PNAS
www.pnas.org

Supplementary Information for

Structural basis for amino acid exchange by a human heteromeric amino acid transporter

Di Wu[1], Tamara N. Grund[1], Sonja Welsch[2], Deryck J. Mills[3], Max Michel[1], Schara Safarian[1*], Hartmut Michel[1*]

*Corresponding authors: Hartmut Michel and Schara Safarian

Email: Hartmut.Michel@biophys.mpg.de; Schara.Safarian@biophys.mpg.de

This PDF file includes:

Figures S1 to S12
Tables S1 to S2
Legends for Movies S1 to S2

Other supplementary materials for this manuscript include the following:

Movies S1 to S2

www.pnas.org/cgi/doi/10.1073/pnas.2008111117

Fig. S1.

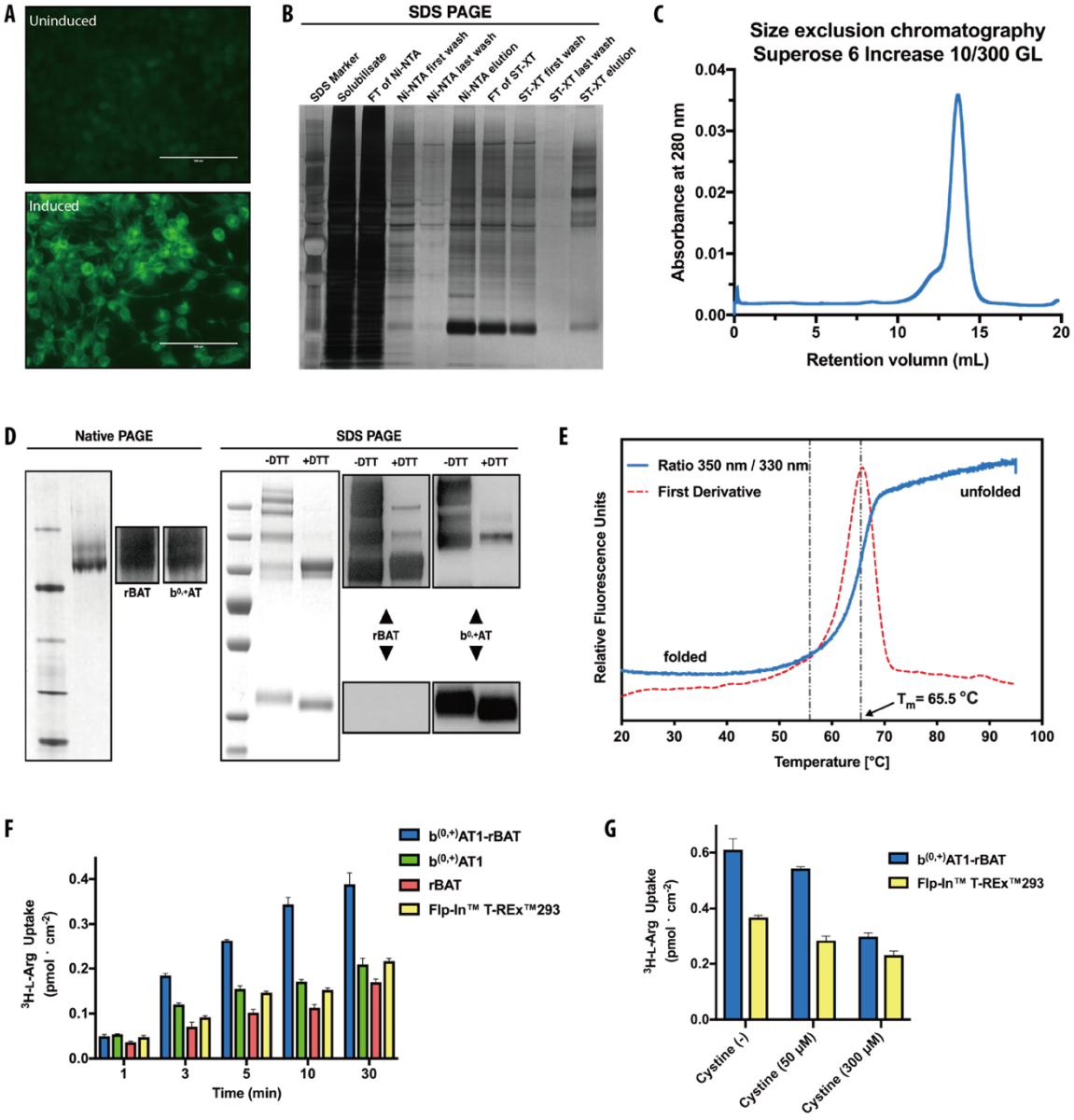

Fig. S1. Sample preparation and validation. **(A)** Fluorescence micrographs of uninduced and induced Flp-In™ T-REx™293-b(0,+)AT1-eGFP-rBAT cells. The results confirmed the successful production and trafficking to the plasma membrane. **(B)** SDS PAGE analyses of fractions from the tandem affinity purification procedure. **(C)** SEC profile of purified b(0,+)AT1-rBAT complex. **(D)** Native PAGE and Western immunoblotting results from the peak fraction of the SEC run. Samples subjected to electrophoresis were additionally treated with10 mM DTT to confirm the inter-subunit disulfide bond. **(E)** DSF profile and T_m value of purified b(0,+)AT1-rBAT complex indicate for a

stable complex preparation. The measