSPARC, as detailed in the corresponding methods Section 2.10, the 3D reconstruction of hTPK1 reached a global resolution of 2.1 Å, with a local resolution of 1.8 Å in the core of the protein (FSC cut-off at 0.143, Fig. 25). This density map was free from any streaking artifacts. hTPK1 could be modelled in full length into the map (amino acids 1-243), with clear density for each residue (Fig. 26). This indicates that the entire polypeptide chain of hTPK1 is structured and that the hTPK1 dimer forms a rigid unit in solution (Fig. 27).

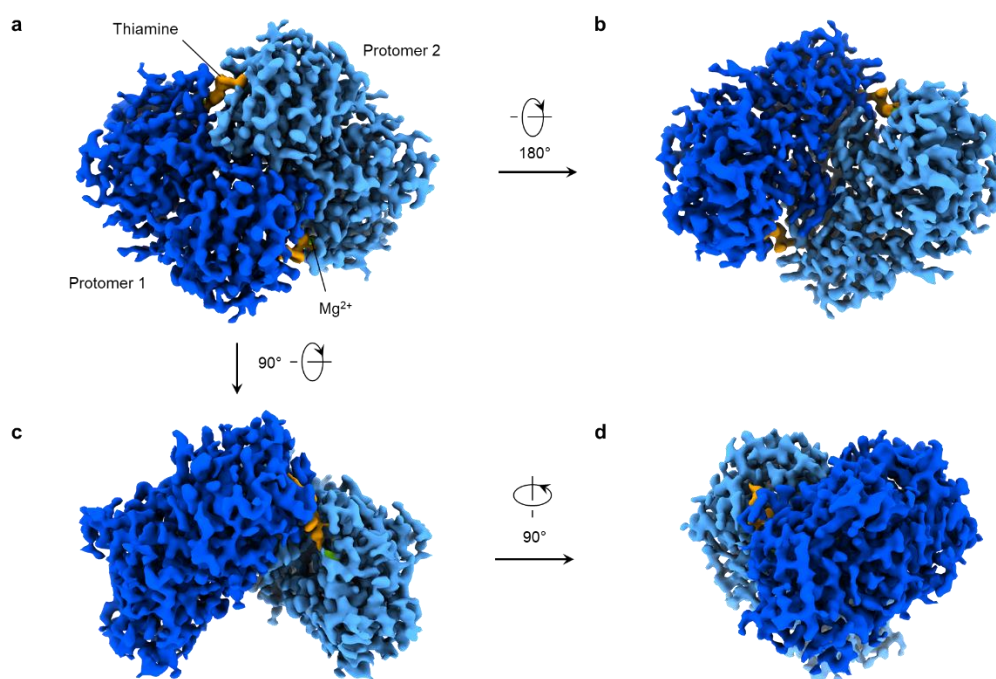




**Fig.25: Cryo-EM data processing for hTPK1.** **a** Data processing workflow of the hTPK1 cryo-EM dataset. Movies were recorded and saved in EER format. The super-resolution pixel size was 0.286 Å. With an upsampling factor of 2, the physical pixel size at specimen level is 0.572 Å. **b** Representative micrograph acquired at a defocus of 1.6 μm. **c** 2D class averages. **d** Final cryo-EM density map of the hTPK1 dimer coloured by local resolution (ChimeraX contour level 1.7). **e** Gold-standard Fourier shell correlation (GSFSC) curves between two half maps. FSC threshold = 0.143. **f** Angular distribution of particles in the final 3D local refinement. **g** Guinier plot for B factor estimation.

## The cryo-EM structure of TPK1

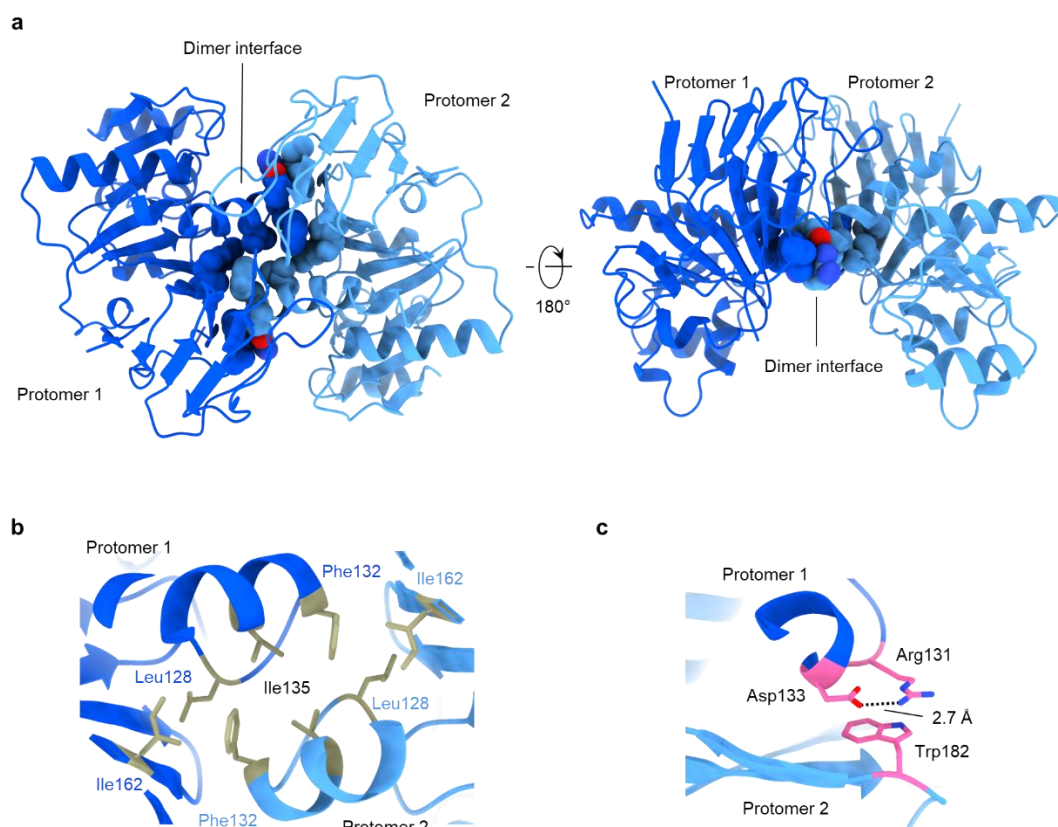
The individual hTPK1 protomers fold respectively into two distinct domains, an N-terminal  $\alpha/\beta$ -domain (NTD, residues 1-159) and a C-terminal  $\beta$ -sandwich domain (CTD, 160-242) (Fig. 28). hTPK1 forms a stable homodimer, with a buried surface area of 1,391  $\text{\AA}^2$  (only protein, atom radius of 1.4  $\text{\AA}$ ). Some key protein-protein interactions between the protomers 1 and 2 (P1 and P2) can be identified in the cryo-EM structure (Fig. 27). Phe132<sup>P1</sup> inserts into a hydrophobic pocket formed by Leu128<sup>P2</sup>, Ile135<sup>P2</sup>, and Ile162<sup>P2</sup>. A  $\pi$ -cation interaction is formed between Trp182<sup>P1</sup> and Arg131<sup>P2</sup>, which is at the same time engaging in a salt bridge with Asp133<sup>P2</sup> (2.7  $\text{\AA}$ ). These interactions are likely important for the stability of the hTPK1 dimer. However, biophysical analysis of the formation and stabilisation of the hTPK1 dimer has, to my knowledge, not been attempted yet. The inspection of the substrate binding site of hTPK1 showed, to my surprise, clear density for a molecule that fits in shape and coordination well to thiamine (Fig. 29). As mentioned above, the hTPK1 sample used to prepare the grids for this cryo-EM data collection has been extensively diluted and washed in ligand-free buffers. Nevertheless, I found this thiamine-like molecule bound in the substrate binding site of the enzyme. To elucidate the identity of this small molecule, we performed untargeted mass spectrometry analysis of the hTPK1 sample in collaboration with Bernhard Drotleff from the EMBL Metabolomics Core Facility. This experiment revealed that the sample contained indeed high amounts of thiamine and to a lesser extend thiamine-pp (Fig. 30 & 31). The implications of this finding for sample preparation will be discussed further in a Section below. For the following discussion of the structure, I will treat hTPK1 as thiamine-bound.



**Fig. 26: Cryo-EM structure of hTPK1.** a–b show different rotations of the final density map of hTPK1. The different protomers of the hTPK1 dimer are coloured in different shades of blue. Ligand density is highlighted in orange. ChimeraX contour level: 1.6. Hide dust level: 10.7.

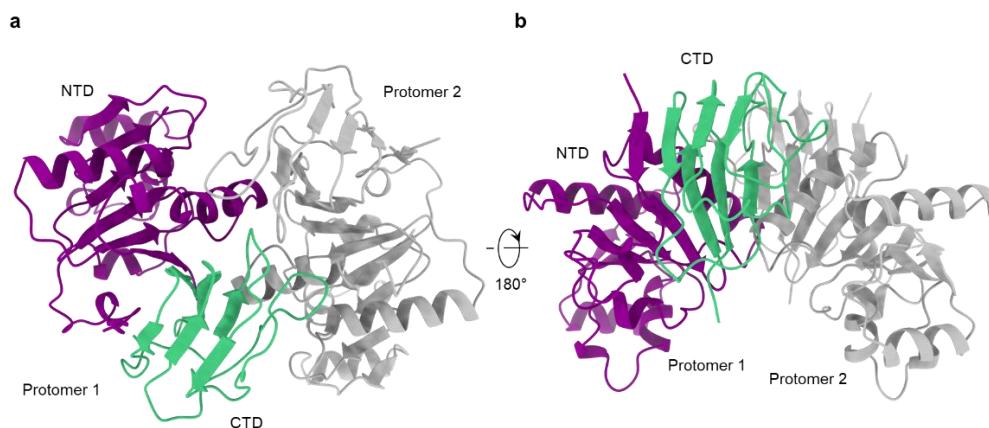


Thiamine binds in a cleft between the two protomers of hTPK1 (Fig. 29). Its pyrimidine ring forms a parallel  $\pi$ - $\pi$  stacking interaction with Trp202. The methyl group of the pyrimidine ring is inserting into a hydrophobic pocket formed by three leucine residues of the CTD (Leu204, Leu209, Leu214), while the adjacent primary amine of the vitamin is in proximity to Asp97. Density for the carboxyl moiety of this residue is, however, lacking in the density map. This is likely a sign of radiation damage, as hTPK1 was exposed to a high dose of electrons ( $70 \text{ e}^-/\text{\AA}^2$ ). Carboxyl groups are known to be among the first moieties to experience radiation damage, dissociating from the side chain in the form  $\text{CO}_2$  (Garman, 2010, Naydenova et al., 2020). Another important interaction of the pyrimidine ring is a hydrogen bond of its ring nitrogen with the hydroxy group of Ser216 ( $2.9 \text{ \AA}$ ). The hydroxyethyl tail of thiamine points towards the catalytically active  $\text{Mg}^{2+}$  ion (Fig. 29). Thiamine adapts a relaxed F-form when binding to hTPK1, meaning that its two aromatic rings are staggered  $90^\circ$  relative to each other (Timm et al., 2001). This stands in contrast to the coordination of the vitamin in the form of TPP in thiamine-dependent enzymes, in which the two aromatic rings are fixated in a parallel, high-energy V-conformation (Huang et al., 2013).



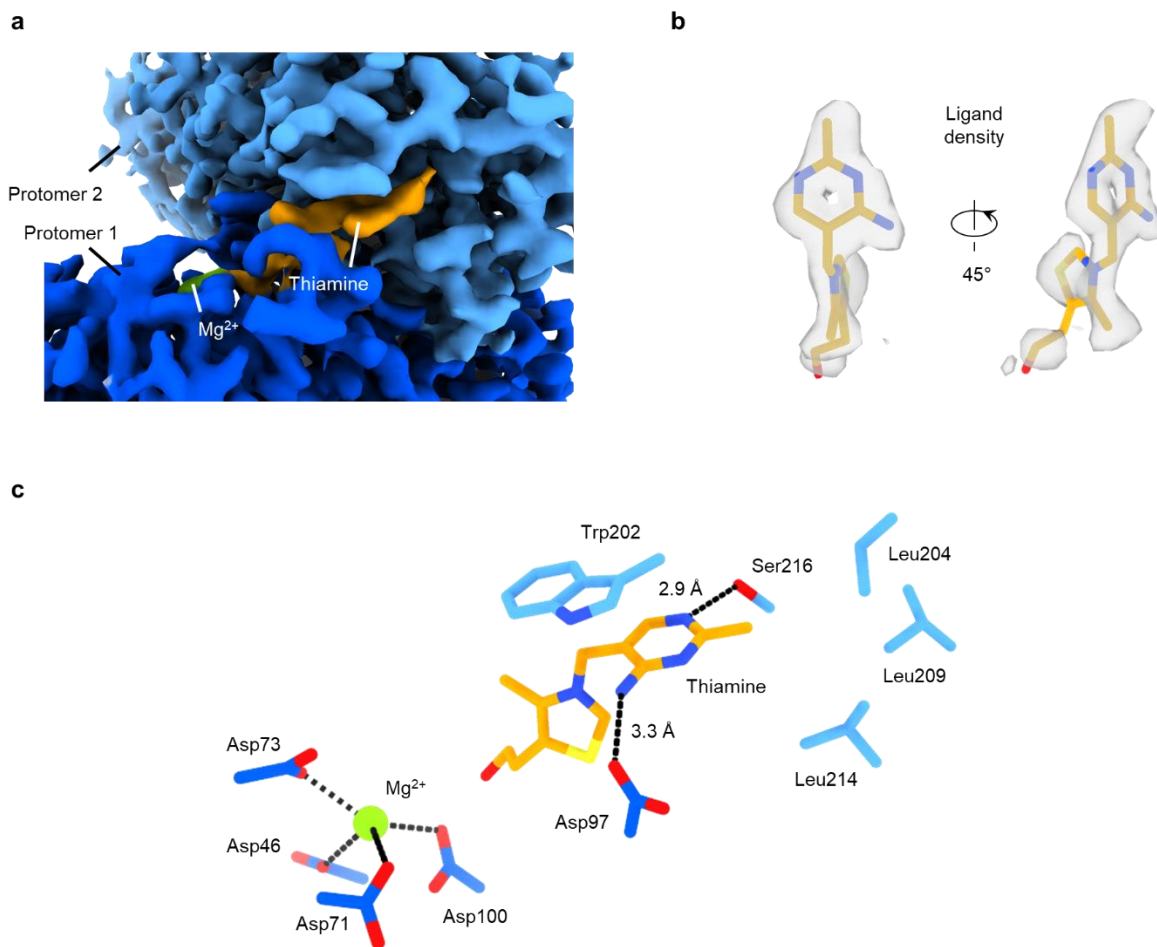
**Fig. 27: Fold of the hTPK1 homodimer.** **a** Backbone representation of the hTPK1 structure determined by cryo-EM. Residues of the dimer interface are shown as spheres. **b** Hydrophobic core of the hTPK1 dimer interface. **c** Polar interaction of the hTPK1 dimer interface.

The kinase activity of hTPK1 is critically dependent on  $Mg^{2+}$ , which is coordinated by four aspartate residues of the NTD, namely Asp46, Asp71, Asp73, and Asp100. These residues are fully conserved (Fig. 33). The density around the ion location is somewhat distorted. Likely reasons for that are radiation damage and the chemical mixture of thiamine and thiamine-pp present in the sample. The latter aspect would lead to a mixture of coordination states of  $Mg^{2+}$  and consequently cause a merge of different ion locations in the averaged density map.

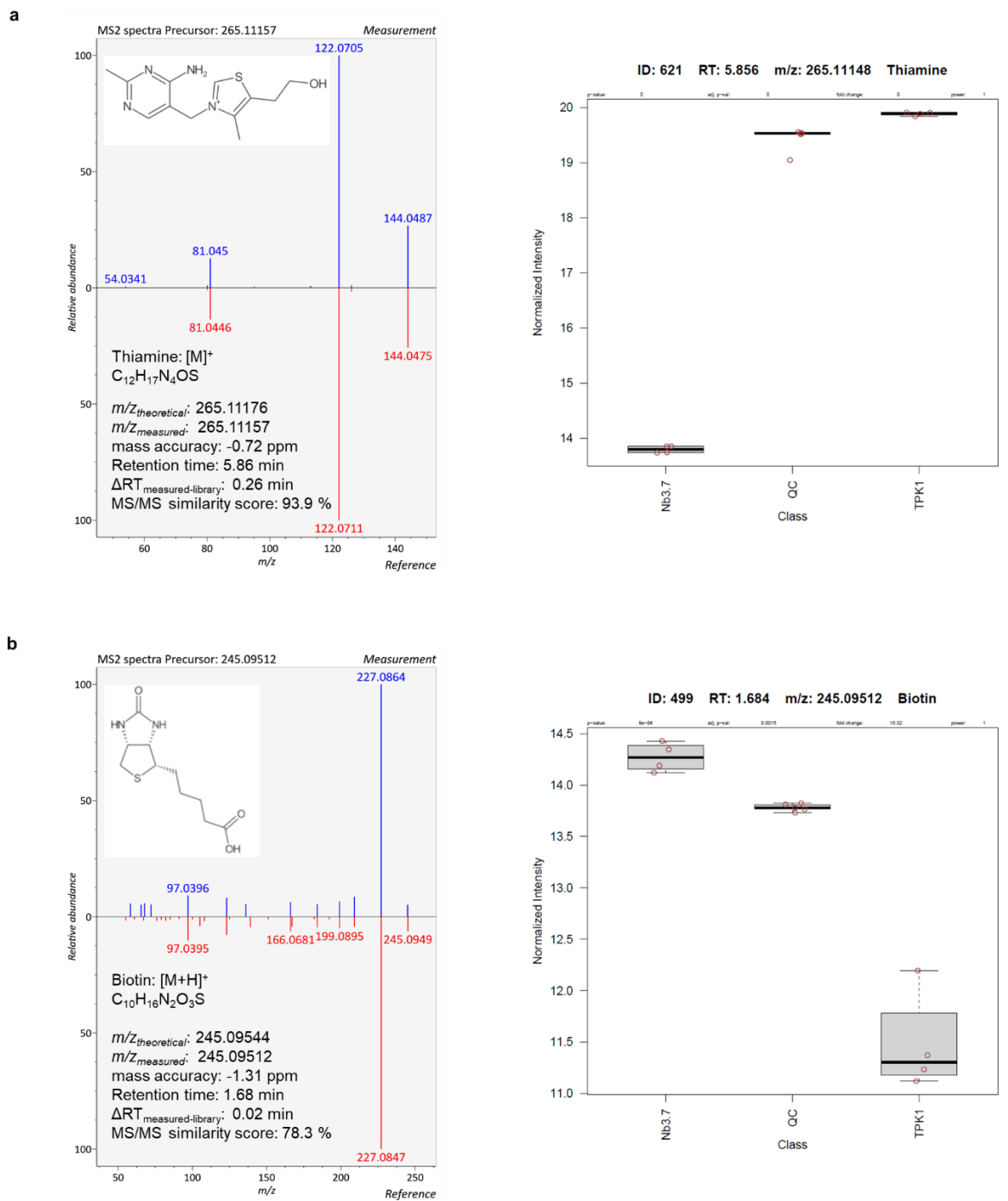


**Fig. 28: Domains of hTPK1.** **a** and **b** show different rotations of the backbone representation of the hTPK1 cryo-EM structure. Each protomer fold in two distinct domains, an N-terminal (NTD), and a C-terminal domain (CTD).

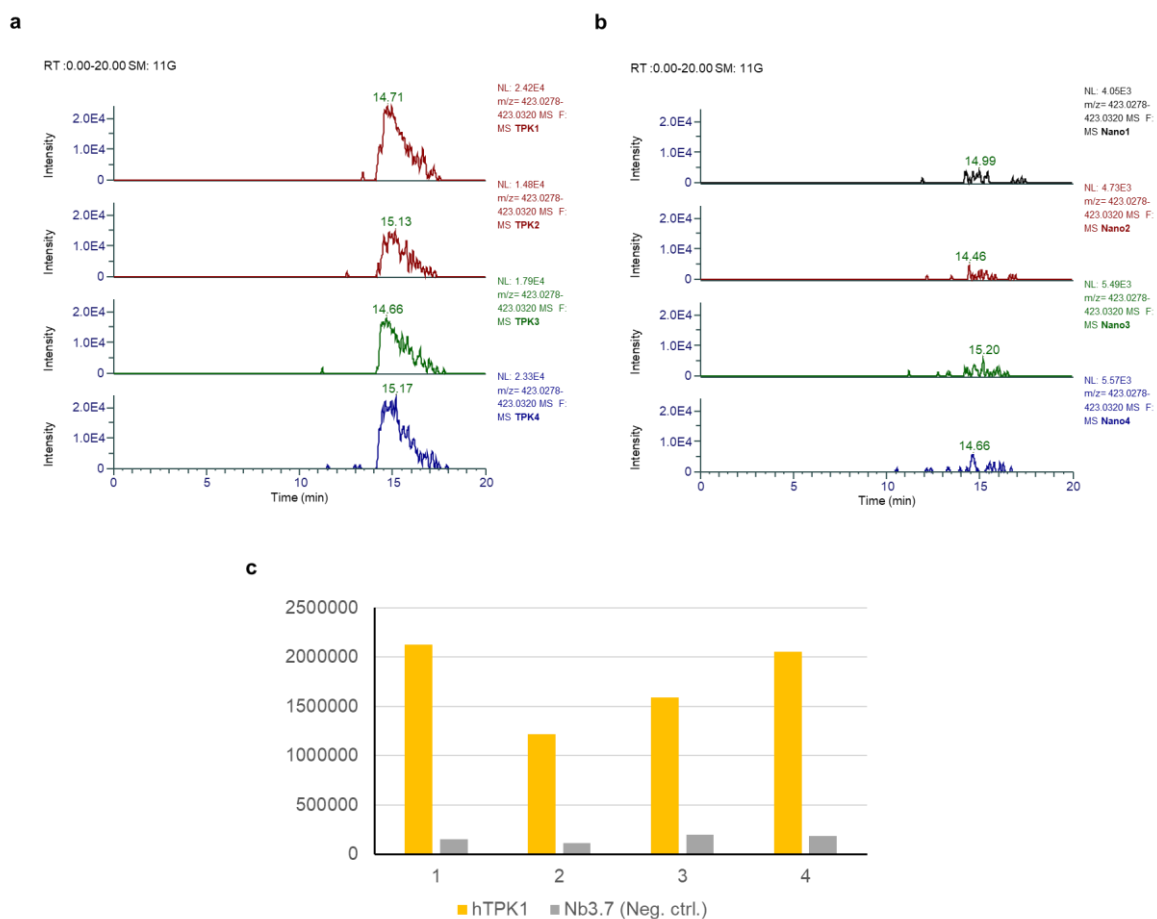
As mentioned above, the density map presented in this study shows likely signs of radiation damage (Russo & Egerton, 2019, Naydenova et al., 2020). Several aspartates and glutamates show only weak density for their carboxyl groups (e.g. Glu9, Asp97, Glu158, Glu159), and the  $Mg^{2+}$  density is distorted. The high dose of  $70 e^{-}/\text{\AA}^2$  is likely necessary to obtain good reconstructions of small soluble proteins like hTPK1. To computationally reduce the impact of radiation damage on high resolution features, like carboxyl groups, the cryo-EM movies have already been dose weighted in the motion correction programmes. A promising strategy to further reduce the effects of radiation damage in the final 3D reconstructions is offered by the use of HexAuFoil® grids (Quantifoil). These gold foil grids have significantly smaller holes (300 nm) than conventional cryo-EM grids (1.2-2.0  $\mu\text{m}$ ). Combined with the hexagonal geometry of their grid bars, they provide a support system for the vitreous sample that shows little to no electron beam induced motion. This allows one to mathematically extrapolate the 3D reconstruction of the target to a hypothetical zero dose (Naydenova et al., 2020). Future cryo-EM data of hTPK1 should be collected on these grids and a corresponding computational compensation of radiation damage should be tried.



**Fig. 29: Substrate binding site of hTPK1.** **a** Cryo-EM structure of hTPK1. The two protomers of hTPK1 are coloured in dark and light blue, respectively. Ligand density is highlighted in orange. ChimeraX contour level: 1.6. **b** Ligand density in the substrate binding site superimposed with the modelled thiamine molecule. ChimeraX contour level: 2.7. Density in a zone of 1.6 Å around the model atoms is shown. **c** Stick representation of the substrate binding site of hTPK1 in the cryo-EM structure model.



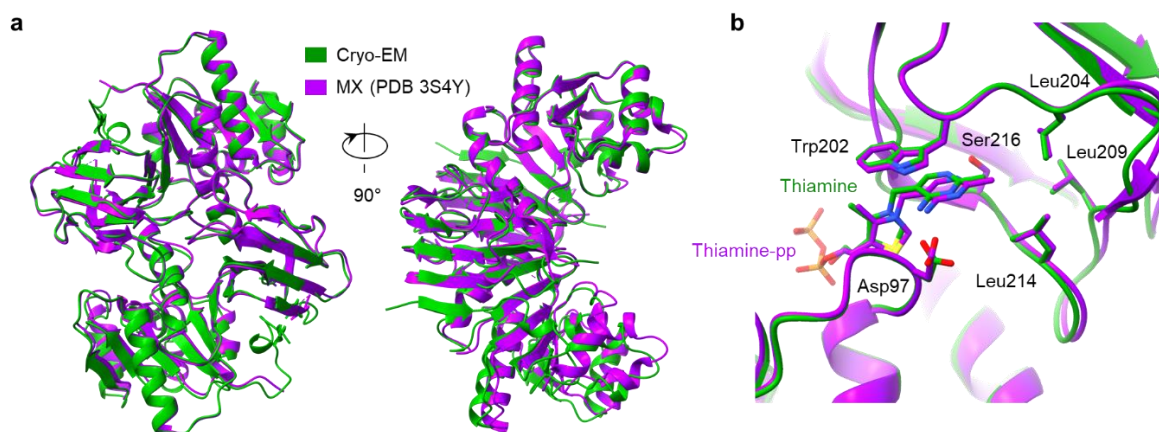
**Fig. 30: Untargeted mass spectrometry analysis of purified hTPK1.** Left panels: Mass spectra found in the hTPK1 sample (blue), compared to reference spectra (red). Right panels: Normalised intensity of the observed mass fingerprints in the hTPK1 sample, a quality control sample (QC), and a negative control sample (Nb3.7). **a** Mass spectrum of thiamine. **b** Mass spectrum of biotin. Analysis performed by Bernhard Drotleff as described in Dekina et al., 2024.



**Fig. 31: Untargeted MS of purified hTPK1 with a focus on thiamine-pp. a** Mass spectrum of thiamine-pp in hTPK1 samples (n =4). **b** Mass spectrum of thiamine-pp in negative control samples (Nb3.7) (n = 4). **c** Thiamine-pp mass peak intensities (integrated signals) in four technical replicates. Analysis performed by Bernhard Drotleff as described in Dekina et al., 2024.

## Comparison of the cryo-EM structure with MX structures of TPK1

Crystal structures of TPK1 have been published before (Table 7). The structures of yeast and mouse TPK1 were solved in the lab of Robert A. Harris at the Indiana University School of Medicine. Another mouse TPK1 structure was published by the same lab in 2006, this time of the enzyme in complex with pyriothiamine pyrophosphate and AMP. Ten years later, in 2011, the crystal structure of the human TPK1 was released in the PDB by the Structural Genomics Consortium (SGC, Shen et al., 2011). On a structure model level, the cryo-EM structure presented in this study is overall very similar to the crystal structures. A noticeable global difference between the results of the two techniques is a slightly stronger bend of the C-shaped hTPK1 dimer in the crystal structures (Fig. 32). This could be caused by the crystal packing imposing more conformational restraints on the protein. In the cryo-EM model, the kinase adopts a mildly more relaxed state. The buried surface area of the homodimer interface is with  $1,391 \text{ \AA}^2$  about identical to that of the crystal structures solved by Harris and colleagues and the SGC ( $\sim 1,400 \text{ \AA}^2$ ). The mode of thiamine binding seems to be structurally well conserved between yeast, mouse, and man (Baker et al., 2001, Timm et al., 2001, Shen et al., 2011).



**Fig. 32: Comparison of crystal and cryo-EM structures of hTPK1.** **a** Superposition of the hTPK1 cryo-EM structure solved in this study and the hTPK1 crystal structure determined by the SGC (Shen et al., 2011). **b** Close-up on the substrate binding site of the kinase.

The coordination of thiamine is also close to identical in the crystal structure (PDB ID 3S4Y) and the cryo-EM structure of the human TPK1 (Fig. 32). On the level of map quality, I would argue that the published crystallographic electron density map of hTPK1 (PDB ID 3S4Y) is still slightly superior to the cryo-EM map. On the one hand, its nominal resolution of 1.8 Å is better than the 2.1 Å reached by cryo-EM. On the other hand, the impact of radiation damage appears to be much lower in the crystal structure. The integrity of both the carboxyl groups and the  $Mg^{2+}$  ion is better preserved in the electron density, compared to the cryo-EM map. Another important feature of the crystal structure of hTPK1 is the clear resolution of water molecules. In cryo-EM, the identification of water molecules and their distinction from random noise are more difficult (Pintilie & Chiu, 2021).

**Table 7: Crystal structures of TPK1.** FL: full length, NA: not available, Nt: N-terminal, SGC: Structural Genomics Consortium.

Species	Sequence	pH	Ligands	Resolution	PDB ID	Publication
Baker's yeast ( <i>S. cerevisiae</i> )	NA, His-tag	5.1	Thiamine	1.8 Å	Not deposited	Baker et al., 2001
Mouse ( <i>M. musculus</i> )	FL Nt-His <sub>20</sub> -tag	7.1	Thiamine	1.9 Å	1IG3	Timm et al., 2001
Mouse ( <i>M. musculus</i> )	FL Nt-His <sub>20</sub> -tag	7.1	Pyriothiamine-pp, AMP, $Mg^{2+}$	2.5 Å	2F17	Liu et al., 2006
Human ( <i>H. sapiens</i> )	FL Nt-His <sub>6</sub> -tag	5.5	Thiamine-pp, $Mg^{2+}$	1.8 Å	3S4Y	Shen et al., 2011 (SGC)

## **On the chemical purity of recombinant protein samples**

The published substrate affinities and other enzymatic parameters of TPK1 vary greatly between different laboratories, as concluded by Sambon et al., 2022. The authors of this paper propose that the underlying cause for this quite significant reproducibility problem lies in the varying and often inadequate choice of ATP and  $Mg^{2+}$  concentrations. In my view, that is certainly a factor to be considered. However, I suggest that another, more concerning factor might contribute to the deviating biophysical parameters measured for TPK1. I determined the cryo-EM structure of hTPK1 on a sample that was purified in the absence of thiamine. The protein has seen extensive dilutions, more than 60 column volumes of high-salt buffer washes on Ni-NTA resin, a size-exclusion chromatography, and the dilution from a 150  $\mu$ M crystallography sample to a 30  $\mu$ M cryo-EM sample (Fig. 24). A far more extensive clean-up in comparison to other published work (Liu et al., 2006, Huang et al., 2019, Sambon et al., 2022). And yet, the cryo-EM map shows clear ligand density that matches the molecular structure of thiamine (Fig. 29). We observe a highly similar density in an X-ray structure of crystals grown from the same hTPK1 sample (MX data not shown). Untargeted mass spectrometry (MS) revealed that this sample did indeed contain thiamine, as well as small amounts of thiamine pyrophosphate (Fig. 30 & 31). That is troubling, as it implies that the published enzymatic parameters are not only incomparable but simply inaccurate. This problem has also general implications for structural biology. Chemical impurity with regard to small molecule ligands is something that cannot be resolved *a posteriori*, neither in cryo-EM, nor in X-ray crystallography. Especially in lower-resolution reconstructions ( $>3.5$  Å), ligand contamination can go unnoticed and might lead to wrong interpretations of the obtained density maps. I would strongly advocate to take this aspect serious, when preparing protein samples for biophysical assays or structural work.

## **Disease-causing mutations in TPK1 from a structural perspective**

Defects in hTPK1 cause thiamine metabolism dysfunction syndrome 5 (THMD5). This syndrome is characterised by Leigh disease-like phenotypes, comprising acute encephalopathy, ataxia, dystonia, and seizures. These symptoms have an onset in early childhood and are accompanied by elevated levels of lactate and oxoglutarate in serum and cerebrospinal fluid. hTPK1 deficiency disorders respond usually well to treatment with high doses of thiamine by oral administration (Sambon et al., 2022). The known disease-causing hTPK1 mutations are scattered all over the enzyme (Table 8) and have varying effects on its stability, substrate binding affinity, and biochemical activity (Huang et al., 2019). The mutations directly affecting the substrate binding site are Asn219Ser and Trp202Gly. Both residues are involved in the coordination of the aminopyrimidine ring of thiamine (Fig. 29). The impact of their on the activity of hTPK1 can presumably be attributed to a perturbation of substrate binding. The effects of the remaining mutants (Table 8) are less clear from a structural point of view. They might impair the folding or the stability of hTPK1, but this remains to be investigated. I propose that cryo-EM, in conjunction with biophysical techniques, would be a suitable tool to study the structural effects of these mutations.

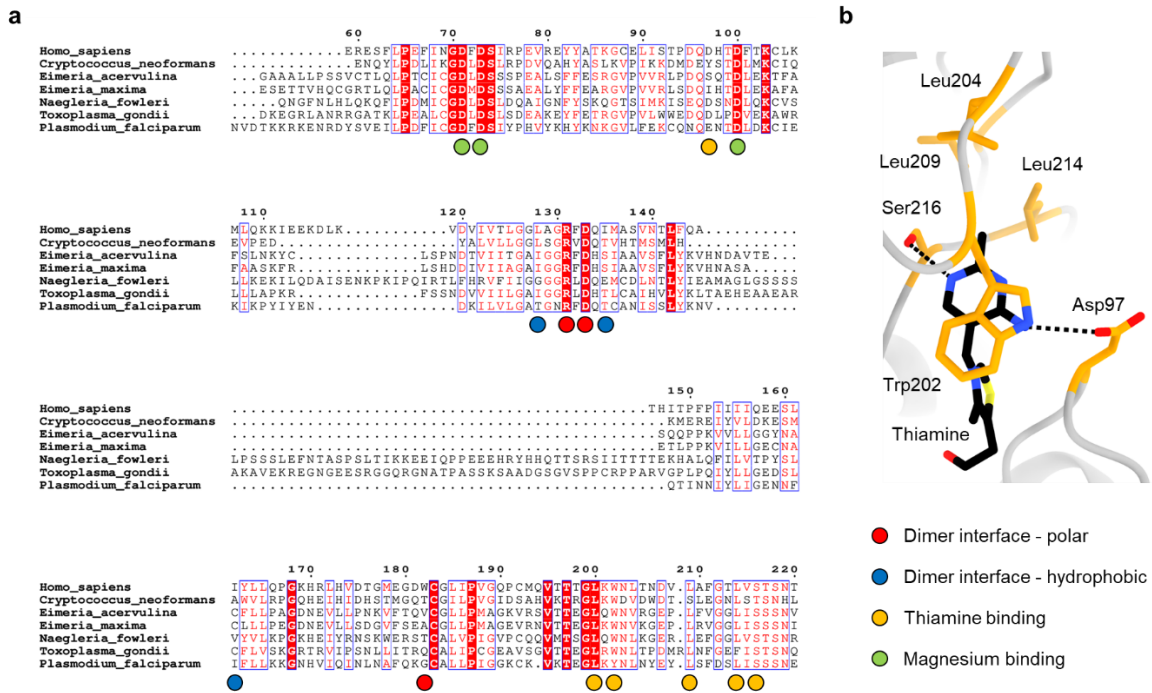
**Table 8: Overview of disease-causing point mutations of hTPK1.** Approximate values adapted from Huang et al., 2019. \*Residues of the substrate binding site of hTPK1. n.d.: not determined.  $k_a$ : Association rate constant.

Mutation	$\Delta T_m^{\text{mutant-wildtype}}$	$k_a$ (% wildtype)	Activity (% wildtype)	Reference
Leu28Ser	+2.0 °C	120%	n.d.	Huang et al., 2019
Leu40Pro	-2.0 °C	n.d.	60%	Mayr et al., 2011, Mahajan & Sidiropoulos, 2017
Asn50His	+1.0 °C	n.d.	40%	Mayr et al., 2011
Ser160Leu	-7.5 °C	25%	240%	Banka et al., 2014
Trp202Gly*	-6.0 °C	$\approx 0\%$	200%	Fraser et al., 2014
Asn219Ser*	-9.0 °C	n.d.	75%	Mayr et al., 2011
Asp222His	-7.0 °C	n.d.	80%	Banka et al., 2014

### TPK1 as drug target

TPK1 has been described as a potential drug target in *Plasmodium falciparum*, the major malaria causing species of the *Plasmodium* genus (Chan et al., 2013). From a structural perspective, I would put this into question. The substrate binding site of TPK1 is highly conserved between humans and *P. falciparum*, as well as other eukaryotic parasites (Fig. 33). It is therefore likely that any drug targeting the parasite TPK1 would also block the human TPK1, with presumably detrimental side effects for the patient. When looking into the TPK1 orthologues of other eukaryotic parasites, the situation is largely similar (Fig. 33). The only thiamine-coordinating residue that is not highly conserved is Asp97. In *Cryptococcus neoformans* and *Eimeria maxima*, this residue is replaced by a tyrosine and an isoleucine, respectively (Fig. 33). Asp97 forms a hydrogen bond with the primary amine of thiamine in the human TPK1. By designing thiamine analogues carrying an apolar substituent in place of this amine, one could potentially generate parasite-selective TPK1 inhibitors.





**Fig. 33: Sequence and structure conservation of TPK1 in human parasites. a** Multiple sequence alignment of the human TPK1 with the TPK1 orthologues from pathogenic eukaryotic parasites. **b** Close-up on the substrate binding site of hTPK1. Asn97 is the only thiamine-coordinating that shows significant variability in certain parasites.

As thiamine-pp is essential for cellular survival, hTPK1 comes into mind as an anti-neoplastic drug target. One could imagine an anti-vitamin B<sub>1</sub> that cannot penetrate the blood-brain barrier (BBB). Our structures of hSLC19A3, the main thiamine transporter of the BBB, as well as hTPK1 could guide such drug design efforts. Negatively charged thiamine analogues could be promising candidates. They are incompatible with the substrate binding site of hSLC19A3 (Supplementary Fig. 1) but should be accepted by hTPK1. An example would be the non-hydrolysable thiamine-S-p. It could bind with high affinity to hTPK1 and prevent the production of thiamine-pp from thiamine and ATP. Thiamine-S-pp, in contrast, would likely be ineffective, as it can probably bind to TDEs and substitute thiamine-pp without major consequences for TDE activity.

## 4. Discussion and Conclusion

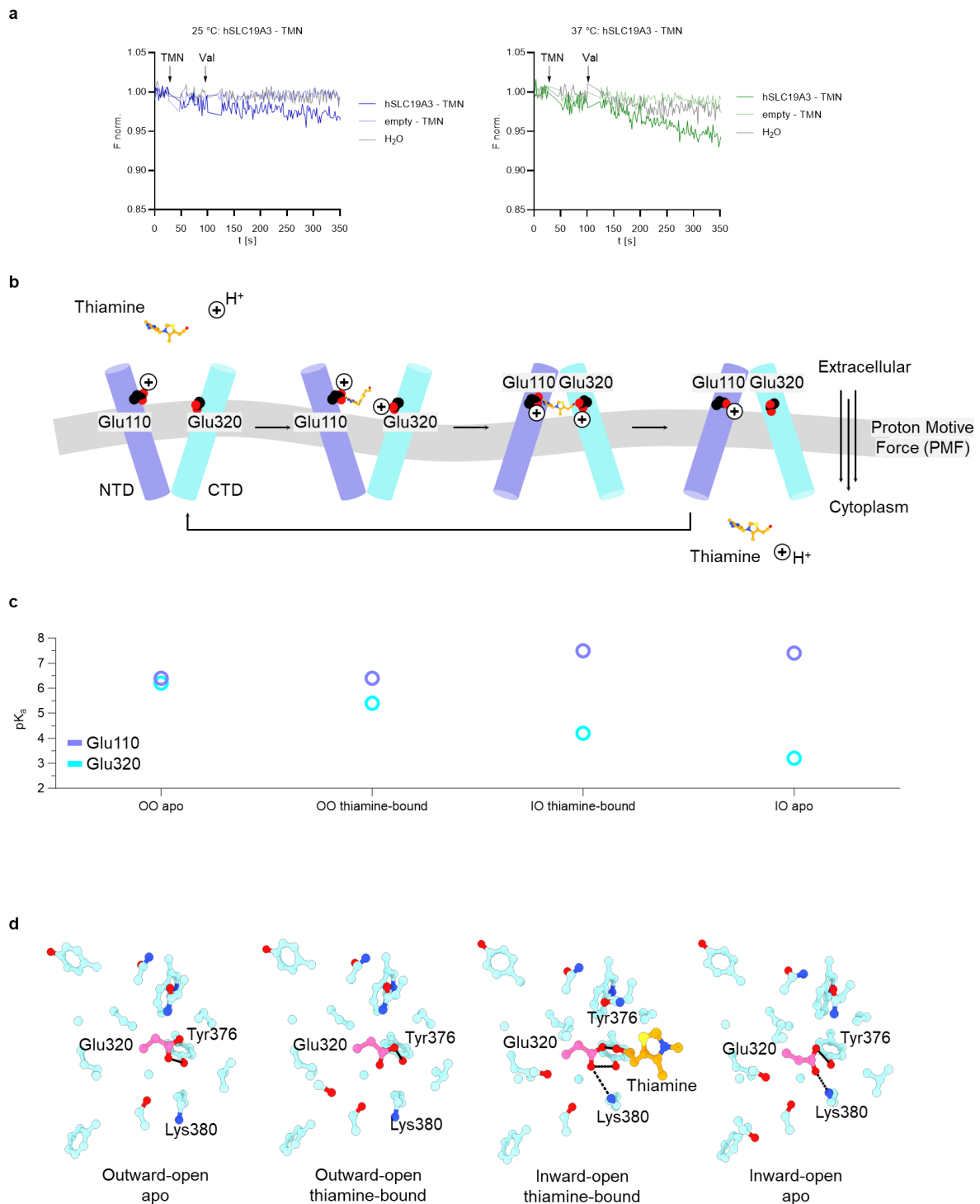
### 4.1 Limitations of the study

The combination of nanobody discovery and cryo-EM led to the determination of high-resolution structures of the human thiamine transporter SLC19A3 in its outward- and inward-open conformational states. This provides important advances in the understanding of thiamine transport in humans and other mammals on a molecular level. There are, however, some limitations of the presented results. It is important to note that all structures of the hSLC19A3 were determined using the detergent mixture of LMNG and CHS as an artificial mimetic of the natural habitat of the transporter, phospholipid membranes. While the rigid parts of the protein, i.e. the NTD and CTD in themselves are unlikely to be affected by this artificial environment, flexible modalities of hSLC19A3 might well behave differently in detergent, compared to a native biological membrane. This is conspicuous for the case of TM12 of hSLC19A3. In the detergent-solubilised transporter, this helix appears rather flexible and protrudes from the micelle in the membrane plane (Fig. 13). In a continuous lipid bilayer this structural configuration is sterically unlikely and TM12 would presumably be more compactly packed to the rest of the CTD, emerging from the bilayer rather orthogonal to the membrane plane. Another important limitation of the experimental work on hSLC19A3 is the affinity determination using thermal shift assays (nanoDSF). The determined  $K_{d,app}$  values for thiamine and other compounds should only be seen as apparent affinities and first approximations. It stands to reason, that the thermodynamic model underlying the fitting of the  $K_{d,app}$  values might have limited validity for the high melting temperatures of hSLC19A3 (60-75 °C). Furthermore, thermal shifts convolute enthalpic and entropic contributions to ligand binding, which complicates an accurate thermodynamic interpretation of the measured titrations. This is showcased by tricyclic antidepressants, which are known to inhibit hSLC19A3, but actually destabilise the protein (Fig. 15). This destabilisation hints to an entropically rather than an enthalpically driven binding event, which is difficult to accurately assess with our simplistic thermodynamic model (Eq. 4). However, as other biophysical techniques, such as ITC and MST, proved unsuitable for the study of hSLC19A3 (data not shown), nanoDSF provided the only experimental access to binding affinities of the purified transporter at the time. In the future, it would be advisable to complement the thermal shift-based affinity determination with more accurate and maybe even kinetic techniques, such as surface plasmon resonance (SPR). A topic, which has not been covered in this study is the interaction of hSLC19A3 with pyridoxine (vitamin B<sub>6</sub>). It is known that pyridoxine can also be transported by hSLC19A3 at low pH (Yamashiro et al., 2020). I omitted the structural investigation of this transport process, predominantly due to time limitations, but also due to the small size of vitamin B<sub>6</sub>, which makes the interpretation of any measured ligand density in cryo-EM maps difficult. Other research groups have succeeded in determining structures of hSLC19A3 in complex with pyridoxine (Dang et al., 2024, Qu et al., 2024). I will discuss their findings in Section 4.3.

Despite the plethora of structural data presented in this study, important questions regarding thiamine transport remain to be answered. A central question for the molecular understanding of hSLC19A3 is its transport mode. While I provide structural snapshots of the different conformational states of the transporter, my data do not allow to make any definitive statements about the kinetics and thermodynamics of the transport cycle of hSLC19A3. To date, it is still under debate whether hSLC19A3 functions as a uniporter or actually needs a co-substrate in the form of protons, ions, or other organic molecules to efficiently drive thiamine transport. In a previous study, thiamine uptake was found to be independent on sodium concentrations, which makes this an unlikely co-substrate (Subramanian et al., 2006). Given the currently published data, making conclusions about the involvement of proton coupling of hSLC19A3 is very difficult. There are studies that support the idea of either proton symport or antiport (Dudeja et al., 2001, Yamashiro et al., 2020), while others find no evidence for it (Liang et al., 2015). An underlying problem of all existing studies on this topic is that they base their conclusions on cell-based thiamine uptake assays. These only assess the influence of extracellular pH on the activity of hSLC19A3. In this setup it is not possible to causally deduce strict proton coupling from the observed pH-dependency of thiamine uptake. Some easily imaginable effects could explain the observed results equally well.

1. Protonation and deprotonation of binding site residues could alter the affinity of hSLC19A3 for thiamine. As I have shown, the substrate binding site of hSLC19A3 comprises three critical glutamate residues, Glu32, Glu110, and Glu320, which can in principle all be present in their protonated or deprotonated form, dependent on the external pH. As the substitutions of these glutamate residues by glutamine show, subtle changes in the hydrogen bonding capacities in these positions can strongly reduce thiamine affinity (Fig. 14). This would result in pH-dependent changes in the observable transport activity of hSLC19A3 in cell-based assays. Protonation-dependent affinity changes would in this context be practically indistinguishable from actual proton coupling.
2. The external pH affects the stability of hSLC19A3. As our work on the transporter could show, the stability of hSLC19A3 is strongly influenced by pH, particularly at pH5 and lower (Fig. 8). This pH-dependent stability profile shows interesting resemblance of the pH-dependent activity profiles published in other studies (Rajgopal et al., 2001). The pH-dependent destabilisation of hSLC19A3 could mask proton coupling of thiamine transport in cell-based assays.

I can by no means provide a significantly better conclusion on this topic today, but I would like to offer a possible model based on our data. In order to settle the question of proton coupling in hSLC19A3 it is necessary to study the transporter in the controlled environment of reconstituted proteoliposomes. We could show that hSLC19A2 and hSLC19A3 can be efficiently reconstituted in proton-tight liposomes. Using the pH sensitive dye pyranine, we found that both transporters mediate very subtle substrate-driven symport of protons at 37 °C (Fig. 34, Spriestersbach, 2022).



**Figure 34: “Speculative model for the molecular basis of proton-driven thiamine uptake.** **a** Pyranine-based proton transport assays in proteoliposomes. Fluorescence quenching indicates acidification of the interior of the liposomes and thus implied proton co-transport (panel adapted from Spriestersbach, 2022). **b** Proposed protonation states of two glutamate residues critical for thiamine binding. Glu110 and Glu320, in the different conformational states of hSLC19A3 are highlighted. **c** The proposed protonation states are based on pK<sub>a</sub> predictions using the experimentally determined structures of hSLC19A3 and the prediction software PROPKA (Bas et al., 2008). Glu110 is predicted to have a relatively constant pK<sub>a</sub> in the neutral range (6.4–7.5). Therefore, this residue is likely protonated throughout the transport cycle. Glu320, in contrast, experiences a strong pK<sub>a</sub> shift from 6.2 in the outward-open state (OO), to 3.2 in the inward-open state (IO). **d** On a structural level, this pK<sub>a</sub> shift is likely linked to the interaction of Glu320 with Lys380. In the outward-open state, these residues are >5 Å apart and Glu320 is placed in a rather hydrophobic environment, which makes its protonation favourable. In the inward-open state, however,

Glu320 and Lys380 form a salt bridge, which stabilises the ionised glutamate side chain and thus promotes the dissociation of the previously bound proton. This predicted conformation dependent protonation of Glu320 suggests that this residue could be involved in connecting thiamine uptake to the proton motive force (PMF). In the proposed model, Glu320 would bind a proton from the extracellular space in the outward-open state and release it in the cytoplasm, when it transitions to the inward-open state. The transporter could then cycle back to the outward-open state with deprotonated Glu320, thereby closing the co-transport cycle of a proton. In this model, co-transport of a proton would be an additional driving force for thiamine uptake.” This model was first published, and the figure adapted from our first preprint on hSLC19A3 on bioRxiv (Gabriel et al., 2024).

These data are, however, very preliminary and the experiments would need to be repeated. Using the determined cryo-EM structures of hSLC19A3, I could perform a first structure-based prediction of the  $pK_a$  values of protonatable moieties in the vicinity of the substrate binding site using the program PROPKA (Bas et al., 2008, Fig. 34). This revealed that most residues do not change their  $pK_a$  over the course of the transport cycle. There is, however, one important outlier. Glu320 is predicted to have a near-neutral  $pK_a$  in the outward-open hSLC19A3, and to become strongly acidic once the transporter transitions to the inward-open state (Fig. 34). This change in predicted  $pK_a$  can be structurally explained by the formation of a salt bridge of Glu320 with Lys380 in the inward-open state. This salt bridge likely stabilises the anionic form of Glu320 and promotes its deprotonation (Fig. 34). Putting these findings together, a possible proton symport mechanism could be mediated by the extracellular protonation of Glu320 in the outward-open state of hSLC19A3, and a subsequent release of that proton into the cytoplasm once the transporter is in its inward-open state. This model is very speculative and would need to be tested experimentally. Also, it is possible that a proton coupling through Glu320 is not necessarily a strict requirement for thiamine transport, but something that can coincide and additionally fuel the uptake of the vitamin. Independent of any debatable co-transport mechanisms, a simple facilitated diffusion of thiamine through hSLC19A3 would be energetically driven by two discernible factors. The concentration gradient of thiamine over the plasma membrane provides an entropic driving force for the uptake of the vitamin. This concentration gradient is fundamentally influenced by the pyrophosphorylation of thiamine in the cytoplasm by TPK1. Thiamine-pp appears to be unable to bind to and is most likely not transported by hSLC19A3 (Supplementary Fig. 1, Zhao & Goldman, 2008). The pyrophosphorylated vitamin would therefore be effectively removed from the transport equilibrium of hSLC19A3 and metabolically trapped inside the cell. Furthermore, the positive charge of thiamine makes the import of the vitamin along the electric potential of the plasma membrane energetically favourable (Jaehme et al., 2018).

## 4.2 Drug interactions of vitamin transporters

Thiamine transport through hSLC19A3 is inhibited by many commonly prescribed drugs. In this study, I identified 14 previously unknown thiamine uptake inhibitors (TUIs). Seven of them were discovered based on their pharmacological similarity to known TUIs. Seven more were identified through virtual screening against the determined cryo-EM structures. I furthermore elucidated the structural basis of the interaction of the inhibitors fedratinib, amprolium, hydroxychloroquine, and domperidone with hSLC19A3. An important aspect of the observed interactions of these drugs with hSLC19A3 is the mode of inhibition. My structural data show that the tested inhibitors bind orthosterically in the substrate binding site of the transporter. For fedratinib and trimethoprim, it is known that they are xenobiotic substrates of hSLC19A3 and thus compete with thiamine for this transport route (Giacomini et al, 2017). Other inhibitors might not be transported and therefore simply compete with thiamine for binding to the transporter. It is also possible that some inhibitors lock the transporter in an occluded or inward-open state. This would effectively prevent access of extracellular thiamine to the substrate binding site, leading to a non-competitive inhibition of the transporter. Another possibility would be an allosteric inhibition of the transporter. So far, however, we find no structural evidence for this interaction mode of TUIs. Further structural and functional studies are needed to determine the specific mechanistic of hSLC19A3 inhibition, particularly in an *in vivo* context. The pharmacological findings on hSLC19A3 presented in this study are a pivotal step in understanding the drug interactions of this vital transporter. In the future, TUIs identified *in vitro* should be assessed for their effects on thiamine transport and metabolism in humans, particularly in those vulnerable for thiamine deficiencies due to pre-existing conditions. Healthy human beings have presumably sufficient compensation mechanisms to account for the temporary blockage of thiamine transporters. These include alternative transport routes for thiamine or thiamine-pp through other membrane transporters, such as OCT1&2 or SLC44A4 (Fig. 3). Additionally, our cells have some storage capacity for vitamin B<sub>1</sub>, which may allow them to survive an inhibition of thiamine uptake for a brief period of time (Mrowicka et al., 2023). However, drug interactions of thiamine transporters can be more detrimental for patient groups vulnerable to thiamine deficiencies. These include patients with one of the following risk factors:

1. Malnutrition (limited food supply, anorexia nervosa, chemotherapy, etc.)
2. Inflammatory bowel disease (e.g. Crohn's disease)
3. Alcoholism
4. Treatment with other TUIs (e.g. fedratinib, tricyclic antidepressants etc.)
5. Genetic defects in thiamine transport and metabolism (including heterozygous carriers), affecting one of the following genes: SLC19A2, SLC19A3, TPK1, and SLC25A19

Collectively, these risk factors may predispose many people for drug-induced thiamine deficiencies. Special attention should be paid to adequate thiamine supplementation and monitoring of thiamine deficiency-associated symptoms in these patient groups. The concept of drug-induced inhibition of an

essential transporter can be transferred to other vitamin transporters, such as the riboflavin (vitamin B<sub>2</sub>) transporters SLC52A1-3, the folate (vitamin B<sub>9</sub>) transporters SLC19A1 and SLC46A1, or the ascorbic acid (vitamin C) transporters SLC23A1-3. I would therefore promote a concept of vitamin uptake inhibition (VUI) that should be routinely examined in the assessment of drug safety. The flip-side of this story could be that vitamin transporters may present viable routes for targeted drug delivery. Most vitamin transporters come in groups of functionally very similar proteins (like SLC19A2 and SLC19A3), which show tissue- and cell-specific expression patterns. By understanding and harnessing their subtle structural differences, it might be possible to use them as specific transport routes for drugs in the human body.

### **4.3 Structure determination of hSLC19A3 - a Sino-European tale**

The year 2024 is an important time mark for our understanding of the molecular biology of vitamin B<sub>1</sub>. After I published the first preprint of my work on 11<sup>th</sup> March, the research groups of Zhe Zhang from Peking University and Chia-Hsueh Lee at St. Jude Children's Research Hospital in Memphis released a Letter to the Editor in Cell Research on the matter of the structure of hSLC19A3 on 19<sup>th</sup> March. They solved cryo-EM structures of hSLC19A3 using an N-terminal BRIL fusion in complex with a BRIL-specific Fab fragment. With this approach, they resolved the inward-open state of hSLC19A3 in complex with thiamine, pyridoxin, and fedratinib, respectively. This publication was eventually followed by a preprint from the research groups of Liguozhang and Pu Gao from Fudan University and the Chinese Academy of Sciences. They solved cryo-EM structures of hSLC19A3 in complex with thiamine, pyridoxine, and metformin, as well as the known thiamine uptake inhibitors fedratinib and amprolium. They also managed to capture the inward- and outward open state of the transporter. Instead of nanobodies, they used hSLC19A3-specific mouse Fab fragments as structural fiducials. This approach allowed them to additionally determine structures of the second high-affinity thiamine transporter in humans, SLC19A2, in its inward-open state. From a scientific point of view, this is an excellent opportunity to compare the results of three independent research teams that used three distinct approaches to determine cryo-EM structures of the human thiamine transporters. On a global level and within the achieved resolution range, the cryo-EM structures of the inward-open state of hSLC19A3 appear to be practically identical in all three studies. Qu et al. also determined structures of the transporter in its outward-open states stabilised by a Fab fragment. At the time of writing, their structure models have not been released. But based on the depictions and descriptions in the published preprint, their outward-open structures are very similar to mine. When it comes to the coordination of ligands in the substrate binding site of hSLC19A3, the studies also come to similar conclusions. I can directly compare my findings with the thiamine- and fedratinib-bound structures of Dang et al., and the thiamine-, fedratinib-, and amprolium-bound structures of Qu et al. While I would argue that my ligand densities are better defined, we all see very similar coordination patterns of the compounds. The studies from

Dang and colleagues and Qu et al. also provide insights that I did not cover in my work. Both papers show the coordination of pyridoxine (vitamin B<sub>6</sub>), the other known endogenous substrate of hSLC19A3. The two research groups find that this small pyridine-derivative intercalated in the aromatic clamp, with its polar substituents engaging in hydrogen bond interactions with Glu32 and Glu110. The obtained resolutions for this ligand are, however, definitely too low to unambiguously assign its orientation in the binding site just based on the cryo-EM density. It seems likely that the rotational pose of this ligand in the plane of its aromatic ring has been primarily modelled based on the surrounding polar and hydrophobic contact points of the transporter. I would argue that orientations other than the ones proposed in these two studies would be equally valid. To unambiguously clarify this question, one would need to determine higher-resolution structures of the pyridoxin-bound hSLC19A3 – ideally at or below 2 Å in the substrate binding site – and complement these with molecular dynamics simulations. Very importantly, Qu and colleagues present the first experimental high-resolution structures of the inward-open state of the second high-affinity thiamine transporter, hSLC19A2. These structures are at a first glance of lower quality than the hSLC19A3 structures. However, the hSLC19A2 cryo-EM structures appear to be in good agreement with the AlphaFold (AF2). hSLC19A2 is consequently also highly similar to hSLC19A3 in regard to its overall structure and its substrate binding site in particular (Supplementary Fig. 14). This supports the idea that the two human high-affinity thiamine transporters SLC19A2 and SLC19A3 are biochemically rather identical.

#### **4.4 Structure determination of small proteins by cryo-EM**

During the beginnings of the ‘resolution revolution’, biological macromolecules with a structured mass below 100 kDa have been inaccessible for structure determination by cryo-EM (Kühlbrandt, 2014). Over the past years, however, this size limit has been repeatedly undercut to ever lower molecular masses (Kimanius et al., 2024). The presented Dissertation demonstrates a continuation of this trend. The hTPK1 homodimer has a combined structured mass of about 50 kDa, and the hSLC19A3:nanobody complexes amount to respectively about 55 kDa. With my EM data, I can show that structured polypeptides of this size can be structurally resolved to a resolution range of 1.8-3.0 Å, if they are either strictly symmetric (C2 symmetry of TPK1) or strictly asymmetric (SLC:nanobody complexes). Pseudosymmetry is still a significant issue in this size range, as the asymmetric signal of particles that small is likely too little for unambiguous pose assignment. For macromolecules below 40-50 kDa, it appears generally necessary to increase their structured mass on a biochemical level. Commonly used strategies include the generation of specific non-covalently binding macromolecular modalities, such as nanobodies or Fab fragments, as well as the genetic fusion of a small target with a structurally well-defined fiducial, such as a BRIL domain. Current strategies often employ a combination of the two approaches, e.g. in the form of BRIL:Fab complexes (Mukherjee et al., 2020) or use of a KOR-ICL3-fusion in complex with specific anti-KOR-ICL3 nanobodies (Che et al., 2020). The physical particle



size limit of 40-50 kDa has not been significantly undercut so far. However, improved instrumentation and data processing might allow to break this limit in the future. On the microscope-level, the cumulative effects of technological advances may push size and resolution limits somewhat further. Data collections at higher magnifications (pixel size  $< 0.5 \text{ \AA}$ ), using better electron sources (cold-FEG), more stable energy filters (SelectrisX), and better cameras can provide higher signal-to-noise ratios, which help with the alignment of small particles. The use of new grid types with increased stability against beam induced motion and thinner, more homogenous vitreous ice – e.g. HexAuFoil grids – will also improve the situation for small targets. Finally, cryo-EM data processing is still being very actively developed and sees significant advances every year. In the foreseeable future, cryo-EM will remain an exciting scientific endeavour, in which biochemical sample preparation, physical data collection and computational data processing go hand in hand to push the limits of structurally accessible biomolecules and spatial resolutions ever further.

## References

- Aburezq, Maryam et al. 2023. “Biotin-Thiamine Responsive Basal Ganglia Disease: A Retrospective Review of the Clinical, Radiological and Molecular Findings of Cases in Kuwait with Novel Variants.” *Orphanet Journal of Rare Diseases* 18(1): 1–13.
- Afonine, Pavel V. et al. 2018. “Real-Space Refinement in PHENIX for Cryo-EM and Crystallography.” *Acta Crystallographica Section D: Structural Biology* 74(6): 531–44.
- Alfadhel, Majid et al. 2019. “Targeted SLC19A3 Gene Sequencing of 3000 Saudi Newborn: A Pilot Study toward Newborn Screening.” *Annals of Clinical and Translational Neurology* 6(10): 2097–2103.
- Baker, L. J., Jill A. Dorocke, Robert A. Harris, and David E. Timm. 2001. “The Crystal Structure of Yeast Thiamin Pyrophosphokinase.” *Structure* 9(6): 539–46.
- Banka, Siddharta et al. 2014. “Expanding the clinical and molecular spectrum of thiamine pyrophosphokinase deficiency: a treatable neurological disorder caused by TPK1 mutations.” *Molecular Genetics and Metabolism* 113(4):301-6.
- Bärland, Natalie et al. 2022. “Mechanistic Basis of Choline Import Involved in Teichoic Acids and Lipopolysaccharide Modification.” *Science Advances* 8(9).
- Bas, Delphine C., David M. Rogers, and Jan H. Jensen. 2008. “Very Fast Prediction and Rationalization of PKa Values for Protein-Ligand Complexes.” *Proteins: Structure, Function and Genetics* 73(3): 765–83.
- Bettendorff, Lucien. 2012. “Present Knowledge in Nutrition - Thiamin.” In *Present Knowledge in Nutrition*, John Wiley & Sons, 261–79.
- Bloch, Joël S. et al. 2021. “Development of a Universal Nanobody-Binding Fab Module for Fiducial-Assisted Cryo-EM Studies of Membrane Proteins.” *Proceedings of the National Academy of Sciences of the United States of America* 118(47).
- Botte, Mathieu et al. 2022. “Cryo-EM Structures of a LptDE Transporter in Complex with Pro-Macrobodies Offer Insight into Lipopolysaccharide Translocation.” *Nature Communications* 13(1): 1–10.
- Bottega, Roberta et al. 2019. “Functional Analysis of the Third Identified SLC25A19 Mutation Causative for the Thiamine Metabolism Dysfunction Syndrome 4.” *Journal of Human Genetics* 64(11): 1075–81.
- Brocchieri, Luciano, and Samuel Karlin. 2005. “Protein Length in Eukaryotic and Prokaryotic Proteomes.” *Nucleic Acids Research* 33(10): 3390–3400.
- Butterworth, Roger F. 1989. “Effects of Thiamine Deficiency on Brain Metabolism: Implications for the Pathogenesis of the Wernicke-Korsakoff Syndrome.” *Alcohol and Alcoholism* 24(4): 271–79.
- Castells-Graells, Roger, and Colleagues. 2023. “Cryo-EM Structure Determination of Small Therapeutic Protein Targets at 3 Å-Resolution Using a Rigid Imaging Scaffold.” *Proceedings of the National Academy of Sciences*.
- Chan, Xie Wah Audrey et al. 2013. “Chemical and Genetic Validation of Thiamine Utilization as an Antimalarial Drug Target.” *Nature Communications* 4.
- Che, Tao et al. 2020. “Nanobody-Enabled Monitoring of Kappa Opioid Receptor States.” *Nature Communications* 11(1).
- Chen, Ligong et al. 2014. “OCT1 Is a High-Capacity Thiamine Transporter That Regulates Hepatic Steatosis and Is a Target of Metformin.” *Proceedings of the National Academy of Sciences of the United States of America* 111(27): 9983–88.

- Chen, Yuanying et al. 2021. "Identification and Functional Analysis of Novel SLC25A19 Variants Causing Thiamine Metabolism Dysfunction Syndrome 4." *Orphanet Journal of Rare Diseases* 16(1): 1–10.
- Ciută, Anca Denise et al. 2023. "Structure of Human Drug Transporters OATP1B1 and OATP1B3." *Nature Communications* 14(1): 1–13.
- Croll, Tristan Ian. 2018. "ISOLDE: A Physically Realistic Environment for Model Building into Low-Resolution Electron-Density Maps." *Acta Crystallographica Section D: Structural Biology* 74(6): 519–30.
- Dang, Yu et al. 2022. "Molecular Mechanism of Substrate Recognition by Folate Transporter SLC19A1." *Cell Discovery* 8(1).
- . 2024. "Substrate and Drug Recognition Mechanisms of SLC19A3." *Cell Research* 34(6): 458–61.
- Dekina, Svitlana, Theodore Alexandrov, and Bernhard Drotleff. 2024. "EMBL-MCF 2.0: An LC-MS/MS Method and Corresponding Library for High-Confidence Targeted and Untargeted Metabolomics Using Low-Adsorption HILIC Chromatography." *Metabolomics : Official journal of the Metabolomic Society* 20(6): 114.
- Deneka, Dawid et al. 2021. "Allosteric Modulation of LRRC8 Channels by Targeting Their Cytoplasmic Domains." *Nature Communications* 12(1): 1–14.
- Diaz, George A. et al. 1999. "Mutations in a New Gene Encoding a Thiamine Transporter Cause Thiamine- Responsive Megaloblastic Anaemia Syndrome." *Nature Genetics* 22(3): 309–12.
- Downing, Kenneth H., and Robert M. Glaeser. 2008. "Restoration of Weak Phase-Contrast Images Recorded with a High Degree of Defocus: The 'Twin Image' Problem Associated with CTF Correction." *Ultramicroscopy* 108(9): 921–28.
- Drew, David, Rachel A. North, Kumar Nagarathinam, and Mikio Tanabe. 2021. "Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS)." *Chemical Reviews* 121(9): 5289–5335.
- Dubochet, Jacques. 1988. "Cryo-Electron Microscopy of Vitrified Specimens." *Quarterly Reviews of Biophysics* 21(2): 129–228.
- Dudeja, Pradeep K. et al. 2001. "Mechanism of Thiamine Uptake by Human Jejunal Brush-Border Membrane Vesicles." *American Journal of Physiology - Cell Physiology* 281(3 50-3): 786–92.
- Dutta, Binita et al. 1999. "Cloning of the Human Thiamine Transporter, a Member of the Folate Transporter Family." *Journal of Biological Chemistry* 274(45): 31925–29.
- Enjo, Fumio et al. 1997. "Isolation and Characterization of a Thiamin Transport Gene, THI10, from *Saccharomyces Cerevisiae*." *Journal of Biological Chemistry* 272(31): 19165–70. <http://dx.doi.org/10.1074/jbc.272.31.19165>.
- Foulon, Veerle et al. 2005. "Breakdown of 2-Hydroxylated Straight Chain Fatty Acids via Peroxisomal 2-Hydroxyphytanoyl-CoA Lyase: A Revised Pathway for the  $\alpha$ -Oxidation of Straight Chain Fatty Acids." *Journal of Biological Chemistry* 280(11): 9802–12.
- Frank, Joachim. 2009. "Single-Particle Reconstruction of Biological Macromolecules in Electron Microscopy-30 Years." *Quarterly Reviews of Biophysics* 42(3): 139–58.
- Fraser, Jamie L. et al. 2014. "Thiamine pyrophosphokinase deficiency causes a Leigh Disease like phenotype in a sibling pair: identification through whole exome sequencing and management strategies." *Molecular Genetics and Metabolism Reports* 1: 66-70.
- Freeze, Hudson H., and Christian Kranz. 2010. "Endoglycosidase and Glycoamidase Release of N-Linked Glycans." *Current Protocols in Molecular Biology* 2010(January): 1–25.

- Fridy, Peter C, Michael P Rout, and Natalia E Ketaren. 2024. “Nanobodies : From High-Throughput Identifi cation to Therapeutic Development Authors.” *Mol Cell Proteomics* 23(12): 100865.
- Fullam, Elizabeth et al. 2012. “Structure and Function of the Transketolase from Mycobacterium Tuberculosis and Comparison with the Human Enzyme.” *Open Biology* 2(JANUARY).
- Gabriel, Florian et al. 2024. “Structural Basis of Substrate Transport and Drug Recognition by the Human Thiamine Transporter SLC19A3.” *bioRxiv*: 2024.03.11.584396.
- . 2024. “Structural Basis of Substrate Transport and Drug Recognition by the Human Thiamine Transporter SLC19A3.” *Nature Communications*.
- Gallagher, John R., Alexander J. Kim, Neetu M. Gulati, and Audrey K. Harris. 2019. “Negative-Stain Transmission Electron Microscopy of Molecular Complexes for Image Analysis by 2D Class Averaging.” *Current Protocols in Microbiology* 54(1): 1–29.
- Garman, Elspeth F. 2010. “Radiation Damage in Macromolecular Crystallography: What Is It and Why Should We Care?” *Acta Crystallographica Section D: Biological Crystallography* 66(4): 339–51.
- Geertsma, Eric R. 2013. “FX Cloning: A Versatile High-Throughput Cloning System for Characterization of Enzyme Variants.” *Methods in Molecular Biology* 978: 133–48.
- Gelová, Zuzana et al. 2024. “Protein Binder Toolbox for Studies of Solute Carrier Transporters.” *Journal of Molecular Biology*: 168665.
- Giacomini, Marilyn M. et al. 2017. “Interaction of 2,4-Diaminopyrimidine-Containing Drugs Including Fedratinib and Trimethoprim with Thiamine Transporters.” *Drug Metabolism and Disposition* 45(1): 76–85.
- Gomes, Filomena, Gilles Bergeron, Megan W. Bourassa, and Philip R. Fischer. 2021. “Thiamine Deficiency Unrelated to Alcohol Consumption in High-Income Countries: A Literature Review.” *Annals of the New York Academy of Sciences* 1498(1): 46–56.
- Götzke, Hansjörg et al. 2019. “The ALFA-Tag Is a Highly Versatile Tool for Nanobody-Based Bioscience Applications.” *Nature Communications* 10(1): 1–12.
- Gowda, Vykuntaraju K., Varunvenkat M. Srinivasan, Kapil Jehta, and Maya D. Bhat. 2019. “Bilateral Striatal Necrosis with Polyneuropathy with a Novel SLC25A19 (Mitochondrial Thiamine Pyrophosphate Carrier OMIM\*606521) Mutation: Treatable Thiamine Metabolic Disorder-A Report of Two Indian Cases.” *Neuropediatrics* 50(5): 313–17.
- Greenberg, Andrew S., and Colleagues. 1995. “A New Antigen Receptor Gene Family That Undergoes Rearrangement and Extensive Somatic Diversification in Sharks.” : 168–73.
- Hall, Justin. 2019. “A Simple Model for Determining Affinity from Irreversible Thermal Shifts.” *Protein Science* 28(10): 1880–87.
- Helliwell, Katherine E., Glen L. Wheeler, and Alison G. Smith. 2013. “Widespread Decay of Vitamin-Related Pathways: Coincidence or Consequence?” *Trends in Genetics* 29(8): 469–78.
- Hillier, James et al. 2024. “Structural Insights into Frizzled3 through Nanobody Modulators.” *Nature Communications* 15(1): 1–13.
- Holton, James M. 2009. “A Beginner’s Guide to Radiation Damage.” *Journal of Synchrotron Radiation* 16(2): 133–42.
- Huang, Wenjie et al. 2019. “Reduced Thiamine Binding Is a Novel Mechanism for TPK Deficiency Disorder.” *Molecular Genetics and Genomics* 294(2): 409–16.
- Huang, Zhiwei et al. 2012. “Discovering Thiamine Transporters as Targets of Chloroquine Using a

- Novel Functional Genomics Strategy.” *PLoS Genetics* 8(11): 4–13.
- Ishida, Nobuhiro, and Masao Kawakita. 2004. “Molecular Physiology and Pathology of the Nucleotide Sugar Transporter Family (SLC35).” *Pflugers Archiv European Journal of Physiology* 447(5): 768–75.
- Jaehme, Michael et al. 2018. “PnuT Uses a Facilitated Diffusion Mechanism for Thiamine Uptake.” *Journal of General Physiology* 150(1): 41–50.
- Jaehme, Michael, and Dirk Jan Slotboom. 2015. “Diversity of Membrane Transport Proteins for Vitamins in Bacteria and Archaea.” *Biochimica et Biophysica Acta - General Subjects* 1850(3): 565–76.
- Jumper, John et al. 2021. “Highly Accurate Protein Structure Prediction with AlphaFold.” *Nature* 596(7873): 583–89. <http://dx.doi.org/10.1038/s41586-021-03819-2>.
- Jurgenson, Christopher T., Tadhg P. Begley, and Steven E. Ealick. 2009. “The Structural and Biochemical Foundations of Thiamin Biosynthesis.” *Annual Review of Biochemistry* 78: 569–603.
- Karlsson, Max et al. 2021. “A Single-Cell Type Transcriptomics Map of Human Tissues.” *Science Advances* 7(31): 1–9.
- Kato, Koji et al. 2015. “Involvement of Organic Cation Transporters in the Clearance and Milk Secretion of Thiamine in Mice.” *Pharmaceutical Research* 32(7): 2192–2204.
- Kimanius, Dari et al. 2024. “Data-Driven Regularization Lowers the Size Barrier of Cryo-EM Structure Determination.” *Nature Methods* 21(7): 1216–21.
- Kohl, Helmut, and Ludwig Reimer. 2008. *Transmission Electron Microscopy: Physics of Image Formation*. 5th ed. Springer Science & Business Media.
- Kono, Satoshi et al. 2009. “Mutations in a Thiamine-Transporter Gene and Wernicke’s-like Encephalopathy.” *New England Journal of Medicine* 360(17): 1792–94.
- Kotov, Vadim et al. 2023. “Plasticity of the Binding Pocket in Peptide Transporters Underpins Promiscuous Substrate Recognition.” *Cell Reports* 42(8).
- Kühlbrandt, Werner. 2014. “The Resolution Revolution.” *Science* 343(6178): 1443–44.
- . 2015. “Structure and Function of Mitochondrial Membrane Protein Complexes.” *BMC Biology* 13(1).
- Kung, Jennifer E et al. 2024. “Disulfide Constrained Fabs Overcome Target Size Limitation for High-Resolution Single-Particle Cryo-EM.” *bioRxiv* (c): 2024.05.10.593593.
- Labay, Valentina et al. 1999. “Mutations in SLC19A2 Cause Thiamine-Responsive Megaloblastic Anaemia Associated with Diabetes Mellitus and Deafness.” *Nature Genetics* 22(3): 300–304.
- Lagarde, William H., Louis E. Underwood, Billie M. Moats-Staats, and Ali S. Calikoglu. 2004. “Novel Mutation in TheSLC19A2 Gene in an African-American Female with Thiamine-Responsive Megaloblastic Anemia Syndrome.” *American Journal of Medical Genetics* 125A(3): 299–305.
- Lander, Gabriel C., and Robert M. Glaeser. 2021. “Conquer by Cryo-EM without Physically Dividing.” *Biochemical Society Transactions* 49(5): 2287–98.
- Langlais, Philip J. 1995. “Alcohol-Related Thiamine Deficiency.” *Alcohol Health & Research World* 19(2): 113–21.
- Liang, Xiaomin et al. 2015. “Metformin Is a Substrate and Inhibitor of the Human Thiamine

- Transporter, THTR-2 (SLC19A3).” *Molecular Pharmaceutics* 12(12): 4301–10.
- Lieberman, M. C., E. Tartaglino, J. C. Fleming, and E. J. Neufeld. 2006. “Deletion of SLC19A2, the High Affinity Thiamine Transporter, Causes Selective Inner Hair Cell Loss and an Auditory Neuropathy Phenotype.” *JARO - Journal of the Association for Research in Otolaryngology* 7(3): 211–17.
- Liu, Jing Yuan, David E. Timm, and Thomas D. Hurley. 2006. “Pyriithiamine as a Substrate for Thiamine Pyrophosphokinase.” *Journal of Biological Chemistry* 281(10): 6601–7.
- Liu, Yuxi, Duc T. Huynh, and Todd O. Yeates. 2019. “A 3.8 Å Resolution Cryo-EM Structure of a Small Protein Bound to an Imaging Scaffold.” *Nature Communications* 10(1): 1–7.
- Ma, Haiyun, Xinyu Jia, Kaiming Zhang, and Zhaoming Su. 2022. “Cryo-EM Advances in RNA Structure Determination.” *Signal Transduction and Targeted Therapy* 7(1): 1–6.
- Mahajan, Abhimanyu, and Sidiropoulos, Christos. 2017. “TPK1 mutation induced childhood onset idiopathic generalized dystonia: Report of a rare mutation and effect of deep brain stimulation.” *Journal of the Neurological Sciences* 15:376:42-43.
- Martynowycz, Michael W. et al. 2021. “Benchmarking the Ideal Sample Thickness in Cryo-EM.” *Proceedings of the National Academy of Sciences of the United States of America* 118(49): 1–7.
- Mayr, Johannes A. et al. 2011. “Thiamine Pyrophosphokinase Deficiency in Encephalopathic Children with Defects in the Pyruvate Oxidation Pathway.” *American Journal of Human Genetics* 89(6): 806–12.
- Mori, Keisuke, Tatsuro Naganuma, and Akio Kihara. 2023. “Role of 2-Hydroxy Acyl-CoA Lyase HACL2 in Odd-Chain Fatty Acid Production via  $\alpha$ -Oxidation in Vivo.” *Molecular Biology of the Cell* 34(9): 1–10.
- Mrowicka, Małgorzata, Jerzy Mrowicki, Grzegorz Dragan, and Ireneusz Majsterek. 2023. “The Importance of Thiamine (Vitamin B1) in Humans.” *Bioscience Reports* 43(10): 1–18.
- Mukherjee, Somnath et al. 2020. “Synthetic Antibodies against BRIL as Universal Fiducial Marks for Single-particle CryoEM Structure Determination of Membrane Proteins.” *Nature Communications* 11(1): 1–14.
- Muyldermans, Serge. 2013. “Nanobodies : Natural Single-Domain Antibodies.”
- Nabokina, Svetlana M. et al. 2014. “Molecular Identification and Functional Characterization of the Human Colonic Thiamine Pyrophosphate Transporter.” *Journal of Biological Chemistry* 289(7): 4405–16.
- Naydenova, Katerina, Peipei Jia, and Christopher J. Russo. 2020. “Cryo-EM with Sub-1 Å Specimen Movement.” *Science* 370(6513): 223–26.
- Neselu, Kasahun et al. 2023. “Measuring the Effects of Ice Thickness on Resolution in Single Particle Cryo-EM.” *Journal of Structural Biology: X* 7(January): 100085.
- Nevers, Yannis, Natasha M. Glover, Christophe Dessimoz, and Odile Lecompte. 2023. “Protein Length Distribution Is Remarkably Uniform across the Tree of Life.” *Genome Biology* 24(1): 1–20.
- NIH. 2024. “Thiamine - Fact Sheet for Health Professionals.” <https://ods.od.nih.gov/factsheets/Thiamin-HealthProfessional/> (November 19, 2024).
- Nogales, Eva, and Sjors H.W. Scheres. 2015. “Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity.” *Molecular Cell* 58(4): 677–89.
- Oishi, Takumi et al. 2023. “Contributions of Histone Tail Clipping and Acetylation in Nucleosome

- Transcription by RNA Polymerase II.” *Nucleic Acids Research* 51(19): 10364–74.
- Oudman, Erik et al. 2021. “Wernicke’s Encephalopathy in Crohn’s Disease and Ulcerative Colitis.” *Nutrition* 86: 111182.
- Ozand, Pinar T. et al. 1998. “Biotin-Responsive Basal Ganglia Disease: A Novel Entity.” *Brain* 121(7): 1267–79.
- Pardanani, Animesh et al. 2015. “Safety and Efficacy of Fedratinib in Patients with Primary or Secondary Myelofibrosis: A Randomized Clinical Trial.” *JAMA Oncology* 1(5): 643–51.
- Pardon, Els et al. 2014. “A General Protocol for the Generation of Nanobodies for Structural Biology.” *Nature Protocols* 9(3): 674–93.
- Pettersen, Eric F. et al. 2021. “UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers.” *Protein Science* 30(1): 70–82.
- Pidathala, Shabareesh et al. 2023. “Mechanisms of Neurotransmitter Transport and Drug Inhibition in Human VMAT2.” *Nature* 623(June).
- Pintilie, Grigore et al. 2020. “Measurement of Atom Resolvability in Cryo-EM Maps with Q-Scores.” *Nature Methods* 17(3): 328–34.
- Pintilie, Grigore, and Wah Chiu. 2021. “Validation, Analysis and Annotation of Cryo-EM Structures.” *Acta Crystallographica Section D: Structural Biology* 77: 1142–52.
- Punjani, Ali, John L. Rubinstein, David J. Fleet, and Marcus A. Brubaker. 2017. “CryoSPARC: Algorithms for Rapid Unsupervised Cryo-EM Structure Determination.” *Nature Methods* 14(3): 290–96.
- Punjani, Ali, Haowei Zhang, and David J. Fleet. 2020. “Non-Uniform Refinement: Adaptive Regularization Improves Single-Particle Cryo-EM Reconstruction.” *Nature Methods* 17(12): 1214–21.
- Qu, Qianhui et al. 2024. “Preprint: Substrate Transport and Drug Interaction of Human Thiamine Transporters SLC19A2 / A3.” *Research Square*: 1–25.
- Quistgaard, Esben M., Christian Löw, Fatma Guettou, and Pär Nordlund. 2016. “Understanding Transport by the Major Facilitator Superfamily (MFS): Structures Pave the Way.” *Nature Reviews Molecular Cell Biology* 17(2): 123–32.
- Rahman, Md Shafiqur et al. 2023. “Structure-Guided Development of Bivalent Aptamers Blocking SARS-CoV-2 Infection.” *Molecules* 28(12): 1–21.
- Rajgopal, Arun, Antoinette Edmondson, I. David Goldman, and Rongbao Zhao. 2001. “SLC19A3 Encodes a Second Thiamine Transporter ThTr2.” *Biochimica et Biophysica Acta - Molecular Basis of Disease* 1537(3): 175–78.
- Rensen, Yvonne C.M., Erik Oudman, Joukje M. Oosterman, and Roy P.C. Kessels. 2021. “Confabulations in Alcoholic Korsakoff’s Syndrome: A Factor Analysis of the Nijmegen–Venray Confabulation List.” *Assessment* 28(6): 1545–55.
- Rosenberg, Marjorie J. et al. 2002. “Mutant Deoxynucleotide Carrier Is Associated with Congenital Microcephaly.” *Nature Genetics* 32(1): 175–79.
- Rösner, Benedikt et al. 2020. “Soft X-Ray Microscopy with 7 Nm Resolution.” *Optica* 7(11): 1602.
- Rothenberger, Sylvia et al. 2022. “The Trispecific DARPIn Ensovibep Inhibits Diverse SARS-CoV-2 Variants.” *Nature Biotechnology* 40(12): 1845–54.
- Ruprecht, Jonathan J., and Edmund R.S. Kunji. 2020. “The SLC25 Mitochondrial Carrier Family: Structure and Mechanism.” *Trends in Biochemical Sciences* 45(3): 244–58.

- Russo, C. J., and R. F. Egerton. 2019. "Damage in Electron Cryomicroscopy: Lessons from Biology for Materials Science." *MRS Bulletin* 44(12): 935–41.
- Said, Hamid M., Krishnaswamy Balamurugan, Veedamali S. Subramanian, and Jonathan S. Marchant. 2004. "Expression and Functional Contribution of HTHTR-2 in Thiamin Absorption in Human Intestine." *American Journal of Physiology - Gastrointestinal and Liver Physiology* 286(3 49-3): 491–98.
- Sakuragi, Takaharu et al. 2021. "The Tertiary Structure of the Human Xkr8–Basigin Complex That Scrambles Phospholipids at Plasma Membranes." *Nature Structural and Molecular Biology* 28(10): 825–34.
- Sambon, Margaux et al. 2022. "Product Inhibition of Mammalian Thiamine Pyrophosphokinase Is an Important Mechanism for Maintaining Thiamine Diphosphate Homeostasis." *Biochimica et Biophysica Acta - General Subjects* 1866(3): 130071.
- Sampedro Vallina, Nestor et al. 2023. "RNA Origami Scaffolds Facilitate Cryo-EM Characterization of a Broccoli-Pepper Aptamer FRET Pair." *Nucleic Acids Research* 51(9): 4613–24.
- Schänzer, Anne et al. 2014. "Stress-Induced Upregulation of SLC19A3 Is Impaired in Biotin-Thiamine- Responsive Basal Ganglia Disease." *Brain Pathology* 24(3): 270–79.
- Scheres, Sjors H.W. 2010. 482 Methods in Enzymology *Classification of Structural Heterogeneity by Maximum-Likelihood Methods*. 1st ed. Elsevier Inc.
- . 2012. "RELION: Implementation of a Bayesian Approach to Cryo-EM Structure Determination." *Journal of Structural Biology* 180(3): 519–30.
- . 2014. "Beam-Induced Motion Correction for Sub-Megadalton Cryo-EM Particles." *eLife* 3: e03665.
- Seo, Ja-young, and Jeong-Hwa Choi. 2021. "Genetic Variations in Thiamin Transferase SLC35F3 and the Risk of Hypertension in Koreans ." *Clinical Nutrition Research* 10(2): 140.
- Shahsavari, Azadeh et al. 2021. "Structural Insights into the Inhibition of Glycine Reuptake." *Nature* 591(7851): 677–81.
- Singleton, Charles K. 1997. "Identification and Characterization of the Thiamine Transporter Gene of *Saccharomyces Cerevisiae*." *Gene* 199(1–2): 111–21.
- Slotboom, Dirk J. 2014. "Structural and Mechanistic Insights into Prokaryotic Energy-Coupling Factor Transporters." *Nature Reviews Microbiology* 12(2): 79–87.
- Smith, Taryn J. et al. 2021. "Thiamine Deficiency Disorders: A Clinical Perspective." *Annals of the New York Academy of Sciences* 1498(1): 9–28.
- Spiegel, Ronen et al. 2009. "SLC25A19 Mutation as a Cause of Neuropathy and Bilateral Striatal Necrosis." *Annals of Neurology* 66(3): 419–24.
- Spiestersbach, Lea. 2022. "Structural and Biophysical Insights into Molecular Transport and Processing of Thiamine in Humans." University of Regensburg.
- Stanfield, Robyn L., Helen Dooley, Martin F. Flajnik, and Ian A. Wilson. 2004. "Crystal Structure of a Shark Single-Domain Antibody V Region in Complex with Lysozyme." *Science* 305(5691): 1770–73.
- Subramanian, Veedamali S., Jonathan S. Marchant, and Hamid M. Said. 2006. "Biotin-Responsive Basal Ganglia Disease-Linked Mutations Inhibit Thiamine Transport via HTHTR2: Biotin Is Not a Substrate for HTHTR2." *American Journal of Physiology - Cell Physiology* 291(5): 851–59.
- Sun, Wen et al. 2020. "A Unique Hormonal Recognition Feature of the Human Glucagon-like



- Peptide-2 Receptor.” *Cell Research* 30(12): 1098–1108.
- Suzuki, Kaoru et al. 2017. “High-Dose Thiamine Prevents Brain Lesions and Prolongs Survival of Slc19a3-Deficient Mice.” *PLoS ONE* 12(6): 1–17.
- Suzuki, Yoshihito et al. 2005. “Preparation of Functional Single-Chain Antibodies against Bioactive Gibberellins by Utilizing Randomly Mutagenized Phage-Display Libraries.” *Bioscience, Biotechnology and Biochemistry* 69(3): 610–19.
- Teng, Grace, and F. Nina Papavasiliou. 2007. “Immunoglobulin Somatic Hypermutation.” *Annual Review of Genetics* 41: 107–20.
- Thomson, Allan D., and E. Jane Marshall. 2006. “The Natural History and Pathophysiology of Wernicke’s Encephalopathy and Korsakoff’s Psychosis.” *Alcohol and Alcoholism* 41(2): 151–58.
- Thul, Peter J. et al. 2017. “A Subcellular Map of the Human Proteome.” *Science* 356(6340).
- Timm, David E., Jingyuan Liu, L. J. Baker, and Robert A. Harris. 2001. “Crystal Structure of Thiamin Pyrophosphokinase.” *Journal of Molecular Biology* 310(1): 195–204.
- Trinco, Gianluca et al. 2021. “Kinetic Mechanism of Na<sup>+</sup>-Coupled Aspartate Transport Catalyzed by GltTk.” *Communications Biology* 4(1).
- Uchański, Tomasz et al. 2021. “Megabodies Expand the Nanobody Toolkit for Protein Structure Determination by Single-Particle Cryo-EM.” *Nature Methods* 18(1): 60–68.
- Uhlén, Mathias et al. 2015. “Tissue-Based Map of the Human Proteome.” *Science* 347(6220).
- Vora, Bianca et al. 2020. “Drug-Nutrient Interactions: Discovering Prescription Drug Inhibitors of the Thiamine Transporter ThTR-2 (SLC19A3).” *American Journal of Clinical Nutrition* 111(1): 110–21.
- Walsh, Christopher T., and Yi Tang. 2019. *The Chemical Biology of Human Vitamins*. Royal Society of Chemistry.
- Wang, Jiaping et al. 2021. “Report of the Largest Chinese Cohort With SLC19A3 Gene Defect and Literature Review.” *Frontiers in Genetics* 12(July): 1–9.
- Webb, Eric, Kathy Claas, and Diana Downs. 1998. “ThiBPQ Encodes an ABC Transporter Required for Transport of Thiamine and Thiamine Pyrophosphate in *Salmonella Typhimurium*.” *Journal of Biological Chemistry* 273(15): 8946–50.
- Wentink, Koen, Christos Gogou, and Dimphna H. Meijer. 2022. “Putting on Molecular Weight: Enabling Cryo-EM Structure Determination of Sub-100-KDa Proteins.” *Current Research in Structural Biology* 4(August): 332–37.
- Wright, Nicholas J. et al. 2022. “Methotrexate Recognition by the Human Reduced Folate Carrier SLC19A1.” *Nature* 609(February).
- Wu, Xudong, and Tom A. Rapoport. 2021. “Cryo-EM Structure Determination of Small Proteins by Nanobody-Binding Scaffolds (Legobodies).” *Proceedings of the National Academy of Sciences of the United States of America* 118(41): 1–9.
- Xu, Yan et al. 2014. “Structures of Bacterial Homologues of SWEET Transporters in Two Distinct Conformations.” *Nature* 515(7527): 448–52.
- Yamashiro, Takahiro, Tomoya Yasujima, Hamid M. Said, and Hiroaki Yuasa. 2020. “PH-Dependent Pyridoxine Transport by SLC19A2 and SLC19A3: Implications for Absorption in Acidic Microclimates.” *Journal of Biological Chemistry* 295(50): 16998–8.
- Yip, Ka Man et al. 2020. “Atomic-Resolution Protein Structure Determination by Cryo-EM.” *Nature* 587(7832): 157–61.

- Yuan, Yafei et al. 2022. “Cryo-EM Structure of Human Glucose Transporter GLUT4.” *Nature Communications* 13(1): 1–8.
- Zang, Xiao Li et al. 2016. “Association of a SNP in SLC35F3 Gene with the Risk of Hypertension in a Chinese Han Population.” *Frontiers in Genetics* 7(JUN): 1–6.
- Zeng, Wen Qi et al. 2005. “Biotin-Responsive Basal Ganglia Disease Maps to 2q36.3 and Is Due to Mutations in SLC19A3.” *American Journal of Human Genetics* 77(1): 16–26.
- Zhang, Kaiming et al. 2019. “Cryo-EM Structure of a 40 KDa SAM-IV Riboswitch RNA at 3.7 Å Resolution.” *Nature Communications* 10(1): 6–11.
- Zhang, Kuixing et al. 2014. “Genetic Implication of a Novel Thiamine Transporter in Human Hypertension.” *Journal of the American College of Cardiology* 63(15): 1542–55.
- Zhang, Qiang et al. 2014. “The Janus Kinase 2 Inhibitor Fedratinib Inhibits Thiamine Uptake: A Putative Mechanism for the Onset of Wernicke’s Encephalopathy.” *Drug Metabolism and Disposition* 42(10): 1656–62.
- Zhang, Qixiang et al. 2022. “Recognition of Cyclic Dinucleotides and Folates by Human SLC19A1.” *Nature* 612(7938): 170–76.
- Zhang, Yongwei, W Uwe Werling, and Winfried Edelmann. 2014. “Seamless Ligation Cloning Extract (SLiCE) Cloning Method Yongwei.” *Methods in Molecular Biology* 1116: 235–244.
- Zhao, Rongbao, and David Goldman. 2008. “Folate and Thiamine Transporters Mediated by Facilitative Carriers (SLC19A1-3 and SLC46A1) and Folate Receptors.” *Bone* 23(1): 1–7.
- Zheng, Shawn Q. et al. 2017. “MotionCor2: Anisotropic Correction of Beam-Induced Motion for Improved Cryo-Electron Microscopy.” *Nature Methods* 14(4): 331–32.
- Zhou, Bing Rui et al. 2019. “Atomic Resolution Cryo-EM Structure of a Native-like CENP-A Nucleosome Aided by an Antibody Fragment.” *Nature Communications* 10(1): 1–7.
- Zimmermann, Iwan et al. 2018. “Synthetic Single Domain Antibodies for the Conformational Trapping of Membrane Proteins.” *eLife* 7: 1–32.
- . 2020. “Generation of Synthetic Nanobodies against Delicate Proteins.” *Nature Protocols* 15(May).
- Zivanov, Jasenko, Takanori Nakane, and Sjors H.W. Scheres. 2019. “A Bayesian Approach to Beam-Induced Motion Correction in Cryo-EM Single-Particle Analysis.” *IUCrJ* 6: 5–17.

# Appendix

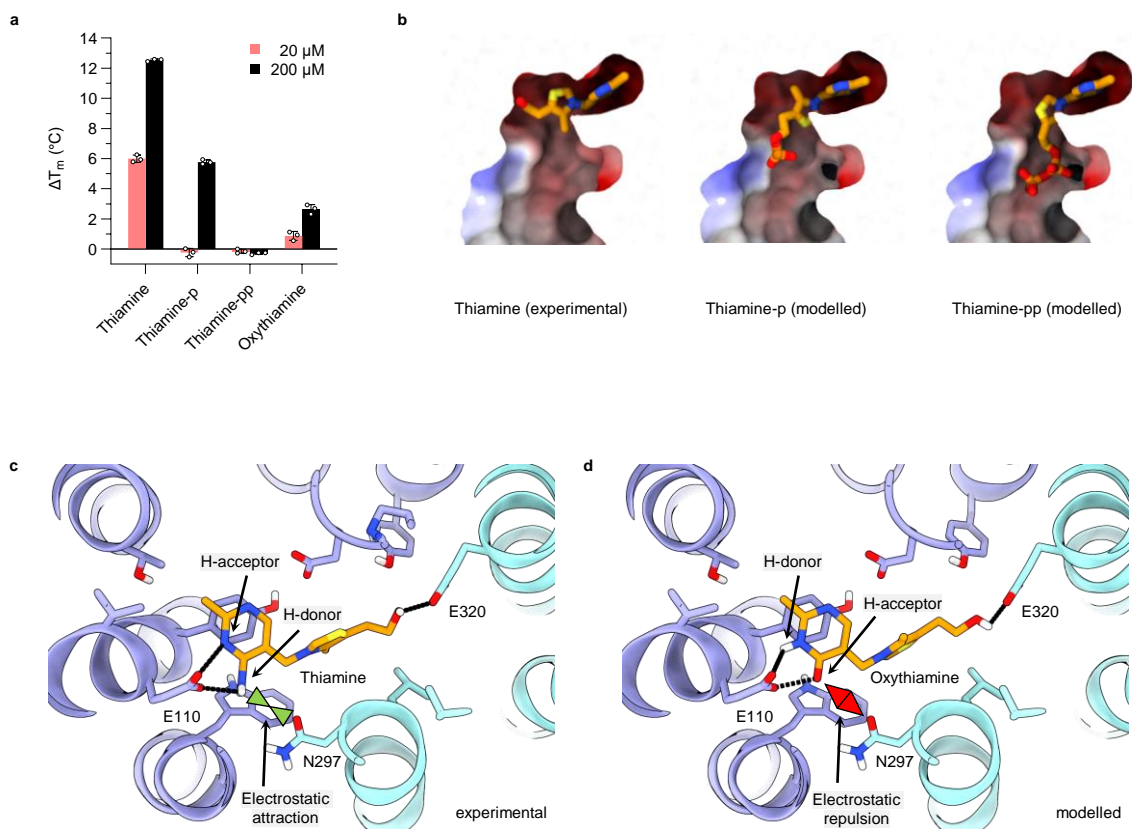
## Content

- Appendix 1: Supplementary Figures
- Appendix 2: List of Hazardous Substances
- Appendix 3: Materials

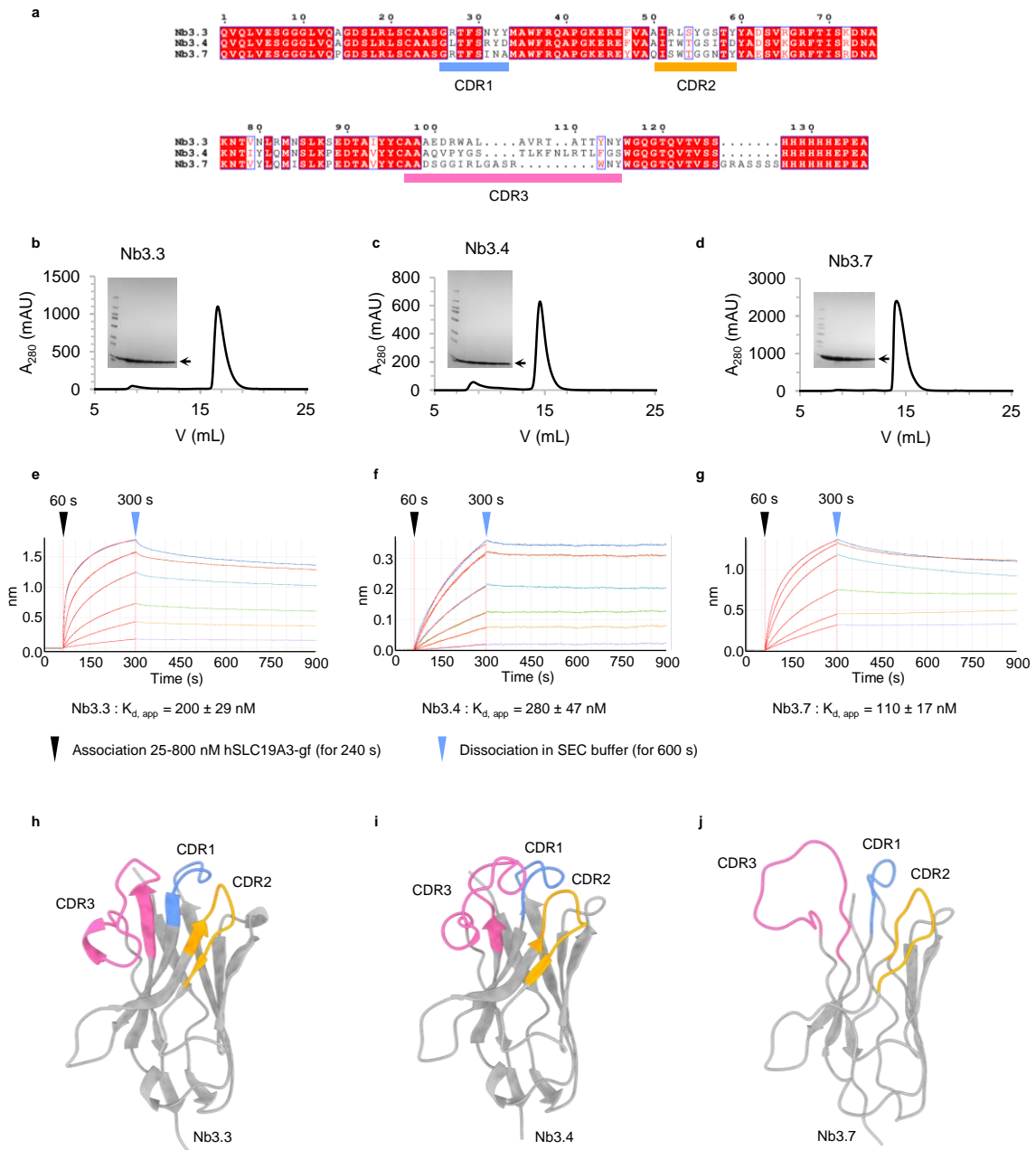
## **Appendix 1: Supplementary Figures**

The Supplementary Figures and the corresponding figure legends were directly taken from my recent publication on the human SLC19A3 (Gabriel et al., 2024, *Nature Communications*).

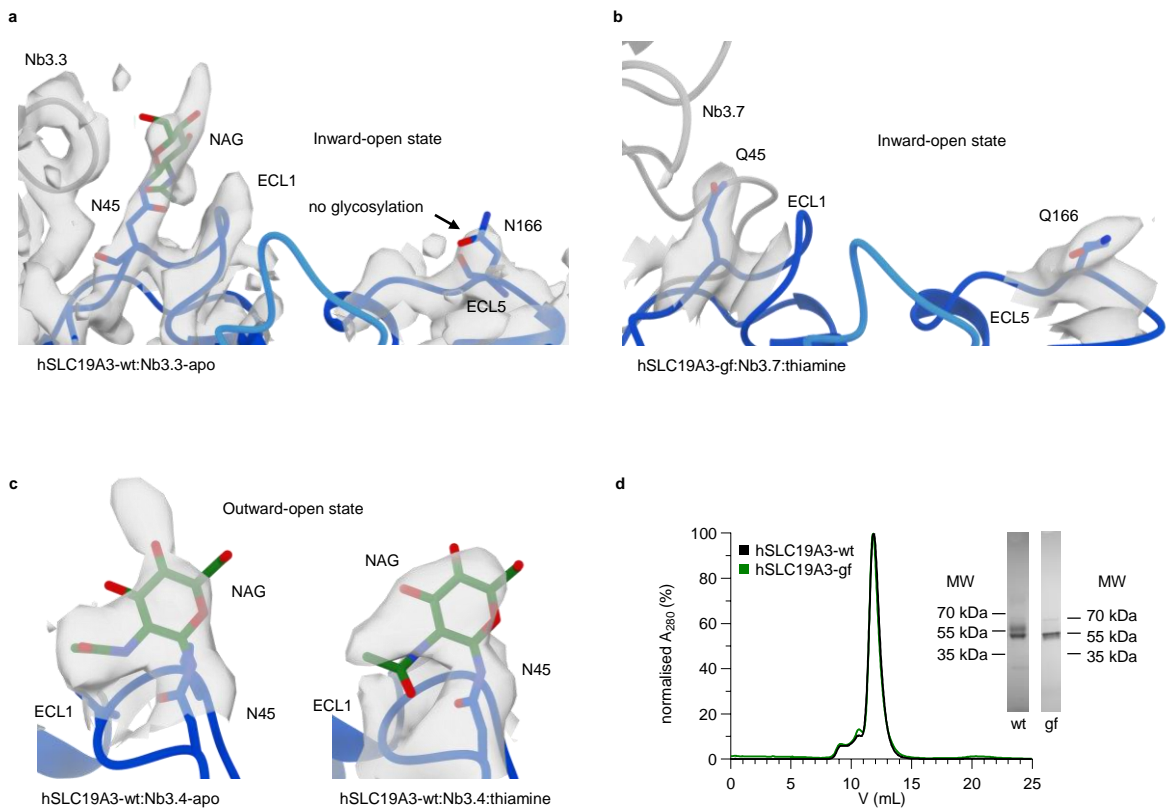
## Supplementary Figures



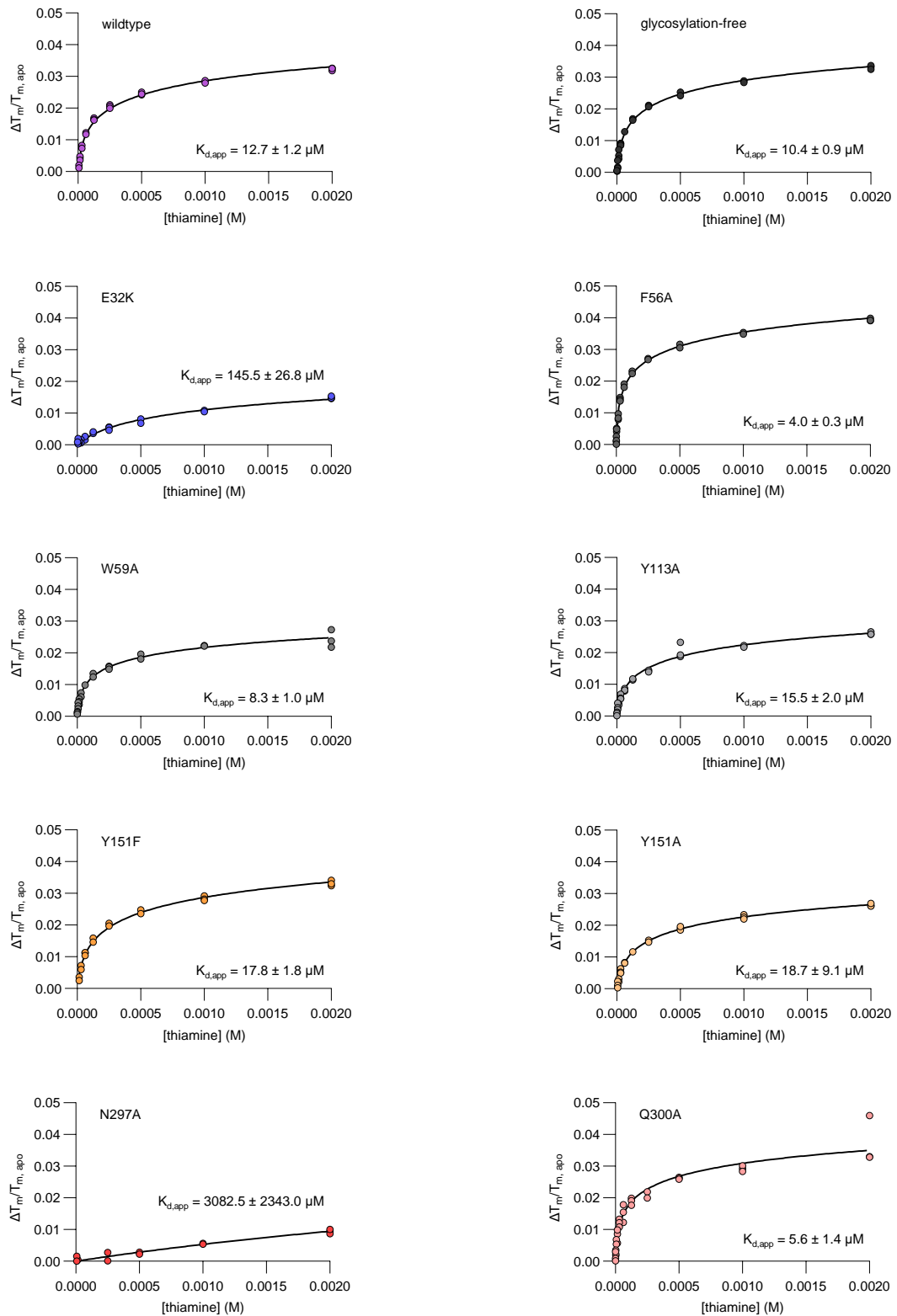
**Supplementary Fig. 1: Recognition of thiamine.** **a**, Thermal shifts induced by the indicated compounds (fluorescence readout of thermal unfolding curve at 350 nm, mean  $\pm$  s.d.,  $n = 3$ ). While thiamine induces a strong stabilisation of hSLC19A3, its phosphorylated derivatives thiamine monophosphate (thiamine-p) shows a strongly reduced affinity for the transporter; and thiamine pyrophosphate (thiamine-pp) does not stabilise the transporter at all. The closely related derivative oxythiamine also stabilises the transporter only weakly. **b**, Structural model for the interaction of the phosphorylated derivatives of thiamine with hSLC19A3 in the inward-open state. The left panel shows the experimentally determined structure of hSLC19A3:Nb3.7:thiamine. The two panels to the right highlight thiamine-p and -pp modelled in the same EM density map using Coot. These models demonstrate the electrostatic mismatch between the negatively charged vestibule of the transporter and the negatively charged phosphate moieties. Furthermore, the important hydrogen bond between the hydroxyethyl tail of thiamine and Glu320 is disrupted by the phosphorylation of thiamine. **c**, Coordination of thiamine in hSLC19A3 in the inward-open state (hSLC19A3:Nb3.7:thiamine). The green arrows indicate the electrostatic attraction between the positively polarised hydrogens of the amino group of thiamine and the side chain carbonyl oxygen of Asn297. **d**, Oxythiamine modelled in the same EM density map. While the hydrogen bonding pattern with the transporter is in principle satisfied, the type of hydrogen bond pairs is different and might be energetically less favourable. Additionally, the carbonyl oxygen of oxythiamine is in electrostatic conflict with the side chain of Asn297. This might explain the observed weak interaction of this compound with hSLC19A3. The protons shown in panels c and d have not been directly observed, but are only inferred from the most likely protonation state of the corresponding chemical moieties at neutral pH.



**Supplementary Fig. 2: Conformation-specific nanobodies against hSLC19A3.** **a**, Multiple sequence alignment of the hSLC19A3-specific nanobodies generated and used in this study. Residue numbering based on Nb3.3. Residue colouring based on conservation. The rearranged and hypermutated complementarity determining regions (CDR1-3) are clearly recognisable as the regions without sequence conservation. All nanobodies carry a C-terminal 6×His-EPEA tag. In Nb3.7, this tag is connected to the nanobody via a linker (GRASSSS). **b-d**, Purification of the nanobodies. Shown are size-exclusion chromatography (SEC) traces (absorbance at  $\lambda = 280$  nm) and SDS-PAGE gels of different fractions of the SEC peaks. **e-g**, Bilayer interferometry (Octet®) measurements of binding of hSLC19A3-gf to the respective nanobody immobilised on anti-penta-His sensors (Octet® HIS1 Biosensors). The sensors were incubated with different concentrations of hSLC19A3-gf; 25 nM (purple), 50 nM (yellow), 100 nM (green), 200 nM (light blue), 400 nM (orange), 800 nM (dark blue). Apparent binding affinities were determined by fitting the maximal response to a one-site binding model. **h-j**, Cryo-EM structures of the Nb3.3 and Nb3.4 when bound to hSLC19A3-wt and Nb3.7 bound to hSLC19A3-gf. The paratopes for the interaction with the transporter are formed by the highly variable CDRs (colour coded).

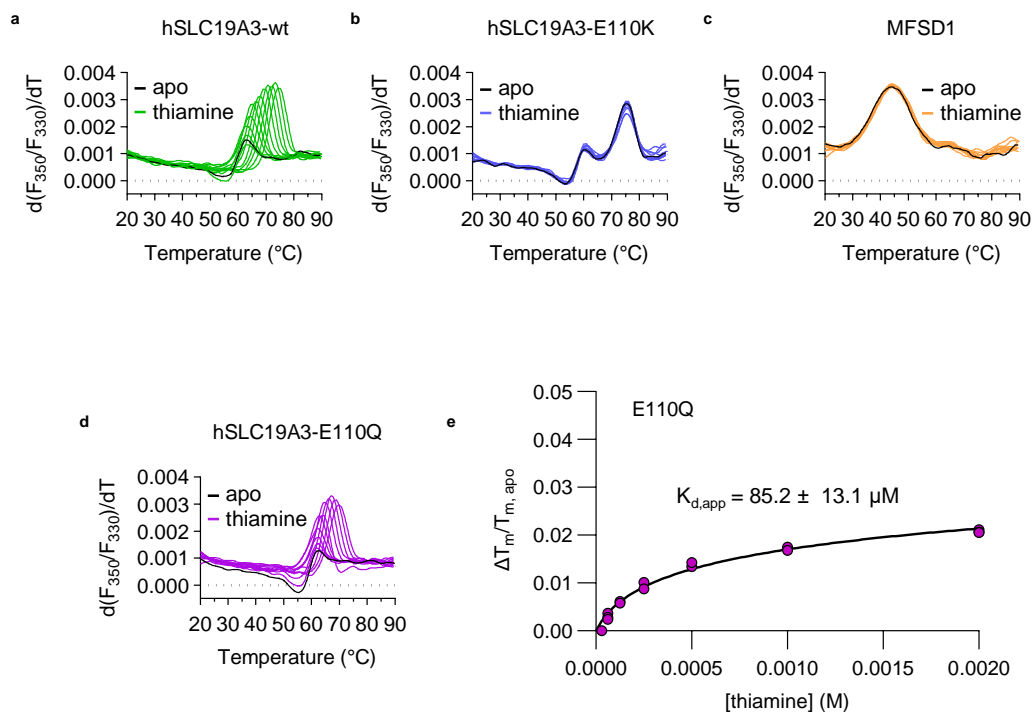


**Supplementary Fig. 3: N-linked glycosylation of hSLC19A3.** **a**, Structure model and cryo-EM density of hSLC19A3-wt:Nb3.3-apo. The wildtype transporter shows additional non-protein density extending from Asn45, but not from Asn166 (the two predicted N-glycosylation sites; map contour level = 1.01) indicative for a glycosylation site. **b**, For the hSLC19A3-gf double mutant (Asn45Gln, Asn166Gln), no extra density is visible. This confirms that the mutation of Asn45 to glutamine prevents N-linked glycosylation of the transporter (map contour level = 0.126). **c**, Glycan densities can be observed on Asn45 in the apo (left) and thiamine-bound (right) outward-open state of wildtype hSLC19A3 (map contour levels = 0.432 and 0.625, respectively). In these structures, there is likewise no extra density on Asn166, indicating that this residue is actually not glycosylated in recombinantly expressed hSLC19A3. **d**, Size-exclusion chromatography (SEC) trace of wildtype hSLC19A3 (wt) and the glycosylation-free hSLC19A3-gf double mutant (gf). The SEC traces are almost identical. SDS-PAGE (insets) show that the higher molecular weight band (slightly above 55 kDa) corresponding to the glycosylated transporter, is absent for hSLC19A3-gf.

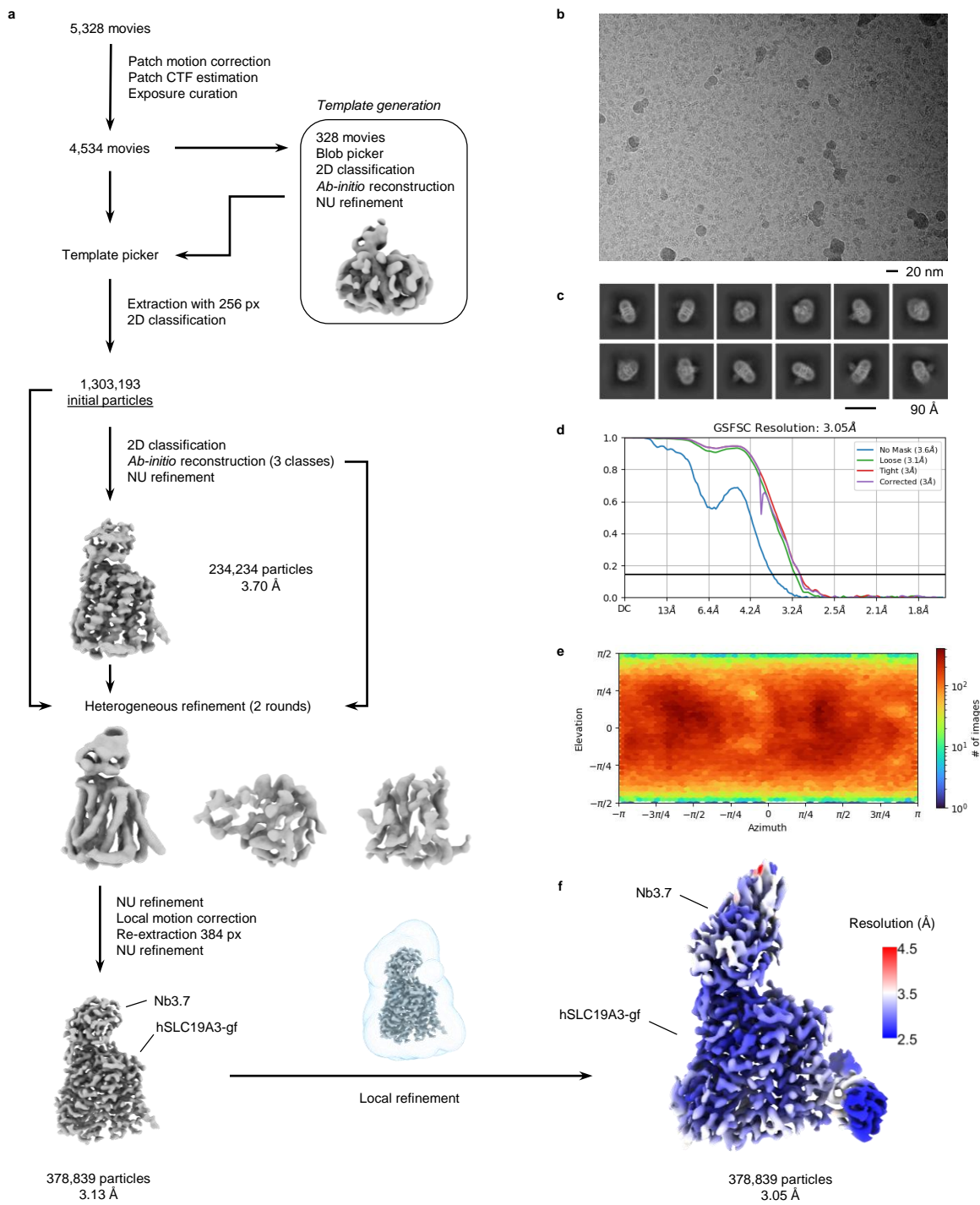


**Supplementary Fig. 4: Thiamine binding affinity of single point mutants of hSLC19A3.** The graphs highlight the measured thermostability shifts of the respective hSLC19A3 mutants in dependence of thiamine concentration (coloured circles,  $n = 3$ , independent measurements). The corresponding melting curves were read out based on the ratio of the fluorescence at 350 nm and 330 nm ( $F_{350}/F_{330}$ ). The depicted curves were fitted as described in the methods section to determine the apparent dissociation constants  $K_{d,app}$ .

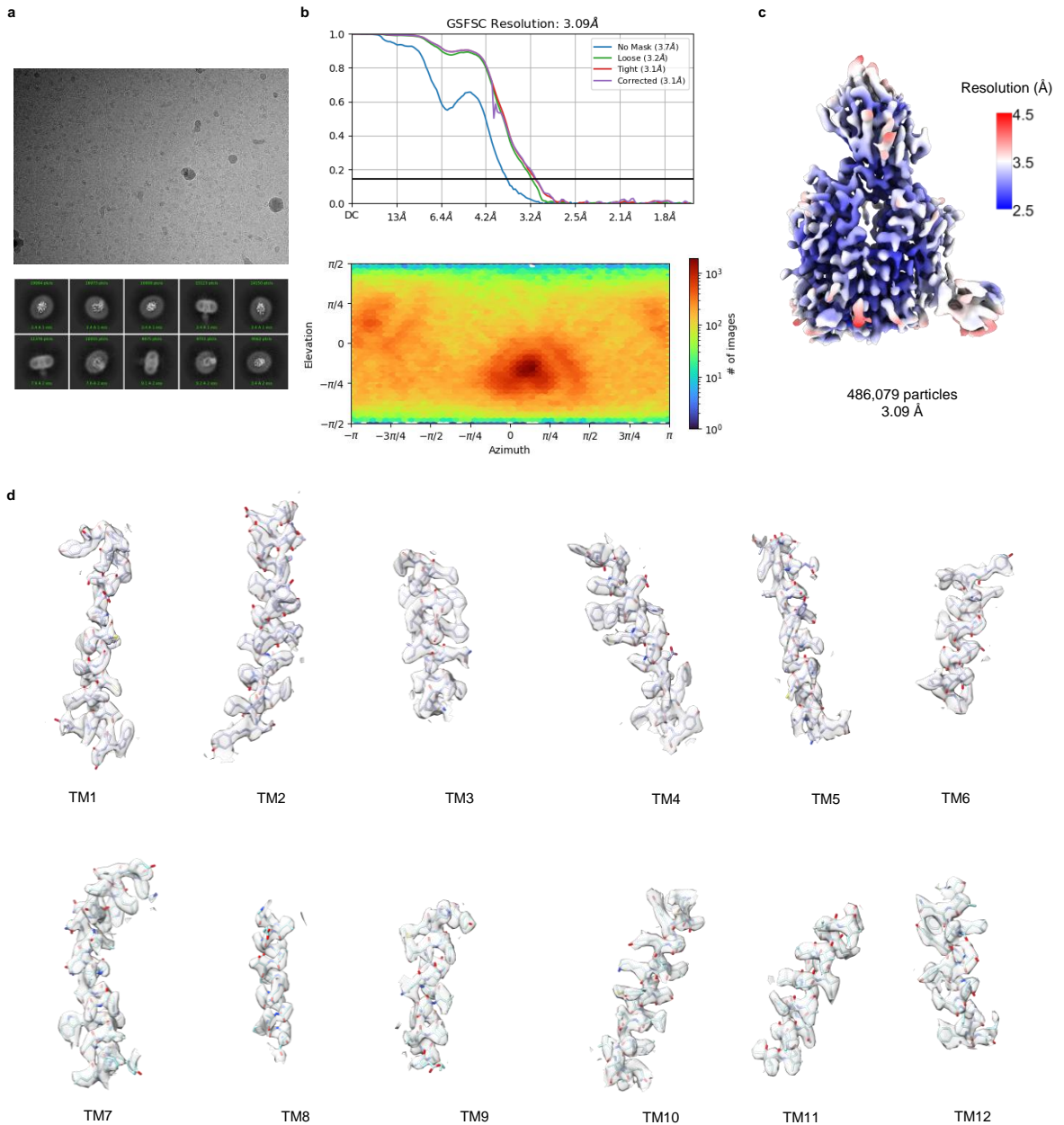




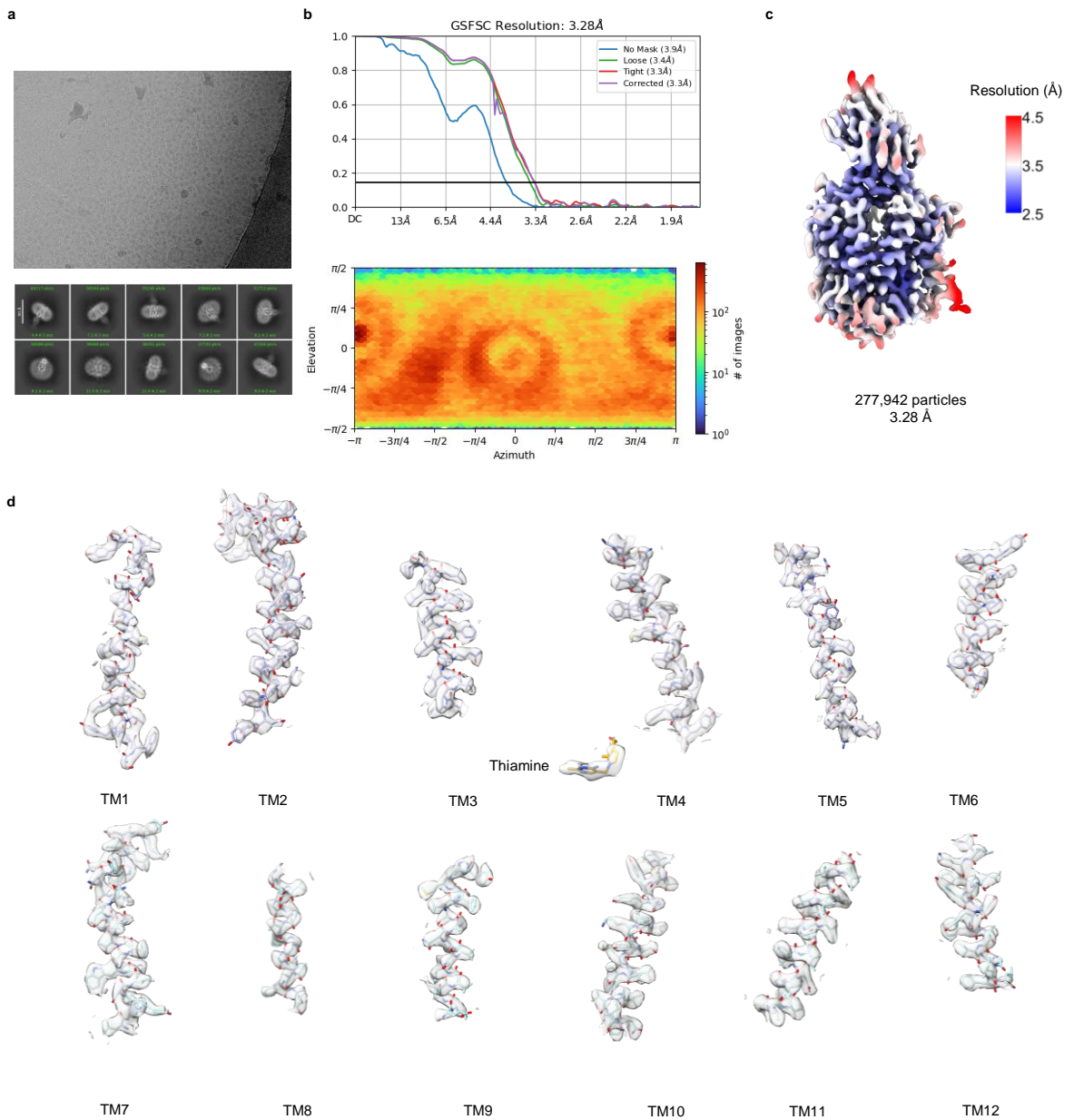
**Supplementary Fig. 5: Thermal shift assays of hSLC19A3 Glu110 mutants and a control solute carrier.** Panels **a-d** show the melting curves of different protein constructs in response to increasing concentrations of thiamine (titration from 1.95  $\mu\text{M}$  to 2 mM). **a**, Thermal shift screen of wildtype hSLC19A3 shows a strong stabilisation of the transporter in the presence of thiamine. **b**, When Glu110 is replaced by a lysine residue, two melting transitions can be observed, none of which responds to thiamine. **c**, The unrelated solute carrier MFSD1 served as a negative control for the thermal shift assays and shows no response to thiamine. **d**, hSLC19A3-E110Q is stabilised by thiamine in a concentration-dependent manner. **e**, Measured thermostability shifts of hSLC19A3-E110Q in dependence of thiamine concentration (coloured circles,  $n = 3$ , independent measurements). The corresponding melting curve was read out based on the ratio of the fluorescence at 350 nm and 330 nm ( $F_{350}/F_{330}$ ). The depicted curve was fitted as described in the methods section to determine the apparent dissociation constants  $K_{d, \text{app}}$ .



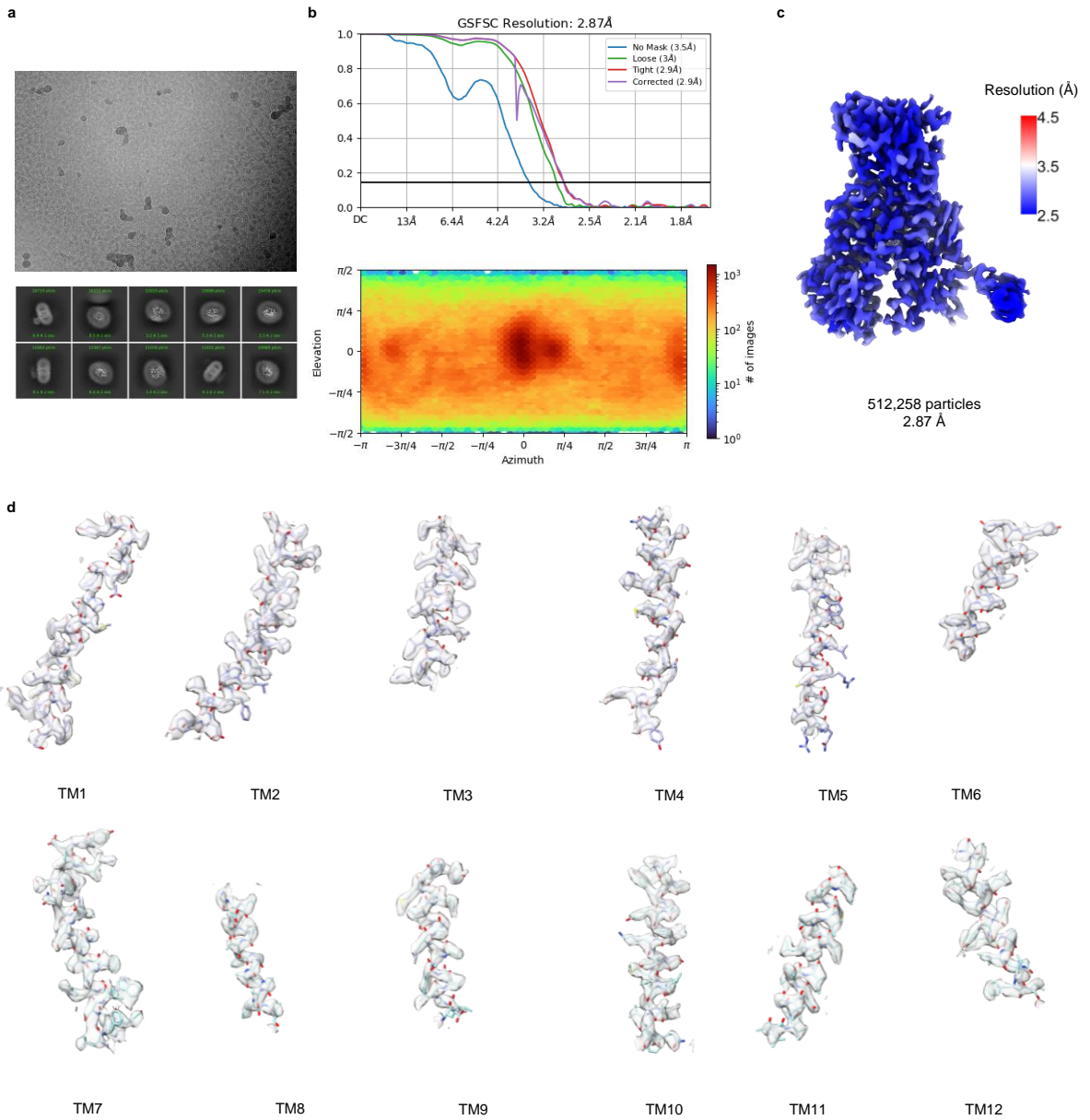
**Supplementary Fig. 6: Representative cryo-EM processing workflow.** **a**, Image processing workflow of the hSLC19A3:Nb3.7:fedratinib data set. The shown workflow was used with minor deviations for the processing of all presented cryo-EM data sets. **b**, Representative micrograph acquired at a defocus of  $-1.6 \mu\text{m}$ . **c**, 2D class averages. **d**, Gold-standard Fourier shell correlation (GSFSC) curves between two half maps. FSC threshold = 0.143. **e**, Angular distribution of particles in the final 3D local refinement. **f**, Cryo-EM density map of hSLC19A3:Nb3.7:fedratinib, surface coloured by local resolution.



**Supplementary Fig. 7: Cryo-EM data processing overview for hSLC19A3-wt:Nb3.4-apo.** **a**, Representative micrograph (top) and 2D class averages (below). **b**, Gold-standard Fourier shell correlation (GSFSC) curves between two half maps (top). FSC threshold = 0.143. Angular distribution of particles in the final 3D local refinement (below). **c**, Local resolution estimation of the final cryo-EM map (map contour level = 0.57). **d**, Close-up on the density for the individual transmembrane helices (map contour level = 0.711).

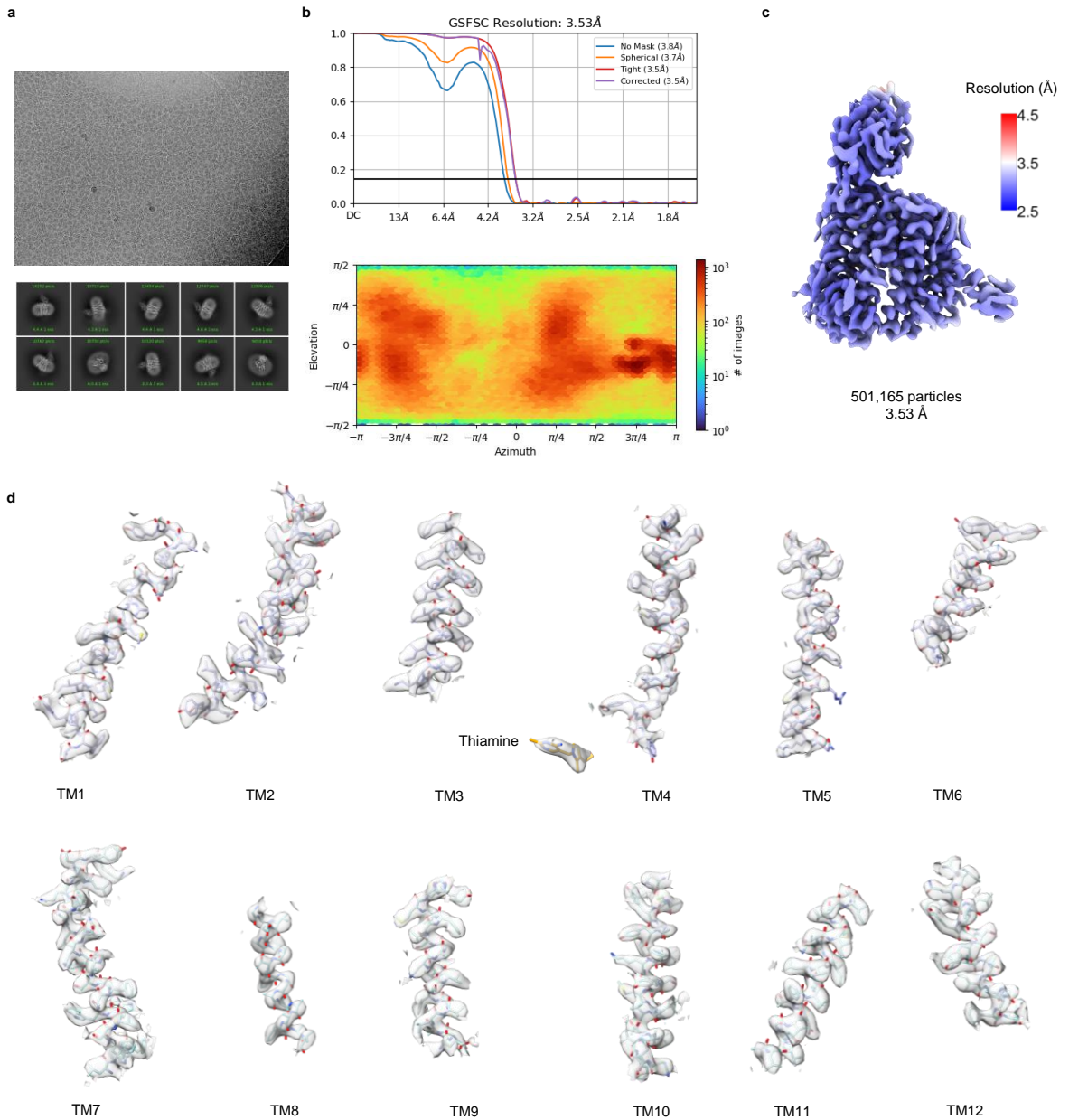


**Supplementary Fig. 8: Cryo-EM data processing overview for hSLC19A3-wt:Nb3.4:thiamine.** **a**, Representative micrograph (top) and 2D class averages (below). **b**, Gold-standard Fourier shell correlation (GSFSC) curves between two half maps (top). FSC threshold = 0.143. Angular distribution of particles in the final 3D local refinement (below). **c**, Local resolution estimation of the final cryo-EM map (map contour level = 0.676). **d**, Close-up on the density for the individual transmembrane helices (map contour level = 1.01).

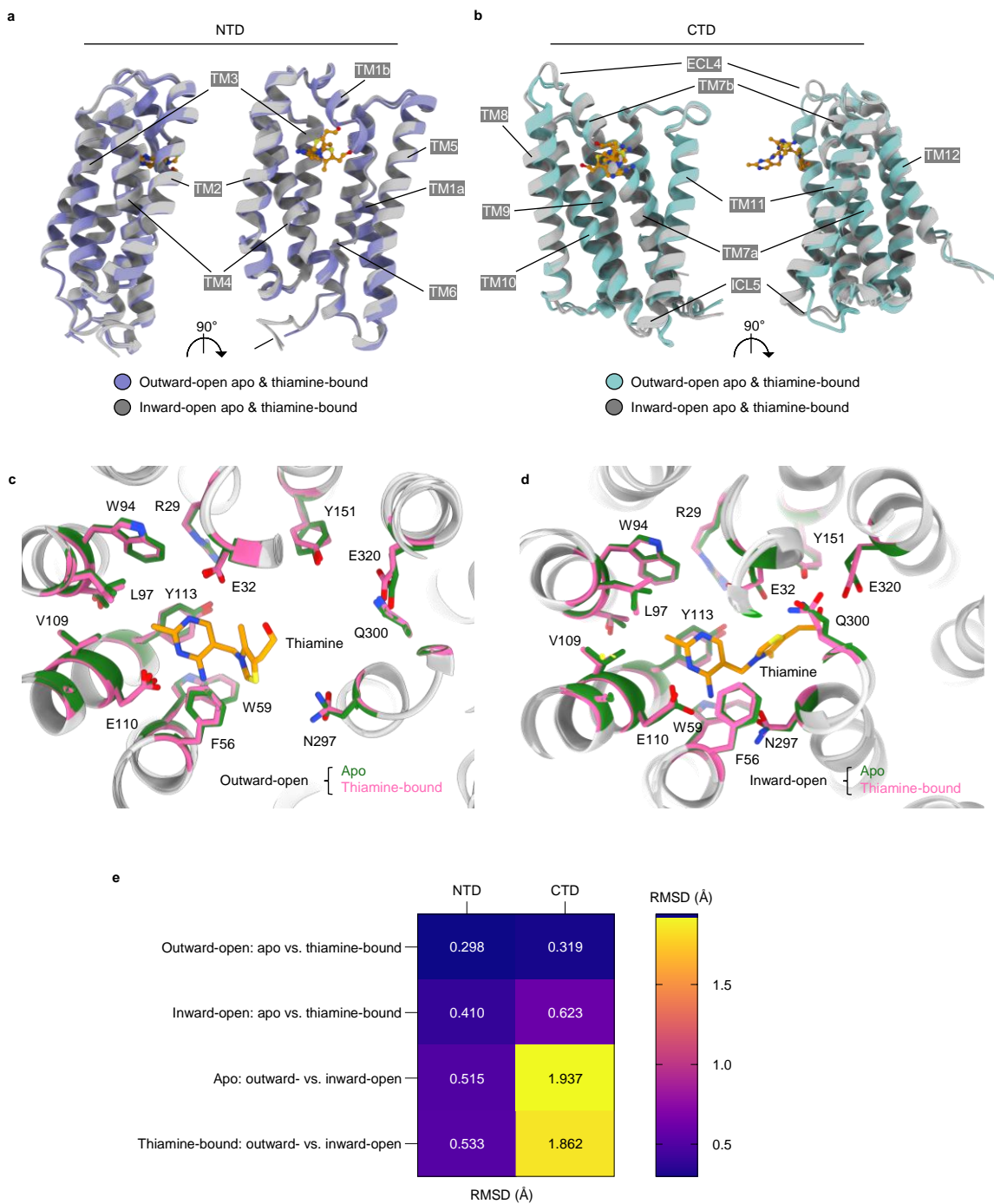


**Supplementary Fig. 9: Cryo-EM data processing overview for hSLC19A3-wt:Nb3.3-apo.** **a**, Representative micrograph (top) and 2D class averages (below). **b**, Gold-standard Fourier shell correlation (GSFSC) curves between two half maps (top). FSC threshold = 0.143. Angular distribution of particles in the final 3D local refinement (below). **c**, Local resolution estimation of the final cryo-EM map (map contour level = 0.951). **d**, Close-up on the density for the individual transmembrane helices (map contour level = 1.07).





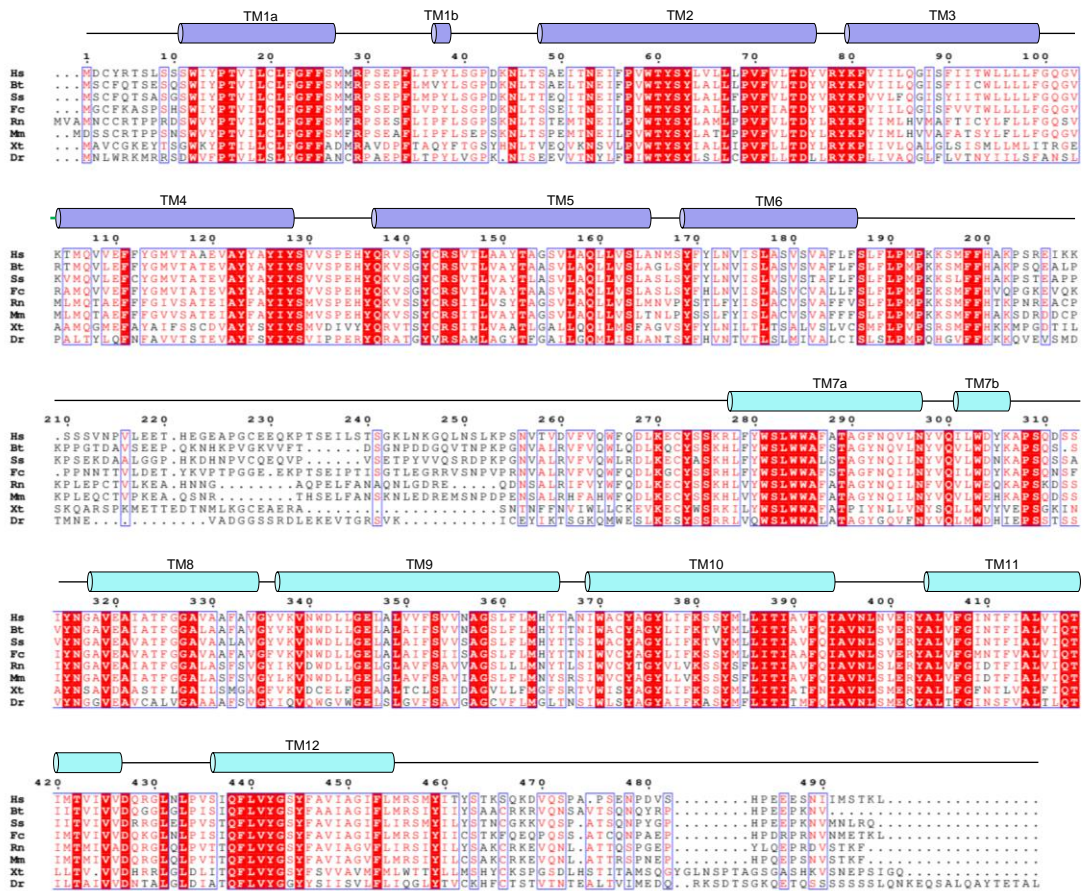
**Supplementary Fig. 10: Cryo-EM data processing overview for hSLC19A3-gf:Nb3.7:thiamine.** **a**, Representative micrograph (top) and 2D class averages (below). **b**, Gold-standard Fourier shell correlation (GSFSC) curves between two half maps (top). FSC threshold = 0.143. Angular distribution of particles in the final 3D local refinement (below). **c**, Local resolution estimation of the final cryo-EM map (map contour level = 0.151). **d**, Close-up on the density for the individual transmembrane helices (map contour level = 0.144).



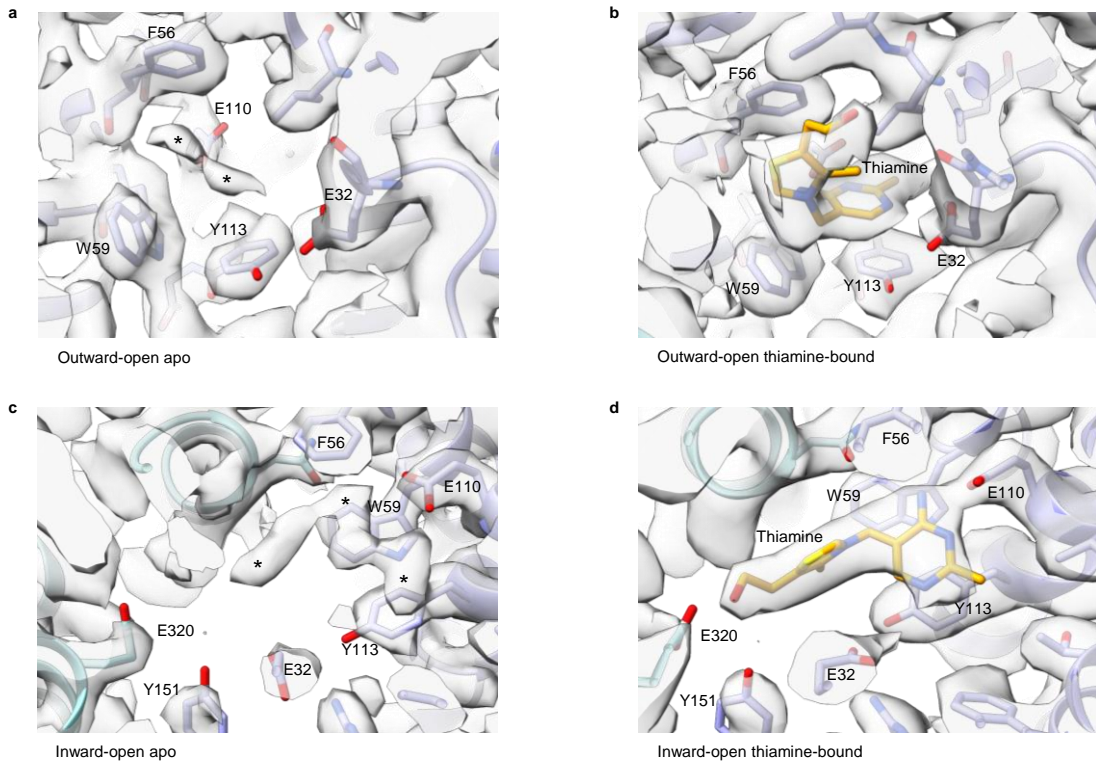
**Supplementary Fig. 11: Structural differences between different conformational states.**

**Supplementary Fig. 11: Structural differences between different conformational states.** **a, b**, Structural superposition of the protein backbone of hSLC19A3 in four different conformations, when aligned to the individual domains, **a**, NTD and **b**, CTD. The coloured structures present the outward-open states (apo and thiamine-bound). The white structures represent the inward-open states (apo and thiamine-bound). The domains move as rigid bodies between the inward- and outward-open state. The only minor intradomain changes that can be observed are rearrangements of ECL4 and ICL5. The movement of ECL4 is likely induced by the binding of Nb3.4. There are no major further rearrangements observable between the apo and thiamine-bound transporter. **c** and **d** show the substrate binding site of hSLC19A3 in the outward- and inward-open state, respectively. A comparison of the apo (green) and the thiamine-bound (pink) structures reveals no major difference between them. This indicates that thiamine is recognised in a key-in-lock manner by the transporter. **e**, Assessment of the overall structural changes of the transporter between different states as root mean square deviation (RMSD) of C<sub>α</sub> atoms of the individually aligned domains (NTD and CTD). Thiamine-binding results in a very low RMSD of 0.3-0.6 Å in both domains. Larger rearrangements can only be observed in the CTD, when transitioning in the rocker-switch like motion between the outward- and inward-open state. The observed higher RMSD values (1.862-1.937 Å) are mainly caused by the rearrangements in ECL4 and ICL5.

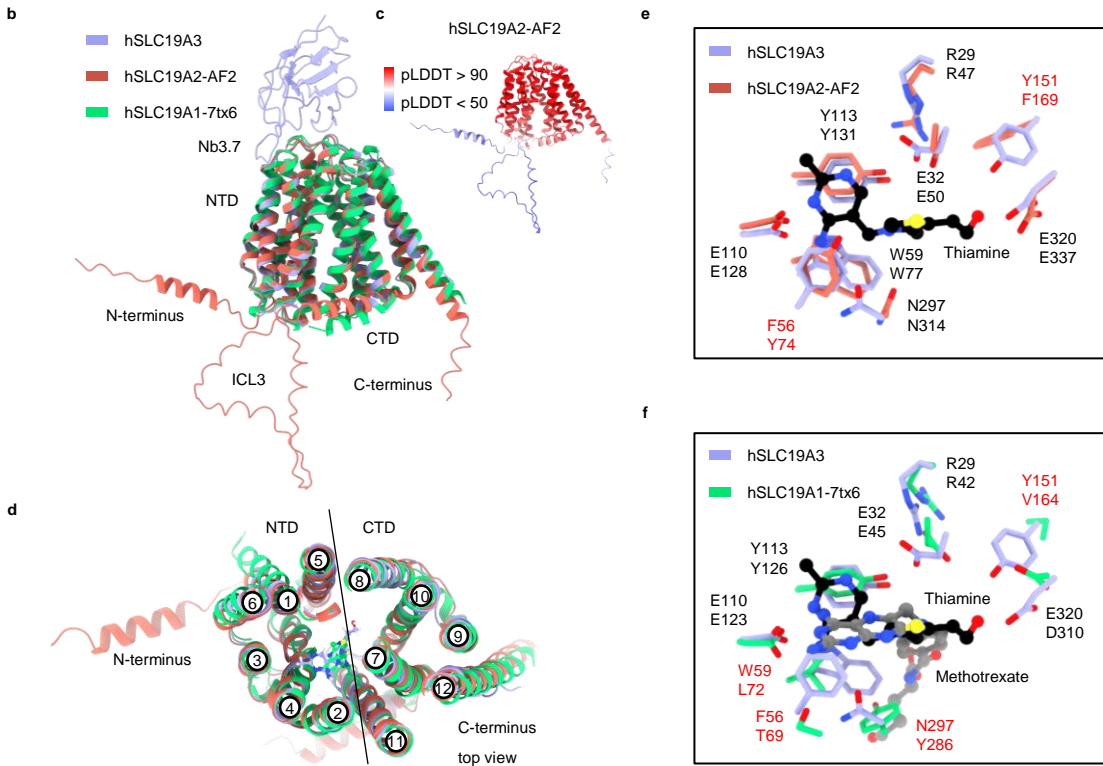
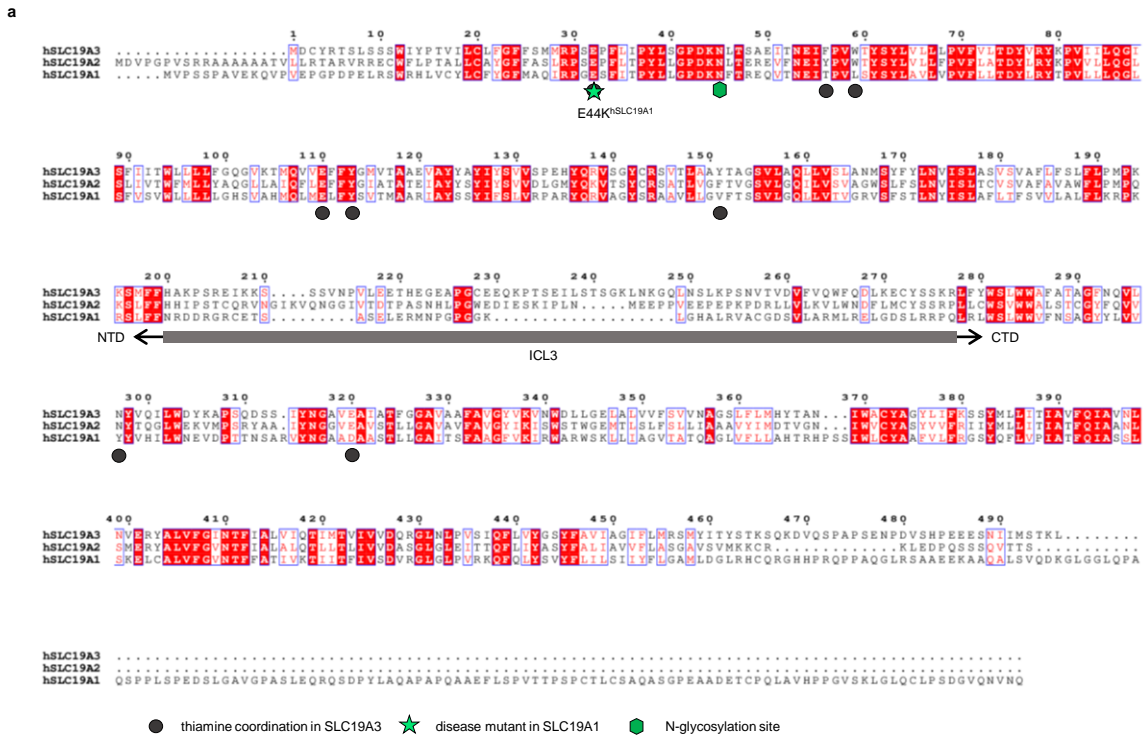




**Supplementary Fig. 12: Sequence conservation of SLC19A3 in vertebrates.** Multiple sequence alignment of SLC19A3 protein sequences of human (*Homo sapiens*, Hs), cattle (*Bos taurus*, Bt), pig (*Sus scrofa*, Ss), cat (*Felis catus*, Fc), rat (*Rattus norvegicus*, Rn), mouse (*Mus musculus*, Mm), Western clawed frog (*Xenopus tropicalis*, Xt) and zebra fish (*Danio rerio*, Dr). Residues are coloured by conservation. The topology of the transporter is indicated by the corresponding transmembrane helices (TMs) above the sequence alignment.



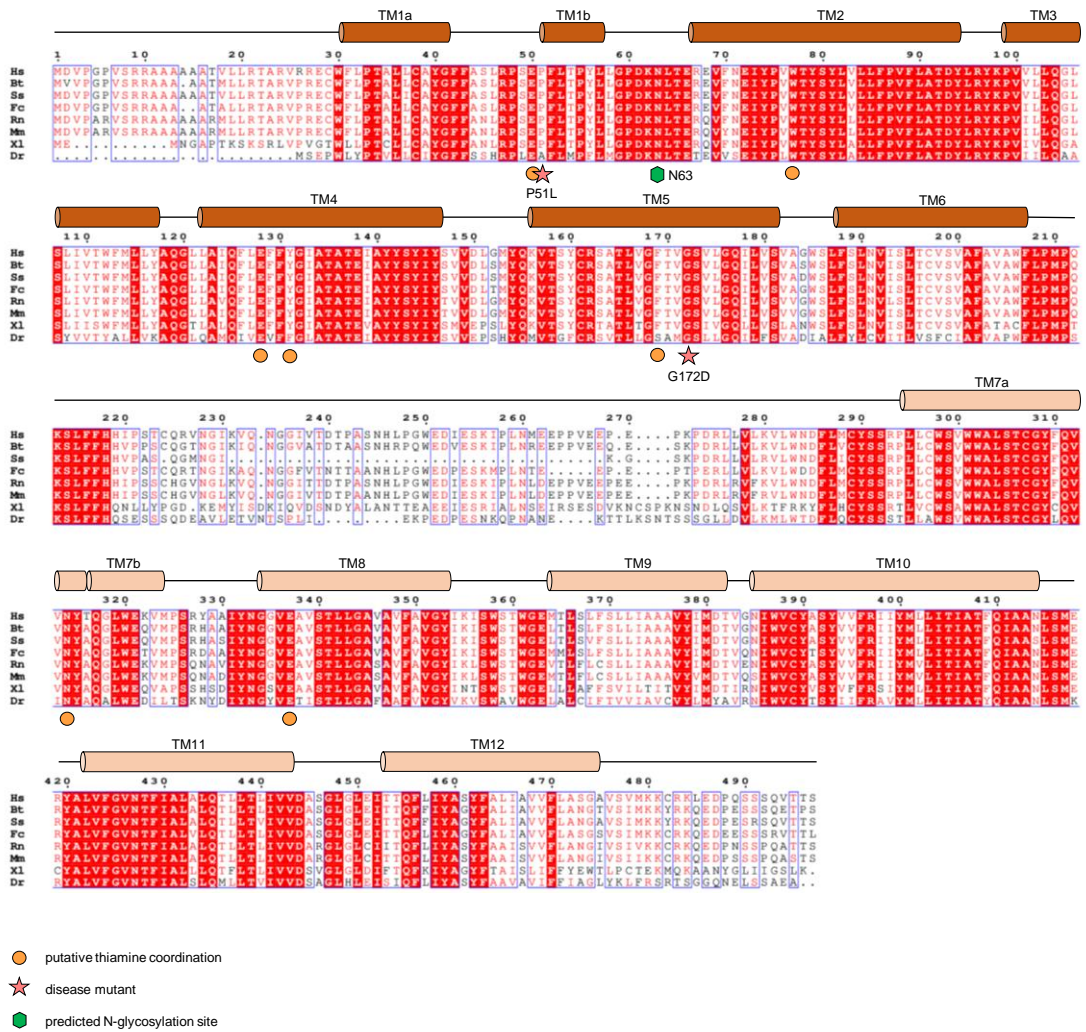
**Supplementary Fig. 13: Cryo-EM density in the substrate binding site of hSLC19A3 in different states of the transport cycle.** Panels **a-d** show the experimentally determined cryo-EM density of hSLC19A3 in the indicated outward- and inward-open states, in its apo and thiamine-bound form, respectively. The density is depicted as a grey surface, at the contour levels listed below. The respective structure models of hSLC19A3 are shown in purple (NTD) and cyan (CTD). Thiamine is depicted in orange. The stars (\*) in the apo structures indicate unassigned density, which is most likely explained by random small molecules that are pulled down with the transporter during purification. ChimeraX contour levels: **a**, 0.365; **b**, 0.372; **c**, 0.517; **d**, 0.104.



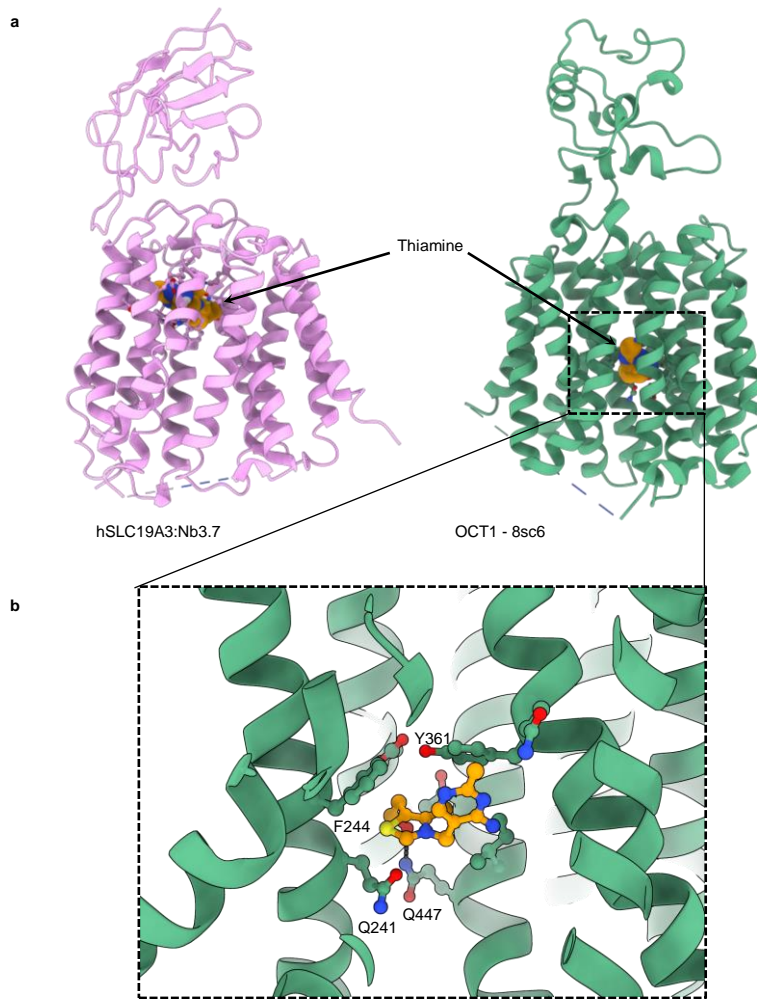
Supplementary Fig. 14: Structure and sequence comparison between the members of the SLC19A family.

**Supplementary Fig. 14: Structure and sequence comparison between the members of the SLC19A family.** **a**, Multiple sequence alignment of human SLC19A3, SLC19A2 and SLC19A1. Residue numbering is based on hSLC19A3. Residue colouring by conservation. Crucial functional or disease related residues are labelled. **b**, Superposition of the cryo-EM structures of hSLC19A3:Nb3.7:thiamine, methotrexate-bound hSLC19A1 (PDB ID 7tx6, Wright et al., 2022) and the structure model of hSLC19A2, predicted by AlphaFold2 (AF2, UniProt identifier: O60779, accessed on 23 January 2024). All structures are in the inward-open state. The fold of the three members of the SLC19A family is highly conserved. On a backbone level, only small differences can be observed. **c**, hSLC19A2 AF2 predicted structure, coloured by the predicted local distance difference test (pLDDT) score. **d**, Extracellular view of **b**. **e**, Close-up on the substrate binding site, comparing hSLC19A3 (cryo-EM structure) and hSLC19A2 (AF2 prediction). The residues of hSLC19A3 are labelled above the corresponding residues of hSLC19A2. Black labels indicate full conservation of the respective residue, whereas red labels indicate amino acid replacement in the corresponding position. The substrate binding site is structurally highly conserved. Only two residue substitutions can be observed from hSLC19A3 to hSLC19A2, namely Phe56 to Tyr and Tyr151 to Phe. **f**, Close-up on the substrate binding site, comparing thiamine-bound hSLC19A3 and methotrexate-bound hSLC19A1. The residues of hSLC19A3 are labelled above the corresponding residues of hSLC19A1. Black labels indicate full conservation of the respective residue, whereas red labels indicate amino acid replacement in the corresponding position. While Glu32, Glu110 and Tyr113 are structurally conserved in hSLC19A1, the aromatic clamp is reshuffled to allow a membrane-perpendicular binding of the methotrexate aromatic ring system in the transporter. This stands in contrast to the membrane-parallel coordination of the aminopyrimidine ring of thiamine by hSLC19A3.

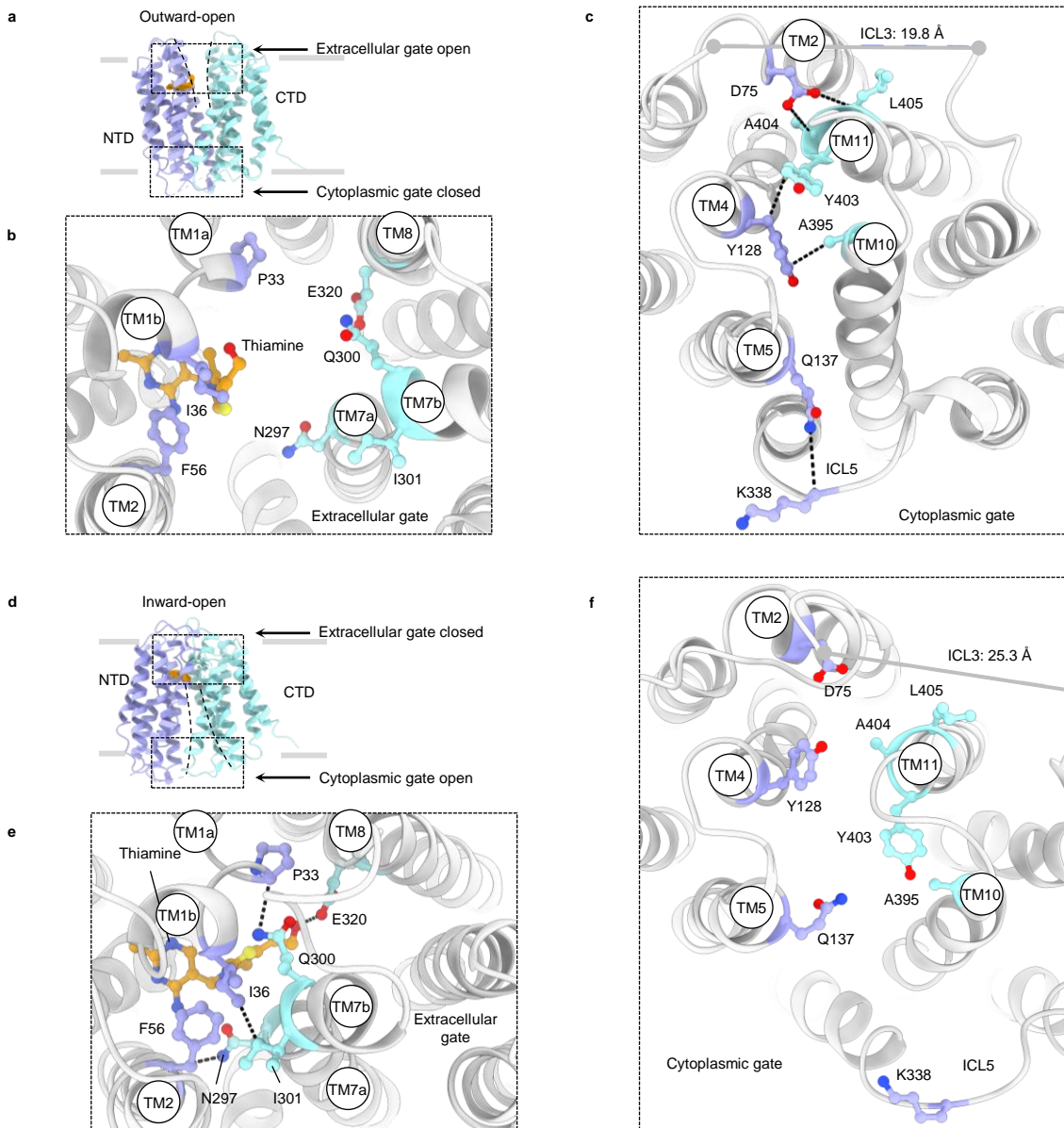




**Supplementary Fig. 15: Sequence conservation of SLC19A2 in vertebrates.** Multiple sequence alignment of SLC19A2 protein sequences of human (*Homo sapiens*, Hs), cattle (*Bos taurus*, Bt), pig (*Sus scrofa*, Ss), cat (*Felis catus*, Fc), rat (*Rattus norvegicus*, Rn), mouse (*Mus musculus*, Mm), Western clawed frog (*Xenopus tropicalis*, Xt) and zebra fish (*Danio rerio*, Dr). Residues are coloured by conservation. The topology of the transporter is indicated by the corresponding transmembrane helices (TMs) above the sequence alignment. The stars indicate mutants of hSLC19A2, which cause thiamine-responsive megaloblastic anemia (TRMA), namely Pro51Leu (Lagard et al., 2004) and Gly172Asp (Labay et al., 1999). Residues likely important for substrate coordination (based on homology to SLC19A3) are labelled with an orange circle.



**Supplementary Fig. 16: Structure comparison between hSLC19A3 and OCT1 (hSLC22A1).** **a**, Cryo-EM structures of hSLC19A3:Nb3.7:thiamine and thiamine-bound OCT1 (SLC22A1, PDB ID 8sc6, Zeng et al., 2023), shown in the membrane plane. Thiamine is highlighted as orange sphere model. The substrate binding site of SLC19A3 is close to the extracellular space, whereas the substrate binding site of OCT1 is close to the centre of the membrane. **b**, Close-up on the substrate binding site of OCT1 coordinating thiamine. Thiamine interacts via  $\pi$ - $\pi$  interactions with Phe244 and Tyr361 and forms a hydrogen bond to Gln447 with its hydroxyethyl tail.

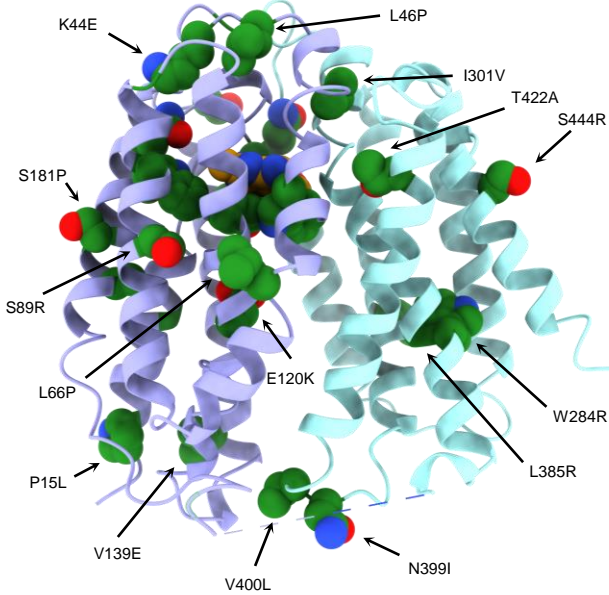


**Supplementary Fig. 17: Molecular gates.** **a-c**, Representation of the molecular gates in the outward-open state of hSLC19A3. While the extracellular gate is open, the cytoplasmic gate is sealed by the apolar interactions between the side chains of Tyr128 on the NTD and Ala395 and Tyr403 on the CTD. The gate is further stabilised by polar interactions between Asp75 and the backbone of Leu405, as well as Gln137 and the backbone of Lys338. **d-f**, Molecular gates of hSLC19A3 in the inward-open state. The extracellular gate is closed. Central polar contacts are formed by Asn297 with the backbone of Phe56 and Gln300 and the backbone of Pro33. The gate is further sealed by the apolar interaction of Ile36 with Ile301. The cytoplasmic gate is open. The structured endpoints of ICL3 move away from each other (from ~20 Å to about 25 Å, measured between C<sub>α</sub> of residues 195 and 271). Selected residues are shown as sticks and black dashes indicate hydrogen bonds and hydrophobic interactions (cut-off at 3.2 Å).

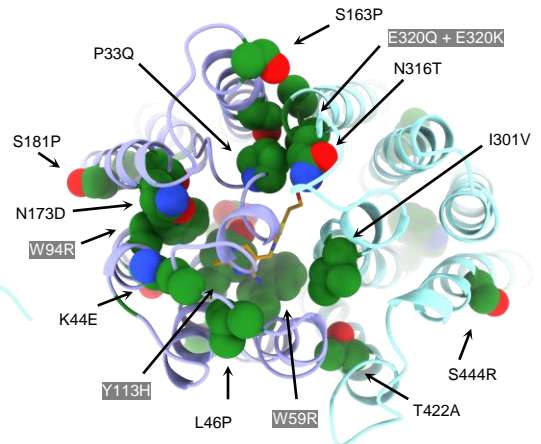
a

Disease	Age of onset	Mutation	Structural module	Specific reference
Classical childhood BTBGD	4.73 ± 0.35 years	P15L		
		G41R		
		W59R	Substrate binding site	Aburezq et al., 2023
		L66P		
		W94R	Substrate binding site	Schänzer et al., 2014
		G100R		
		Y113H	Substrate binding site	
		V139E		
		S155L		
		S163P		
		N173D		
		S176Y		Kevelam et al., 2013
		S181P		Kevelam et al., 2013
		W284R		
		G317E		
Infantile phenotype	3.00 ± 0.48 months	E320K	Substrate binding site	Wesol-Kucharska et al., 2021
		A321V		
		L385R		Kevelam et al., 2013
		N399I		
		V400L		
		T422A		
		P33Q	EC gate	
Mixed type		L46P		
		E120K		
		A183V		
		I301V	EC gate	
		N316T		
WLE	10-20 years	L385R		
		S444R		
Mixed type		G23V		
		S89R		
		K44E		Kono et al., 2009
WLE	10-20 years	E320Q	Substrate binding site	Kono et al., 2009

b



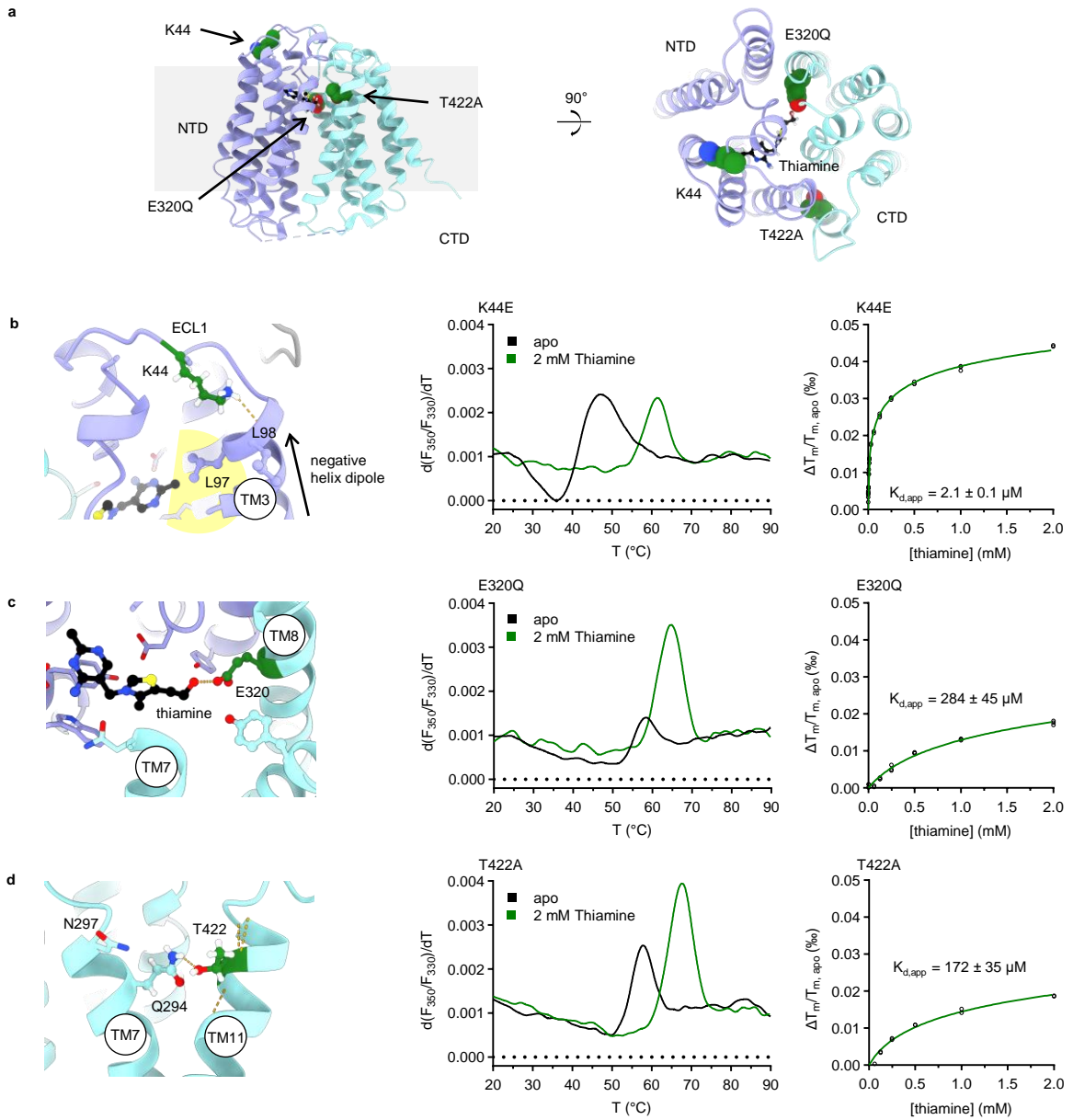
c



Supplementary Fig. 18: Disease-causing mutations of hSLC19A3.

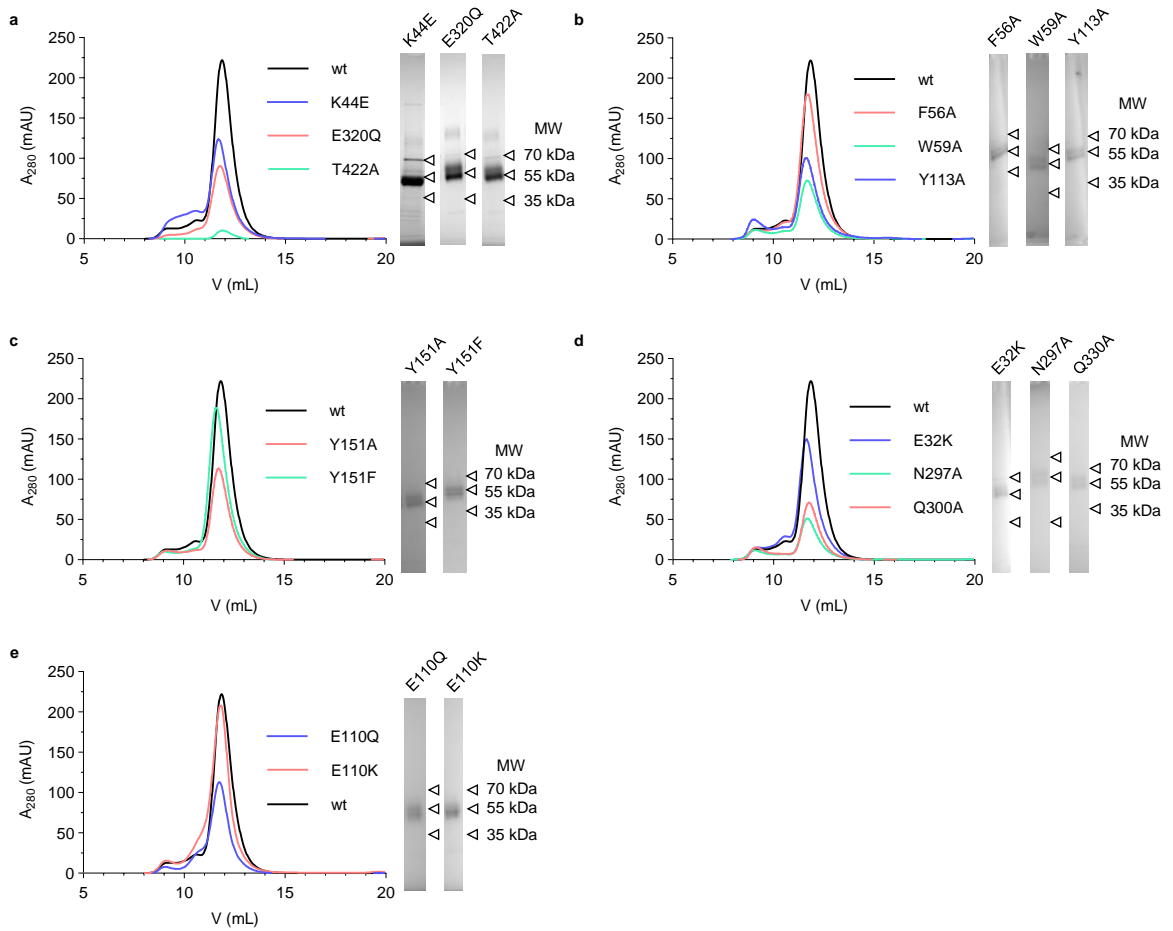


**Supplementary Fig. 18: Disease-causing mutations of hSLC19A3.** **a**, List of disease-causing point mutants of hSLC19A3. Most of these mutants cause biotin- and thiamine-responsive basal ganglia disease (BTBGD). This disease can be subdivided in classical childhood BTBGD, BTBGD with infantile phenotype and mixed type BTBGD, based on the age of onset (Wang et al., 2021). The heterozygous combination of the two latter mutations (K44E and E320Q) causes in contrast Wernicke's like encephalopathy (WLE), with an onset of symptoms in the second decade of life (Kono et al., 2009). Unless otherwise specified, see Wang et al., 2021 for a review of the mutations. Based on our structural data, we can assign a structural function to several of the mutated residues, e.g. when they are part of the substrate binding site or the molecular gates of the transporter. **b**, Cryo-EM structure of inward-open hSLC19A3 bound to thiamine (orange), shown in the membrane plane. Disease causing mutations as listed in the table above are highlighted in dark green sphere in the structure. **c**, The same complex, shown from an extracellular perspective. Substrate binding site residues are highlighted in grey.

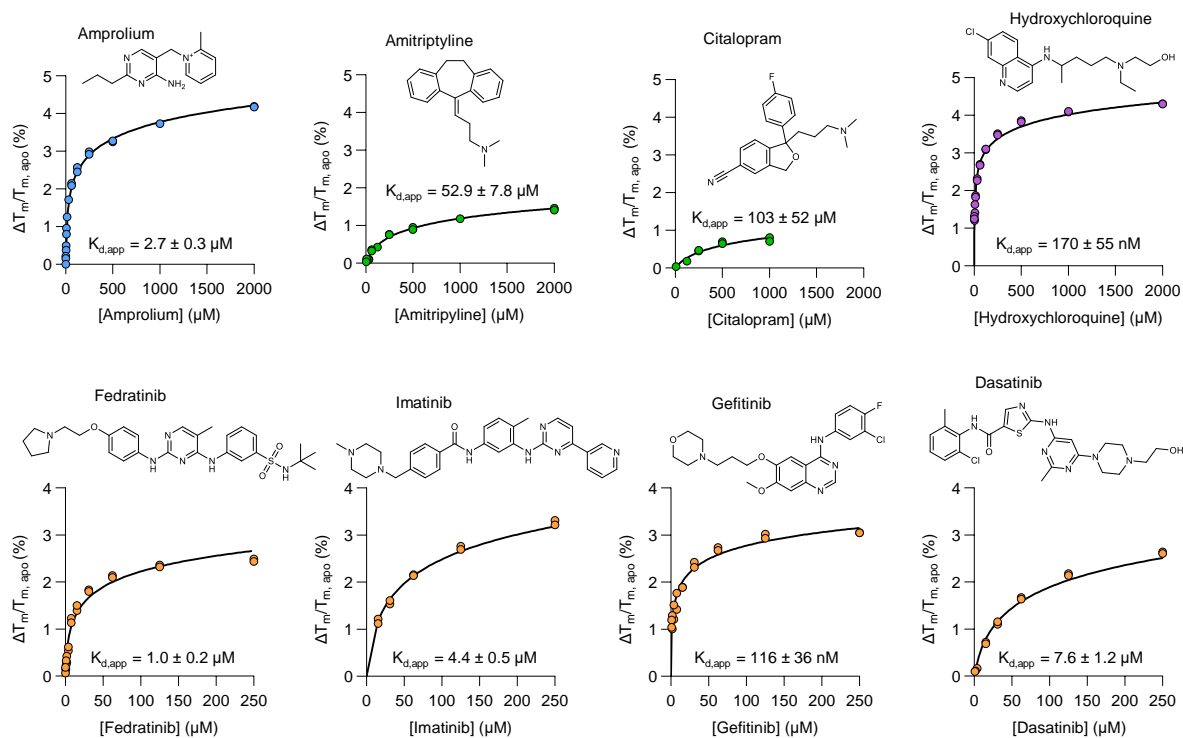


**Supplementary Fig. 19: Thermostability and thiamine binding affinity of selected disease-causing mutations of hSLC19A3.**

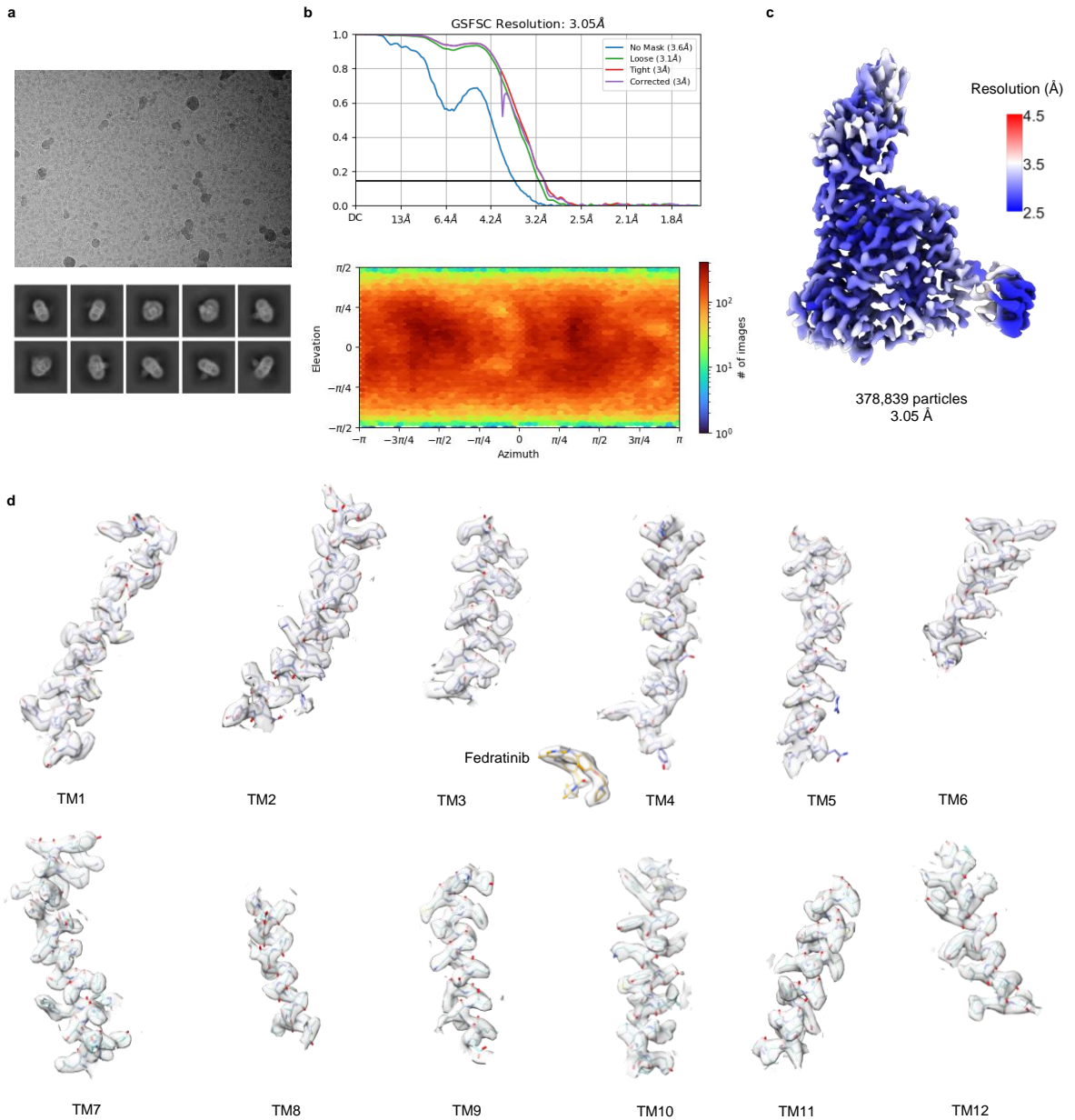
**Supplementary Fig. 19: Thermostability and thiamine binding affinity of selected disease-causing mutations of hSLC19A3.** **a**, Localisation of the studied disease-causing mutations. Individuals that carry aSLC19A3-Lys44Glu and -Glu320Gln allele at the same time develop Wernicke's like encephalopathy (Kono et al., 2009). Thr422Ala is the most reported single point mutation of SLC19A3 causing biotin- and thiamine-responsive basal ganglia disease (BTBGD) in homozygous individuals. **b-d**, The panels below show structural close-ups to the affected residues in the cryo-EM structure of hSLC19A3:Nb3.7:thiamine (left), representative melting curves (middle; first derivative of the ratio of the fluorescence at 350 nm and 330 nm (F350/F330)), and thiamine concentration dependent thermal shifts used for the determination of thiamine binding affinities (right). **b**, Lys44 forms polar contacts with the C-terminal end of TM3. Its positively charged side chain engages with the backbone carbonyl of Leu98 and presumably with the negative helix dipole of TM3. The thermostability of this mutant is strongly reduced ( $\Delta T_m^{K44E-wt} = -15$  °C), when compared to wildtype. However, the mutated transporter binds thiamine with a six-fold higher affinity than wildtype hSLC19A3. While this mutant appears to retain its substrate binding capacity, it has been shown that it cannot be trafficked properly to the plasma membrane (Kono et al., 2009). Our data, in the form of a reduced thermostability and lacking glycosylation of hSLC19A3-Lys44Glu (Supplementary Fig. 19a) suggest that the mutant might be quite unstable to pass protein quality control in the endoplasmic reticulum (ER). **c**, hSLC19A3-Glu320Gln, in contrast, localises to the plasma membrane (Kono et al., 2009); the thiamine uptake via this mutant is, however, strongly impaired. Our biophysical data display, that the stability of the mutant is only moderately affected ( $\Delta T_m^{E320Q-wt} = -4$  °C). The affinity for thiamine was, however, markedly reduced from a  $K_{d,app}$  of 12.7  $\mu$ M to 284  $\mu$ M. This underpins the importance of the hydrogen bond between Glu320 and the hydroxyethyl tail of thiamine for high-affinity binding. **d**, Thr422 forms a hydrogen bond with Gln294, which might be important for proper helix packing. The thermostability of the Thr422Ala mutant of hSLC19A3 is decreased and binds thiamine with > ten-fold lower affinity. The lower binding affinity could render the transporter unreceptive for physiological thiamine levels. It might, however, be able to transport thiamine at higher concentrations, which is supported by the success of high-dose thiamine treatment of patients carrying this mutation. The protons shown in this figure have not been directly observed, but are only inferred from the most likely protonation state of the corresponding chemical moieties at neutral pH.



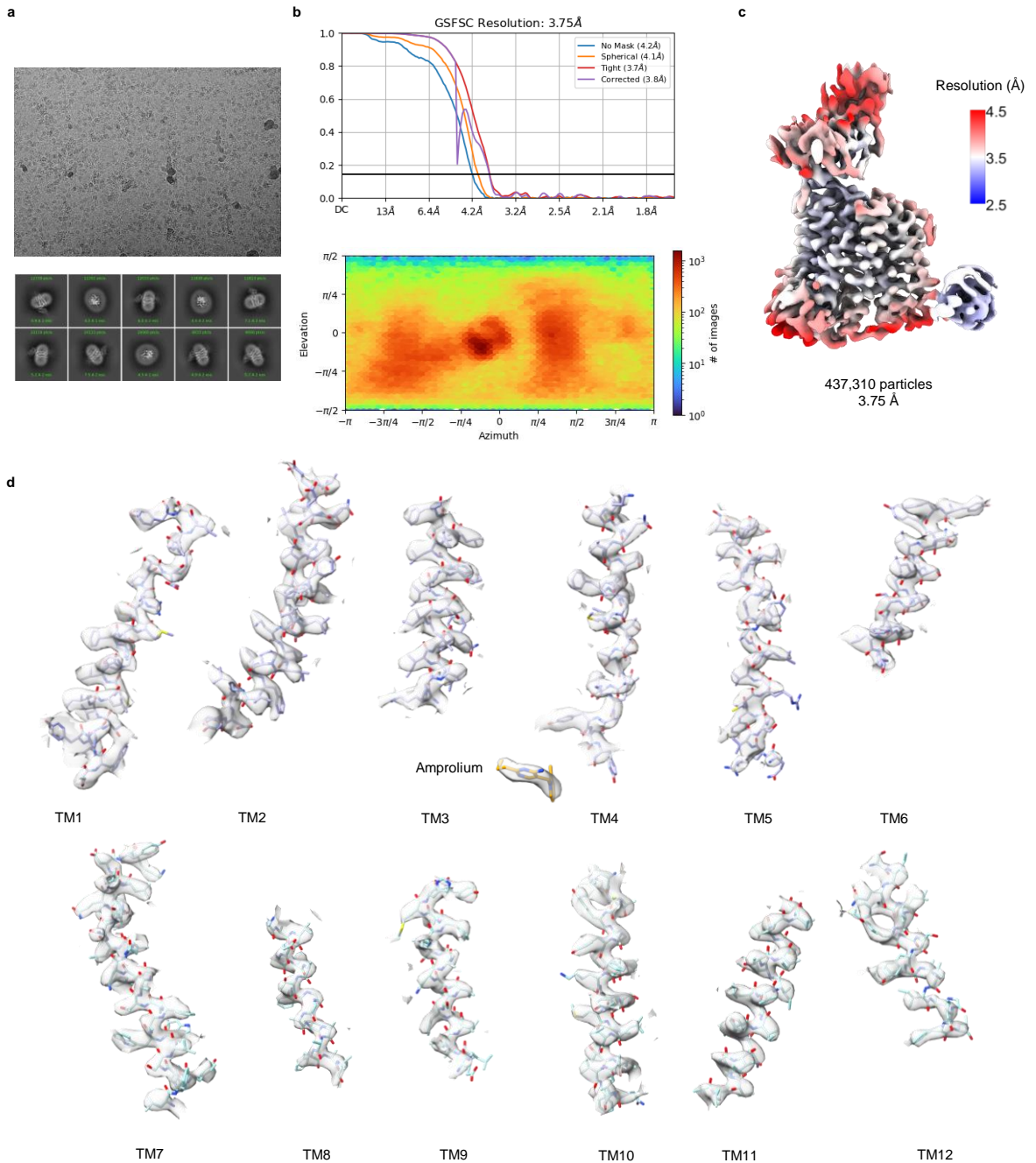
**Supplementary Fig. 20: Purification of single point mutants of hSLC19A3.** a-e, Size-exclusion chromatography (SEC) traces are shown in the respective chromatograms (absorbance at  $\lambda = 280$  nm), while the insets show SDS-PAGE analysis of the pooled SEC-peak fraction. For **a**, concentrated protein was loaded on the SDS-PAGE gels, whereas for the rest of the shown gels, the unconcentrated peak fraction was loaded directly. Note in **a**, left inset, that the disease-causing mutant Lys44Glu lacks the additional higher molecular weight band slightly above 55 kDa, which represents glycosylated protein. It has been shown previously, that this mutant is retained in the endoplasmic reticulum (ER, Kono et al., 2009). Our data exhibit that hSLC19A3-Lys44Glu is likely not properly glycosylated in the ER which could prevent trafficking to the plasma membrane. For further details, see Supplementary Fig. 18. White triangles indicate the molecular weight markers.



**Supplementary Fig. 21: Drug interactions of hSLC19A3.** Determination of the apparent binding affinity ( $K_{d,app}$ ) of the TUIs, identified in the thermal shift screen, using the readouts as described above. Shown are the chemical structures and the binding isotherms that were fitted according to a modified script of Hall's method (Hall, 2019, Kotov et al., 2023) to the experimentally determined compound concentration-dependent thermal shifts ( $n = 3$ , independent measurements, shown as coloured circles).

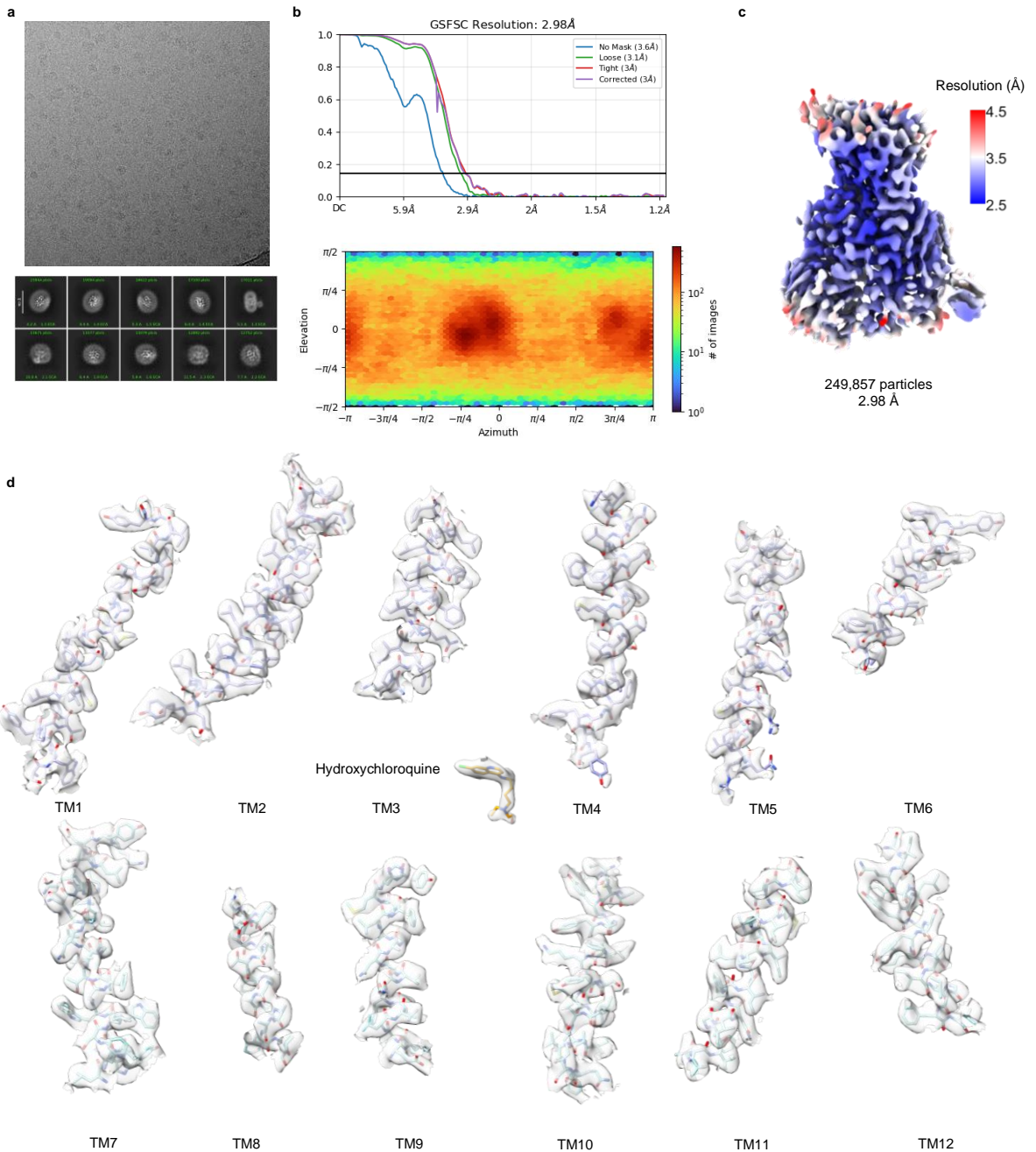


**Supplementary Fig. 22: Cryo-EM data processing overview for hSLC19A3-gf:Nb3.7:fedratinib.** **a**, Representative micrograph (top) and 2D class averages (below). **b**, Gold-standard Fourier shell correlation (GSFSC) curves between two half maps (top). FSC threshold = 0.143. Angular distribution of particles in the final 3D local refinement (below). **c**, Local resolution estimation of the final cryo-EM map (map contour level = 0.858). **d**, Close-up on the density for the individual transmembrane helices (map contour level = 0.563).



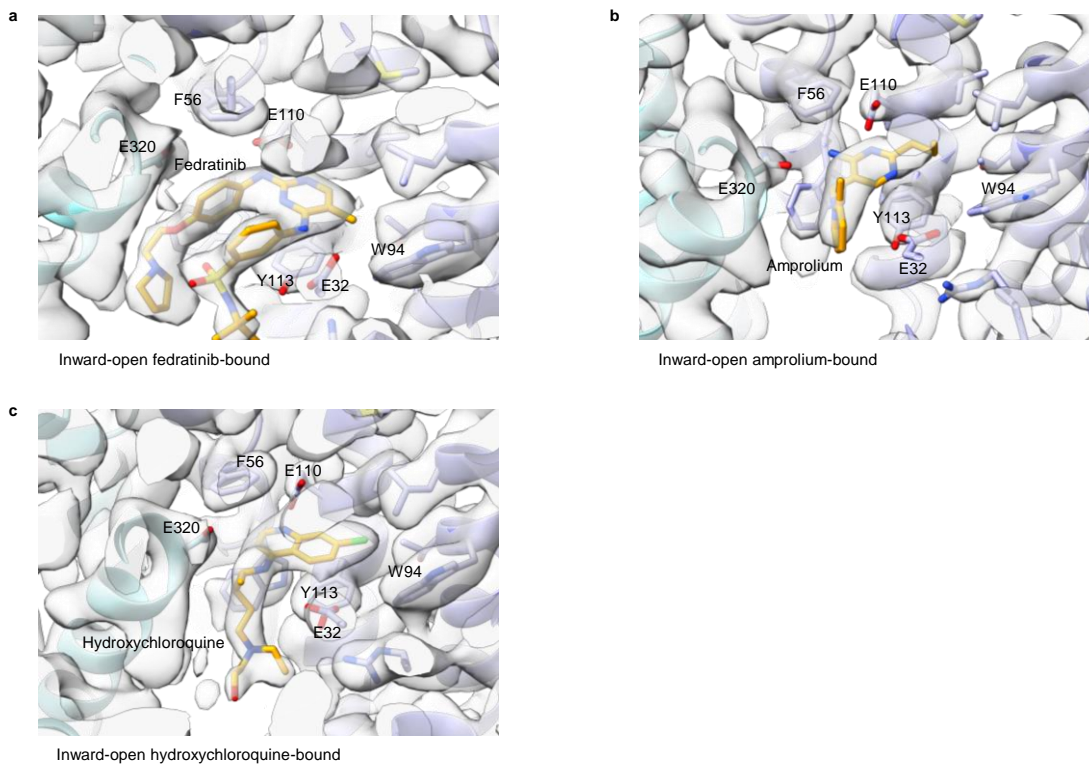
**Supplementary Fig. 23: Cryo-EM data processing overview for hSLC19A3-gf:Nb3.7:amprolium.** **a**, Representative micrograph (top) and 2D class averages (below). **b**, Gold-standard Fourier shell correlation (GSFSC) curves between two half maps (top). FSC threshold = 0.143. Angular distribution of particles in the final 3D local refinement (below). **c**, Local resolution estimation of the final cryo-EM map (map contour level = 0.102). **d**, Close-up on the density for the individual transmembrane helices (map contour level = 0.134).



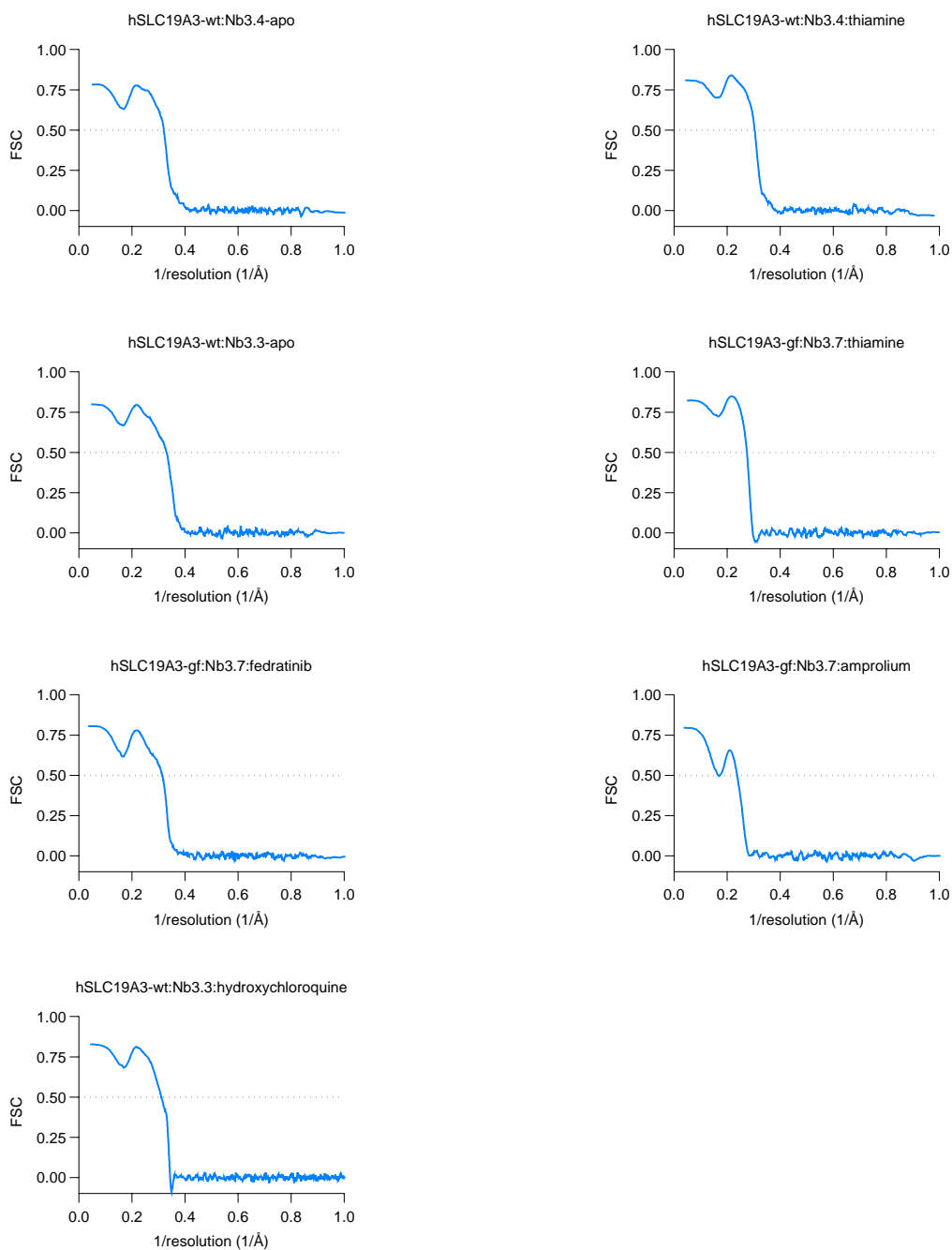


**Supplementary Fig. 24: Cryo-EM data processing overview for hSLC19A3-gf-Nb3.7:hydroxychloroquine.** **a**, Representative micrograph (top) and 2D class averages (below). **b**, Gold-standard Fourier shell correlation (GSFSC) curves between two half maps (top). FSC threshold = 0.143. Angular distribution of particles in the final 3D local refinement (below). **c**, Local resolution estimation of the final cryo-EM map (map contour level = 0.741). **d**, Close-up on the density for the individual transmembrane helices (map contour level = 0.443).





**Supplementary Fig. 25: Cryo-EM density in the substrate binding site of hSLC19A3 in complex with different inhibitors.** Panels **a-c** show the experimentally determined cryo-EM density of hSLC19A3 bound to the respectively indicated inhibitors. The density is depicted as a grey surface, at the contour levels listed below. The respective structure models of hSLC19A3 are shown in purple (NTD) and cyan (CTD). The inhibitors are depicted in orange. ChimeraX contour levels: **a**, 0.652; **b**, 0.16; **c**, 0.445.



**Supplementary Fig. 26: Map-versus-model FSC curves of the cryo-EM structures of hSLC19A3.** The blue curves show the FSC of the different cryo-EM density maps against the respective structure models. The FSCs were calculated using the Mtriage tool in PHENIX. The dashed line indicates the correlation cut-off of 0.5 used as measure for the model resolutions (Supplementary Table 1).

**Supplementary Table 1:****Cryo-EM data collection, refinement, and validation statistics**

	hSLC19A3- wt:Nb3.4-apo  (EMDB-19716) (PDB 8S4U)	hSLC19A3- wt:Nb3.4:Thia- mine  (EMDB-19750) (PDB 8S5U)	hSLC19A3- wt:Nb3.3-apo  (EMDB-51088) (PDB 9G5K)	hSLC19A3- gf:Nb3.7:Thia- mine  (EMDB-19754) (PDB 8S61)
<b>Data collection and processing</b>				
Magnification	105,000x	105,000x	105,000x	105,000x
Voltage (kV)	300	300	300	300
Electron exposure (e-/Å <sup>2</sup> )	45	62	45	45
Defocus range (µm)	0.8-1.6	0.8-1.6	0.8-1.6	0.8-1.6
Pixel size (Å)	0.83	0.85	0.83	0.83
Symmetry imposed	C1	C1	C1	C1
Initial particle images (no.)	1,084,118	1,602,998	1,619,839	1,264,486
Final particle images (no.)	486,079	277,942	512,258	501,165
Map resolution (Å)	3.09	3.28	2.87	3.53
FSC threshold	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.6-3.5	2.8-4.5	2.4-3.0	2.9-3.4
<b>Refinement</b>				
Initial model used (PDB code)	AF_ AFQ9BZV2F1	AF_ AFQ9BZV2F1	AF_ AFQ9BZV2F1	AF_ AFQ9BZV2F1
Model resolution (Å)	3.09	3.28	2.87	3.53
FSC threshold	0.143	0.143	0.143	0.143
Model resolution range (Å)	2.6-3.5	2.8-4.5	2.4-3.0	2.9-3.4
Map sharpening <i>B</i> factor (Å <sup>2</sup> )	-65.6	-90.1	-80.0	-111.1
Model composition				
Non-hydrogen atoms	3,981	3,999	4,042	3,990
Protein residues	502	502	507	502
Ligands	1	1	1	1
<i>B</i> factors (Å <sup>2</sup> )				
Protein	29.14	27.76	27.31	10.58
Ligand	N/A	24.38	N/A	10.30
R.m.s. deviations				
Bond lengths (Å)	0.006	0.006	0.004	0.005
Bond angles (°)	0.981	0.990	0.953	0.929
Validation				
MolProbity score	1.67	1.51	1.36	1.72
Clashscore	6.91	5.01	5.21	5.41
Poor rotamers (%)	0.24	0.47	0.70	0.00
Ramachandran plot				
Favored (%)	95.77	96.37	97.60	93.35
Allowed (%)	4.23	3.63	2.40	6.65
Disallowed (%)	0.00	0.00	0.00	0.00

	hSLC19A3- gf:Nb3.7: Fedratinib  (EMDB-19752) (PDB 8S5W)	hSLC19A3- gf:Nb3.7: Amprolium  (EMDB-19755) (PDB 8S62)	hSLC19A3- gf:Nb3.7: Hydroxy- chloroquine  (EMDB-19753) (PDB 8S5Z)
<b>Data collection and processing</b>			
Magnification	105,000x	105,000x	215,000x
Voltage (kV)	300	300	300
Electron exposure (e-/Å <sup>2</sup> )	45	46	58
Defocus range (µm)	0.8-1.6	1.0-2.0	0.5-1.7
Pixel size (Å)	0.83	0.83	0.572
Symmetry imposed	C1	C1	C1
Initial particle images (no.)	1,303,193	932,296	576,620
Final particle images (no.)	378,839	437,310	249,857
Map resolution (Å)	3.05	3.75	2.98
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	2.5-3.5	3.2-4.5	2.5-3.7
<b>Refinement</b>			
Initial model used (PDB code)	AF_ AFQ9BZV2F1	AF_ AFQ9BZV2F1	AF_ AFQ9BZV2F1
Model resolution (Å)	3.05	3.75	2.98
FSC threshold	0.143	0.143	0.143
Model resolution range (Å)	2.5-3.5	3.2-4.5	2.5-3.7
Map sharpening <i>B</i> factor (Å <sup>2</sup> )	-80.1	-190.9	-68.2
Model composition			
Non-hydrogen atoms	4,009	3,990	4,066
Protein residues	502	502	507
Ligands	1	1	1
<i>B</i> factors (Å <sup>2</sup> )			
Protein	30.02	58.38	41.95
Ligand	33.10	51.76	44.73
R.m.s. deviations			
Bond lengths (Å)	0.006	0.006	0.006
Bond angles (°)	1.025	1.090	0.993
Validation			
MolProbity score	1.71	1.89	1.51
Clashscore	5.26	8.55	5.69
Poor rotamers (%)	0.24	0.00	0.47
Ramachandran plot			
Favored (%)	93.55	93.55	96.81
Allowed (%)	6.45	6.45	3.19
Disallowed (%)	0.00	0.00	0.00

## Appendix 2: List of Hazardous Substances

**Supplementary Table 2:** List of hazardous substances. For more information see Regulation (EC) No 1272/2008 of the European Agency for Safety and Health at Work. Meaning of the GHS pictograms: GHS01: Explosive, GHS02: Flammable, GHS03: Oxidising, GHS04: Compressed gas, GHS05: Corrosive, GHS06: Toxic, GHS07: Harmful, GHS08: Health hazard, GHS09: Environmental hazard.

Name	GHS pictogram	GHS hazardous statements (H)	GHS precautionary statements (P)
2-Propanol	GHS02, GHS07	225-319-336	210-233-240-305+351+338-403+235
Acetic acid	GHS02, GHS05	226, 290, 314	210, 280, 303+361+353, 305+351+338, 310
Ampicillin disodium salt	GHS08	317, 334	261, 280, 302+352, 342+331
Calcium chloride	GHS07	319	305+351+338
Chloramphenicol	GHS08	351	280
cOmplete™ Protease Inhibitor Cocktail	GHS05	314	260, 280, 301+330+331, 303+361+353, 304+340+310, 305+351+338+310
EDTA tetra sodium salt	GHS07, GHS08	319, 332, 373	280, 304+340-312, 305+351+338-337+313
Ethanol	GHS02, GHS07	225, 319	210, 240, 305+351+338, 403+223
Ethidium Bromide	GHS06, GHS08	302, 330, 341	206, 281, 284, 310
GelGreen™ Nucleic Acid Gel Stain	GHS07	227	210, 280, 403+233, 501
Hydrochloric acid	GHS05, GHS07	290, 314, 335	260, 280, 303+361+353, 304+340+310, 305+351+338

Imidazole	GHS05, GHS07, GHS08	302, 314, 360D	260, 280, 301+330+331, 303+361+353, 305+351+338, 308+313
InstantBlue™	GHS05	301, 317, 334	261, 208, 301+310, 342+311
Kanamycin sulphate	GHS07	360	201, 280, 308+313
MES	GHS07	315, 319, 335	261, 305+351+338
Methanol	GHS02, GHS06, GHS08	225, 331, 311, 301, 370	210, 233, 208, 302+352, 304+340, 308+310, 403+235
Nickel (II) chloride hexahydrate	GHS06, GHS08, GHS09	350i, 360D, 341, 301, 331, 372, 315, 317, 334, 410	273, 281, 302+352, 304+340, 309+310
Pentane	GHS02, GHS07, GHS08, GHS09	225, 304, 336, 411	210, 240, 273, 301+330+331, 403+233
SDS	GHS02, GHS05, GHS07	228, 302+332, 315, 318, 335, 412	210, 261, 280, 301+312+330, 305+351+338+310, 370+378
Sekusept Plus	GHS05, GHS07, GHS09	302+332, 314, 400	273, 280, 303+361+353, 305+351+338, 310
Sodium hydroxide	GHS05	290, 314	280, 301+330+331, 305+351+338, 308+310
Sodium hypochlorite	GHS05, GHS09	H302, H314, H410	260, 264, 273, 280, 301+330+331, 303+361+353, 304+340, 305+351+338, 310, 321, 363, 391, 405, 501
Tetracycline hydrochloride	GHS07, GHS08	302, 315, 361, 362, 413	201, 202, 260, 263, 264, 273, 280, 301+312, 302+352, 308+313, 330, 332+313, 362+364
Tris base	GHS07	315, 319, 335	280, 302+352, 305+351+338
Tris(2-carboxyethyl) phosphine hydrochloride	GHS05	314	280, 305+351+338, 310
Valinomycin	GHS06	300, 310, 330	362, 364, 280, 301+310+330, 302+352+310, 361+364

---

## Appendix 3: Materials

**Supplementary Table 3:** List of reagents and resources.

Reagent or Resource	Source
<b>Chemicals and Media</b>	
2-(N-morpholino) ethane sulfonic acid (MES)	Carl Roth GmbH + Co. KG
2-Propanol	Carl Roth GmbH + Co. KG
4-amino-salicylic-acid	Fraunhofer ITMP
Adenine	Sigma-Aldrich
Agarose SERVA for DNA electrophoresis	SERVA Electrophoresis GmbH
Alfuzosin (S)	Fraunhofer ITMP
Amitriptyline-hydrochloride	Sigma-Aldrich
Ammonium sulfate	Carl Roth GmbH + Co. KG
Ampicillin sodium salt	Carl Roth GmbH + Co. KG
Amprolium	Sigma-Aldrich
Arabinose	Carl Roth GmbH + Co. KG
Aztreonam	Fraunhofer ITMP
Baricitinib	MedChemExpress
Bepidril (R)	Fraunhofer ITMP
Bezafibrate	Fraunhofer ITMP
Biotin	Sigma-Aldrich
Bovine serum albumin	Carl Roth GmbH + Co. KG
Buphenine (R,R,S)	Fraunhofer ITMP
Carbamylcholine	Fraunhofer ITMP
Carbonyl cyanide m-chlorophenyl-hydrazone (CCCP)	Sigma-Aldrich
Celecoxib	Fraunhofer ITMP
Chloramphenicol	Carl Roth GmbH + Co. KG
Chloroquine (S)	Fraunhofer ITMP
Cholesterol Hemisuccinate TRIS Salt	Anatrace
Cimaterol (R)	Fraunhofer ITMP
Citalopram hydrobromide	Selleckchem
Cytosine	Sigma-Aldrich
Dasatinib	Sigma-Aldrich
Decyl Maltose Neopentyl Glycol (DMNG)	Anatrace
Desthiobiotin	Iba Lifesciences GmbH
Dipotassium hydrogen phosphate	SERVA Electrophoresis GmbH
Disodium hydrogen phosphate	Carl Roth GmbH + Co. KG
Dobutamine (S)	Fraunhofer ITMP
Domperidone	Biomol
Domperidone	Fraunhofer ITMP

Dopamine hydrochloride	Selleckchem
Doxepine hydrochloride	Sigma-Aldrich
Ethanol	Carl Roth GmbH + Co. KG
Ethylenediamine tetraacetic acid (EDTA)	Carl Roth GmbH + Co. KG
Etoricoxib	Fraunhofer ITMP
Famciclovir	Fraunhofer ITMP
Fedratinib	Sigma-Aldrich
Fenbufen	Fraunhofer ITMP
Fluoxetine hydrochloride	Selleckchem
Fluphenazine	Fraunhofer ITMP
Fluvastatin (3S, 5R)	Fraunhofer ITMP
FreeStyle™ 293 Expression Medium	Thermo Fisher Scientific
Gefitinib	Selleckchem
Glycerol	Carl Roth GmbH + Co. KG
Glycine	Carl Roth GmbH + Co. KG
Hydroxychloroquine sulfate	Sigma-Aldrich
Imatinib mesylate	Sigma-Aldrich
Imidazole	Carl Roth GmbH + Co. KG
Imipramine hydrochloride	Sigma-Aldrich
Ketanserine	Fraunhofer ITMP
Lauryl Maltose Neopentyl Glycol (LMNG)	Anatrace
Lauryldimethylamine-N-oxide (LDAO)	Anatrace
LB-Agar (Lennox), granulated	Carl Roth GmbH + Co. KG
Levallorphan	Fraunhofer ITMP
Loperamide	Fraunhofer ITMP
Luria Broth granulated	VWR International
Magnesium chloride	Carl Roth GmbH + Co. KG
Metformin hydrochloride	Sigma-Aldrich
Moroxydine	Fraunhofer ITMP
Naftopidil (S)	Fraunhofer ITMP
n-Decyl-β-D-maltopyranoside (DM)	Anatrace
n-Dodecyl-β-maltoside	GLYCON Biochemicals GmbH
NEBuffer™ r3.1	New England Biolaby
Nefazodone	Fraunhofer ITMP
Nefazodone hydrochloride	Biomol
n-Nonyl-β-D-glucopyranoside (NG)	Anatrace
n-Octyl-β-D-glucopyranoside (β-OG)	Anatrace
Oxibendazole	Fraunhofer ITMP
Oxythiamine chloride hydrochloride	Sigma-Aldrich
Pantothenic-acid	Fraunhofer ITMP
PEG6000	Sigma-Aldrich
PEI MAX™ – Transfection Grade Linear Polyethylenimine Hydrchloride (MW 40,000)	Polysciences, Inc.



Phenylpropanolamine	Fraunhofer ITMP
Potassium chloride	Carl Roth GmbH + Co. KG
Potassium dihydrogen phosphate	Carl Roth GmbH + Co. KG
Potassium ferricyanide (III)	Merck KGaA
Practolol	Fraunhofer ITMP
Praziquantel (R)	Fraunhofer ITMP
Propanol	Carl Roth GmbH + Co. KG
Protease inhibitor (cOMplete, EDTA-free)	Roche Diagnostic GmbH
Pyranine	Sigma-Aldrich
Pyridoxine	Sigma-Aldrich
Pyrimethamine	Sigma-Aldrich
rCutSmart™	New England Biolabs
Rotiphorese® 10x TBE Buffer	Carl Roth GmbH + Co. KG
SDS Pellets	Carl Roth GmbH + Co. KG
Serotonin hydrochloride	Selleckchem
Sertraline hydrochloride	Sigma-Aldrich
Sodium chloride	Carl Roth GmbH + Co. KG
Sodium sulfate	Carl Roth GmbH + Co. KG
Sotalol (S)	Fraunhofer ITMP
Sulfazalazine	Fraunhofer ITMP
T4 DNA Ligase Reaction Buffer	New England Biolabs
Terfenadine (R)	Fraunhofer ITMP
Thiamine hydrochloride	Sigma-Aldrich
Thiamine monophosphate chloride dihydrate	Sigma-Aldrich
Thiamine pyrophosphate chloride	Sigma-Aldrich
Thymine	Sigma-Aldrich
Tofacitinib	Selleckchem
Trimethoprim	Fraunhofer ITMP
Tris(2-carboxyethyl) phosphine- hydrochloride (TCEP)	Soltec Ventures
Triton X-100	Carl Roth GmbH + Co. KG
Tween® 20	Carl Roth GmbH + Co. KG
Tryptone	Carl Roth GmbH + Co. KG
Valinomycin	Sigma-Aldrich
Xamoterol (R)	Fraunhofer ITMP
Yeast extract	Carl Roth GmbH + Co. KG
Zardaverine	Fraunhofer ITMP

---

### Consumables

Dynabeads™ MyOne™ C1	Thermo Fisher Scientific, Inc.
Dynabeads™ MyOne™ T1	Thermo Fisher Scientific, Inc.
GeneRuler 1 kb DNA ladder	Thermo Fisher Scientific, Inc.
Microplate, 96 well, PP, F-Bottom black	Greiner Bio-One GmbH
MicroSpin™ G-25 Columns	Cytiva

Ni-NTA Agarose	Invitrogen
Nunc <sup>®</sup> Maxisorp <sup>™</sup> 96 Well Plates	Merck
PageRuler <sup>™</sup> Plus Prestained Protein Ladder	Thermo Fisher Scientific, Inc.
PD midiTrap G-25	GE Healthcare
Polycarbonate membrane, 400 nm pore diameter	AVESTIN, Inc.
QUANTIFOIL <sup>®</sup> R 1.2/1.3, R 2/1 Au 300 mesh grids	Quantifoil Micro Tools GmbH
Quick Coomassie Stain	Protein Ark
SERVAGel <sup>™</sup> TG PRiME <sup>™</sup> 4-12 %	SERVA Electrophoresis GmbH
Spin-X <sup>®</sup> UF concentrator (6 mL, 500 µL; 10 k, 100 k MWCO)	CORNING
Strep-Tactin <sup>®</sup> Sepharose <sup>®</sup>	IBA Lifesciences GmbH
Trans-Blot Turbo Transfer Pack Mini format 0.2 µm PVCF	BIO-RAD Laboratories, Inc.
Whatman 1 Filter papers, 55 mm diameter	Cytiva

---

### Commercial Kits

AffinityScript Multiple Temperature Reverse Transcriptase Kit	Agilent
NucleoBond PC 10000 EF, Giga Kit	Macherey-Nagel
NucleoBond Xtra Maxi Kit	Macherey-Nagel
NucleoSpin Gel and PCR Clean-up Mini Kit	Macherey-Nagel
PUREfrefx 2.1 Kit	PUREfrefx
QIAprep Spin Miniprep Kit (250)	QIAGEN
QIAquick Gel Extraction Kit (250)	QIAGEN
QIAquick PCR Purification Kit (250)	QIAGEN
RNeasy Kit	QIAGEN

---

### Organisms

<i>E. coli</i> (DH5α)	EMBL Hamburg
<i>E. coli</i> (MC1061)	Lucigen, cat. no. 60514-1
<i>E. coli</i> (SS320)	Lucigen, cat. no. 60512-1
<i>E. coli</i> (TG1)	Lucigen, cat. no. 60502-1
<i>E. coli</i> (WK6)	EMBL Hamburg
Human embryonic kidney cells (HEK293F <sup>™</sup> , Expi293F <sup>™</sup> )	Thermo Fisher Scientific

---

### Oligonucleotides

For primer sequences for Sanger sequencing see Supplementary table XX.

For PCR primers see Supplementary table XX.

---

### Base vectors

pDX_init	Addgene #110101
pMESy4	Vrije Universiteit Brussel
pSb_init	Addgene #110100
pXLG vectors	EMBL Hamburg

---

### Enzymes

Antarctic phosphatase	New England Biolabs
-----------------------	---------------------

BspQI	New England Biolabs
DNase I	AppliChem GmbH
DpnI	New England Biolabs
EcoRV-HF	New England Biolabs
EndoH	EMBL Hamburg
HRV 3C	EMBL Hamburg
NotI-HF	New England Biolabs
Phusion DNA polymerase	EMBL Hamburg
T4 DNA Ligase	New England Biolabs
TEV	EMBL Hamburg
Trypsin	Merck

---

### Software

AlphaFold2	Google DeepMind
ChatGPT 4o mini <sup>1</sup>	OpenAI
ChemDraw 23.1.1	Reivity Signals
Coot (v.0.9.8.1)	MRC LMB
cryoSPARC (v.4.6.0)	Structura Biotechnology
EPU	Thermo Fisher Scientific, Inc.
GraphPad Prism (v.10)	Dotmatics
ISOLDE	Tristan Croll
Phenix (v.1.21.1)	Phenix
Refmac Servalcat	CCP-EM
RELION (v.5.0)	Sjors Scheres
SerialEM	University of Colorado
SnapGene	Dotmatics
UCSF ChimeraX (v.1.6.1)	University of California

---

### Devices and Equipment

1260 Infinity liquid chromatography system	Agilent Technologies, Inc.
Aekta pure	GE Healthcare
Avanti JXN-26 Centrifuge	Beckman Coulter GmbH
Centrifuge 5424 R	Eppendorf
Centrifuge 5810 R	Eppendorf
Centrifuge 5920 R	Eppendorf
Flacon 4i Electron Detector	Thermo Fisher Scientific, Inc.
JLA-8.1 rotor	Beckman Coulter GmbH
K3 electron counting direct detection camera	Gatan, Inc.
Leica DM IRB	Leica Microsystems
LiposoFast	AVESTIN, Inc.
Mastercycler nexus	Eppendorf
Maxisafe 2020 Biological Safety Cabinet	Thermo Fisher Scientific, Inc.
Mini Star silverline Microcentrifuge	VWR International

MLA-130 rotor	Beckman Coulter GmbH
Nanodrop 2000c Spectrophotometer	Thermo Fisher Scientific, Inc.
New Brunswick innova 42 incubator	Eppendorf
New Brunswick innova 44 incubator	Eppendorf
New Brunswick S41i	Eppendorf
OD600 DiluPhotometer™	IMPLEN
Optima MAX XP Ultracentrifuge	Beckman Coulter GmbH
Optima XE-90 Ultracentrifuge	Beckman Coulter GmbH
Prometheus NT.48 nanoDSF	NanoTemper Technologies GmbH
Prometheus™ Series Capillaries	NanoTemper Technologies GmbH
Rotary Evaporator R-114	Buechi
SevenCompact pH/lon	Mettler-Toledo GmbH
Superdex™ 200 Increase 5/150	Cytiva
Superdex™ 200 Increase 10/300 GL	GE Healthcare
TC20™ Automated Cell Counter	BIO-RAD Laboratories, Inc.
Tecan Spark 20M	TECAN
ThermoMixer F1.5	Eppendorf
Titan Krios Cryo-TEM G3i	Thermo Fisher Scientific, Inc.
Titan Krios Cryo-TEM G4i	Thermo Fisher Scientific, Inc.
Trans-Blot Turbo Transfer System	BIO-RAD Laboratories, Inc.
Type 45-Ti rotor	Beckman Coulter GmbH
UVP ChemStudio	Analytic Jena GmbH
Vitrobot	Thermo Fisher Scientific, Inc.
Vortex-Genie 2	Scientific Industries, Inc.

---

<sup>1</sup>ChatGPT was used to a very limited extent to improve the phrasing of individual sentences.

**Supplementary Table 4:** Primers for Sanger sequencing of plasmid DNA.

Primer name	Vector	Reading of insert	DNA sequence
FG116-CMV-for	pXLG	From N-term	CGCAAATGGGCGGTAGGCGTG
FG117-WHV-5R	pXLG	From C-term	AGCAGCGTATCCACATAGCG
FG172-MP57	pMESy4	From N-term	TTATGCTTCCGGCTCGTATG
FG173-GIII	pMESy4	From C-term	CCACAGACAGCCCTCATAG
FG188-pBAD-for	pSb_init	From N-term	ATGCCATAGCATTTTTATCC
FG189-pBAD-rev	pSb_init	From C-term	GATTTAATCTGTATCAGG

**Supplementary Table 5:** PCR primers for cloning of cDNA of the main protein constructs used in this study, and exemplary primers used for site directed mutagenesis (marked by an \*). One-letter tag code: C: C-terminal, N: N-terminal, 3: HRV 3C protease cleavage site, T: TEV protease cleavage site, H: His<sub>6</sub>-tag, S: TwinStrep-tag, G: eGFP, A: Avi-tag.

Primer name	DNA sequence
FG023-h19A3-for	ACAGGTGTCCAGGCGGCCCGCCATGGATTGTTACAGAACTTCACTAAG
FG024-h19A3-CTGH-rev	CACGCCCTGAAAATAAAGATTCTCGAGTTTTGTTGACATGATGATATT
FG066-h19A3-C3S-rev	CGGGCCCTGAAACAGAACTTCCAGGAGTTTTGTTGACATGATGATATTAC
FG068-h19A3-CA-rev	AGCCTCGAAGATGTCGTTCCAGGCCGAGTTTTGTTGACATGATGATATTAC
FG069-h19A3-CA3S-rev	CGGGCCCTGAAACAGAACTTCCAGTTCGTGCCATTCGATTTCTGAG
FG074-h19A2-for	TCCACAGGTGTCCAGGCGGCCCGCCATGGATGTGCCCGGCC
FG075-h19A2-C3S-rev	CGGGCCCTGAAACAGAACTTCCAGTGAAGTGGTTACTTGAGAACTTGATTG
FG077-h19A2-CA-rev	TTCGATTTTCTGAGCCTCGAAGATGTCGTTCCAGGCCTGAAGTGTTACTTGAGAACTTGATTGT
FG078-h19A2-CA3S-rev	AAACAGAACTTCCAGTTCGTGCCATTCGATTTTCTGAGCCTCGAAGAT
FG084-hTPK1-pSb-for	ATATATGCTCTTCAAGTGAGCATGCCTTTACCCCGTT
FG085-hTPK1-pSb-rev	TATATAGCTCTTCCCTGCCGGGCCCTGAAACAGAACTTCCAGGCTTTTGATGGCCATGGTCC
FG090-h19A3-K44E-ifor*	TGGACCAGATGAAAACCTGACCAGTGCAGAGATAAC
FG091-h19A3-K44E-irev*	GTCAGGCTTTTATCTGGTCCAGATAAATATGGGATAAGGAACGG
FG100-h19A3-gf-ifor1*	ACCAGATAAACAGCTGACCAGTGCAGAGATAACAAATGA
FG101-h19A3-gf-irev1*	CTGGTCAGCTGTTTATCTGGTCCAGATAAATATGGGATAAGG
FG102-h19A3-gf-ifor2*	ATCCCTGGCGCAGATGTCGTACTTTTACCTCAA
FG103-h19A3-gf-irev2*	TACGACATCTGCGCCAGGGATACCAAGAGTT

**Supplementary Table 6:** Amino acid sequences of the main recombinant proteins expressed and purified in this study. One-letter tag code: C: C-terminal, N: N-terminal, 3: HRV 3C protease cleavage site, T: TEV protease cleavage site, H: His<sub>6</sub>-tag, S: TwinStrep-tag, G: eGFP, A: Avi-tag, M: Myc-tag, E: EPEA-tag,  $\alpha$ : ALFA-tag. The tag sequences are highlighted in bold font. Point mutants (relative to the wildtype sequence) are marked in red.

#### **hSLC19A3-wt-C3S**

MDCYRTSLSSSWIYPTVILCLFGFFSMMRPSEPFLIPYLSGPDKNLTSAEITNEIFPVWVTYSYLVLLLPLVFVLT  
YVRYKPVIIILQGISFIITWLLLLFGQGVKTMQVVEFFYGMVTAEEVAYYAYIYSVVSPEHYQRVSGYCRSVTLAA  
YTAGSVLAQLLVSLANMSYFYLNVISLASVSVAFSLFSLFLPMPKKSMMFFHAKPSREIKKSSSVNPVLEETHEGEA  
PGCEEQKPTSEILSTSGKLNKGQLNSLKPSNVTVDVVFVQWFDLKECYSSKRLFYWVSLWAFATAGFNQVLNYVQ  
ILWDYKAPSQDSSIYNGAVEAIATFGGAVAAFAVGYVKVNWDLGELALVVFVSVNAGSLFLMHYTANIWACYAG  
YLIFKSSYMLLITIAVFQIAVNLNVERYALVFGINTFIALVIQTIMTVIVVDQRGLNLPVSIQFLVYGSYFAVIA  
GIFLMRSMYITYSTKSQKDVQSPAPSENPDVSHPEEESNIIMSTKL**LEVLFQGPSSGWSHPQFEKGGGSGGGSGG**  
**SAWSHPQFEK**

#### **hSLC19A3-wt-CA3S**

MDCYRTSLSSSWIYPTVILCLFGFFSMMRPSEPFLIPYLSGPDKNLTSAEITNEIFPVWVTYSYLVLLLPLVFVLT  
YVRYKPVIIILQGISFIITWLLLLFGQGVKTMQVVEFFYGMVTAEEVAYYAYIYSVVSPEHYQRVSGYCRSVTLAA  
YTAGSVLAQLLVSLANMSYFYLNVISLASVSVAFSLFSLFLPMPKKSMMFFHAKPSREIKKSSSVNPVLEETHEGEA  
PGCEEQKPTSEILSTSGKLNKGQLNSLKPSNVTVDVVFVQWFDLKECYSSKRLFYWVSLWAFATAGFNQVLNYVQ  
ILWDYKAPSQDSSIYNGAVEAIATFGGAVAAFAVGYVKVNWDLGELALVVFVSVNAGSLFLMHYTANIWACYAG  
YLIFKSSYMLLITIAVFQIAVNLNVERYALVFGINTFIALVIQTIMTVIVVDQRGLNLPVSIQFLVYGSYFAVIA  
GIFLMRSMYITYSTKSQKDVQSPAPSENPDVSHPEEESNIIMSTKL**GLNDIFEAQKIEWHELEVLFQGPSSGWSH**  
**PQFEKGGGSGGGSGGSAWSHPQFEK**

#### **hSLC19A3-gf-C3S**

MDCYRTSLSSSWIYPTVILCLFGFFSMMRPSEPFLIPYLSGPDK**Q**LTSAEITNEIFPVWVTYSYLVLLLPLVFVLT  
YVRYKPVIIILQGISFIITWLLLLFGQGVKTMQVVEFFYGMVTAEEVAYYAYIYSVVSPEHYQRVSGYCRSVTLAA  
YTAGSVLAQLLVSLA**Q**MSYFYLNVISLASVSVAFSLFSLFLPMPKKSMMFFHAKPSREIKKSSSVNPVLEETHEGEA  
PGCEEQKPTSEILSTSGKLNKGQLNSLKPSNVTVDVVFVQWFDLKECYSSKRLFYWVSLWAFATAGFNQVLNYVQ  
ILWDYKAPSQDSSIYNGAVEAIATFGGAVAAFAVGYVKVNWDLGELALVVFVSVNAGSLFLMHYTANIWACYAG  
YLIFKSSYMLLITIAVFQIAVNLNVERYALVFGINTFIALVIQTIMTVIVVDQRGLNLPVSIQFLVYGSYFAVIA  
GIFLMRSMYITYSTKSQKDVQSPAPSENPDVSHPEEESNIIMSTKL**LEVLFQGPSSGWSHPQFEKGGGSGGGSGG**  
**SAWSHPQFEK**

#### **hSLC19A3-gf-CA3S**

MDCYRTSLSSSWIYPTVILCLFGFFSMMRPSEPFLIPYLSGPDK**Q**LTSAEITNEIFPVWVTYSYLVLLLPLVFVLT  
YVRYKPVIIILQGISFIITWLLLLFGQGVKTMQVVEFFYGMVTAEEVAYYAYIYSVVSPEHYQRVSGYCRSVTLAA  
YTAGSVLAQLLVSLA**Q**MSYFYLNVISLASVSVAFSLFSLFLPMPKKSMMFFHAKPSREIKKSSSVNPVLEETHEGEA  
PGCEEQKPTSEILSTSGKLNKGQLNSLKPSNVTVDVVFVQWFDLKECYSSKRLFYWVSLWAFATAGFNQVLNYVQ  
ILWDYKAPSQDSSIYNGAVEAIATFGGAVAAFAVGYVKVNWDLGELALVVFVSVNAGSLFLMHYTANIWACYAG  
YLIFKSSYMLLITIAVFQIAVNLNVERYALVFGINTFIALVIQTIMTVIVVDQRGLNLPVSIQFLVYGSYFAVIA  
GIFLMRSMYITYSTKSQKDVQSPAPSENPDVSHPEEESNIIMSTKL**GLNDIFEAQKIEWHELEVLFQGPSSGWSH**  
**PQFEKGGGSGGGSGGSAWSHPQFEK**

### hSLC19A2-wt-C3S

MDVPGPVSRRAAAAAATVLLRTARVRRECWFLLPTALLCAYGFFASLRPSEPFLTPYLLGPDKNLTEREVFNEIYP  
VWTYSYLVLLFPVFLATDYLRYPVLLQGLSLIVTWFMLLYAQGLLAIQFLEFFYGIATATEIAYYSYIYSVVD  
LGMYPQVTSYCRSATLVGFTVGSVLGQILVSVAGWSLFSLNVISLTCVSVAFAVAWFLPMPQKSLFFHHIPSTCQ  
RVNGIKVQNGGIVTDTASNHLPGWEDIESKIPLNMEEPVVEEPEPKPDRLLVVKVWLNDFLMCYSSRPLLCWSV  
WWALSTCGYFQVVNYTQGLWEKVMPSRYAAIYNGGVEAVSTLLGAVAVFAVGYIKISWSTWGEMTSLFSLLIAA  
AVYIMDTVGNIWVCYASYVVFRIIYMLLITITATFQIAANLSMERYALVFGVNTFIALALQTLTLIVVDASGLGL  
EITTQFLIYASYFALIAVVFLASGAVSVMKKCRKLEDPQSSSQVTT**SLEVLFGQPSSGWSHPQFEKGGGSGGGSG**  
**GSAWSHPQFEK**

### hSLC19A2-wt-CA3S

MDVPGPVSRRAAAAAATVLLRTARVRRECWFLLPTALLCAYGFFASLRPSEPFLTPYLLGPDKNLTEREVFNEIYP  
VWTYSYLVLLFPVFLATDYLRYPVLLQGLSLIVTWFMLLYAQGLLAIQFLEFFYGIATATEIAYYSYIYSVVD  
LGMYPQVTSYCRSATLVGFTVGSVLGQILVSVAGWSLFSLNVISLTCVSVAFAVAWFLPMPQKSLFFHHIPSTCQ  
RVNGIKVQNGGIVTDTASNHLPGWEDIESKIPLNMEEPVVEEPEPKPDRLLVVKVWLNDFLMCYSSRPLLCWSV  
WWALSTCGYFQVVNYTQGLWEKVMPSRYAAIYNGGVEAVSTLLGAVAVFAVGYIKISWSTWGEMTSLFSLLIAA  
AVYIMDTVGNIWVCYASYVVFRIIYMLLITITATFQIAANLSMERYALVFGVNTFIALALQTLTLIVVDASGLGL  
EITTQFLIYASYFALIAVVFLASGAVSVMKKCRKLEDPQSSSQVTT**SGLNDIFEAQKIEWHELEVLFGQPSSGWS**  
**HPQFEKGGGSGGGSGGSAWSHPQFEK**

### hSLC19A2-wt-Ca3S

MDVPGPVSRRAAAAAATVLLRTARVRRECWFLLPTALLCAYGFFASLRPSEPFLTPYLLGPDKNLTEREVFNEIYP  
VWTYSYLVLLFPVFLATDYLRYPVLLQGLSLIVTWFMLLYAQGLLAIQFLEFFYGIATATEIAYYSYIYSVVD  
LGMYPQVTSYCRSATLVGFTVGSVLGQILVSVAGWSLFSLNVISLTCVSVAFAVAWFLPMPQKSLFFHHIPSTCQ  
RVNGIKVQNGGIVTDTASNHLPGWEDIESKIPLNMEEPVVEEPEPKPDRLLVVKVWLNDFLMCYSSRPLLCWSV  
WWALSTCGYFQVVNYTQGLWEKVMPSRYAAIYNGGVEAVSTLLGAVAVFAVGYIKISWSTWGEMTSLFSLLIAA  
AVYIMDTVGNIWVCYASYVVFRIIYMLLITITATFQIAANLSMERYALVFGVNTFIALALQTLTLIVVDASGLGL  
EITTQFLIYASYFALIAVVFLASGAVSVMKK**SRLEEEELRRRLTELEVLFGQPSSGWSHPQFEKGGGSGGGSGGSA**  
**WSHPQFEK**

### Nb2.1-CHE

QVQLVESGGGLVQPGESLRLSCAASGFTFRNYIMNWVRQAPGKGLELVSTIYRSGGATTYADSVKGRFTISRDDA  
RNTLYLQMNLLKPEDTAVYYCVTRDVTNGLVKGGTQVTVSS**HHHHHHEPEA**

### Nb2.2-CHE

QVQLVESGGGLVQPGGSLRLSCAASGFTFSNYWMSWVRQVPGKLEWISNINS DGSITSYADSVKGRFTSSRDNA  
RNTLYLQMNLLKPEDTAVYYCATRDLYGIHRGGTQVTVSS**HHHHHHEPEA**

### Nb3.3-CHE

QVQLVESGGGLVQAGDSLRLSCAASGRTFSNYMAWFRQAPGKEREVFAAIRLSYGSTYYADSVRGRFTISKDNA  
KNTVNLRMNSLKSEDTAIYYCAAADRWALAVRTATTYNYWGGTQVTVSS**HHHHHHEPEA**

### Nb3.4-CHE

QVQLVESGGGLVQAGDSLRLSCAASGLTFSRYDMAWFRQAPGKEREVAAITWTGSITDYADSVKGRFTISRDN  
KNTIYLMNSLKPEDTAVYYCAAQVPYGSTLKFNLRTLFGSWGQGTQVTVSS**HHHHHHEPEA**

### Nb3.7-CHE

QVQLVESGGGLVQPGDSLRLSCAASGRTFSINAMAWFRQAPGKEREYVAQISWTGGNTYYAESVKGRFTISRDN  
KNTVYLMISLKPEDTAVYYCAADSGGIRLGASRWNYWGQGTQVTVSSGRASSSS**HHHHHHEPEA**

### Nb<sub>ALFA</sub>-C3MH

QVQLQESGGGLVQPGGSLRLSCTASGVTISALNAMAMGWYRQAPGERRVMVAAVSERGNAMYRESVQGRFTVTRD  
FTNKMVSLQMDNLKPEDTAVYYCHVLEDRVDSFHDIWGQGTQVTVSS**LEVLFOGPAGRAGEQKLISEEDLNSAVD**  
**HHHHH**

### hTPK1-C3MH

EHAFTPLEPLLSTGNLKYCLVILNQPLDNYFRHLWNKALLRACADGGANRLYDITEGERESFLPEFINGDFDSIR  
PEVREYYATKGCCELISTPDQDHTDFTKCLKMLQKKIEEKDLKVDVIVTLGGLAGRFDQIMASVNTLFQATHITPF  
PIIIIQEESLIYLLQPGKHLHVDTGMEGDWCGLIPVGQPCMQVTTTGLKWNLTNDVLAFGTLVSTSNYDGSV  
VTVETDHPDLLWTMAIKS**LEVLFOGPAGRAGEQKLISEEDLNSAVDHHHHH**



## **Acknowledgements**

For their support and individual contributions to the research for this Dissertation, I want to thank the following people (current affiliations in parentheses):

### **Supervised students:**

Lea Priestersbach (EMBL Hamburg)

Antonia Fuhrmann (Universität zu Lübeck)

### **Löw group at EMBL Hamburg:**

Katharina Jungnickel (Cryo-EM Facility CSSB Hamburg)

Siavash Mostafavi (EMBL Hamburg)

Tânia Custódio (EMBL Hamburg)

Christian Löw (EMBL Hamburg, Bernhard Nocht Institut für Tropenmedizin)

### **Sample Preparation and Characterisation Facility at EMBL Hamburg:**

Stephan Niebling (EMBL Hamburg)

Angelica Struve (EMBL Hamburg)

### **Cryo-EM training and support:**

Joseph Bartho (Structural and Computational Biology Unit, EMBL Heidelberg)

Cornelia Cazez (Cryo-EM Facility CSSB Hamburg)

Carolin Seuring (Cryo-EM Facility CSSB Hamburg)

### **Nanobody discovery – Training and collaboration:**

Alison Pirro-Lundqvist (Vrije Universiteit Brussel)

Els Pardon (Vrije Universiteit Brussel)

Jan Steyaert (Vrije Universiteit Brussel)

### **Structure-based discovery:**

Björn Windshügel (Fraunhofer Institute for Translational Medicine)

Tania Szal (Fraunhofer Institute for Translational Medicine)

### **Support for mass spectrometry-based and radiolabelled uptake assays:**

Bernhard Drotleff (EMBL Metabolomics Core Facility)

Markus Heine (Universitätsklinikum Hamburg-Eppendorf, UKE)

### **Thesis advisory committee:**

Maria Garcia (EMBL Hamburg)

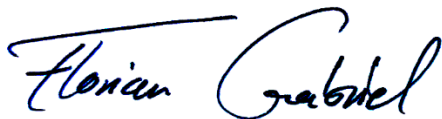
Sebastian Eustermann (EMBL Heidelberg)

Daniel Wilson (Universität Hamburg)

## Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben. Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Hamburg, 18. November 2024

A handwritten signature in black ink, reading "Florian Grubel". The signature is written in a cursive style with a long horizontal stroke at the beginning.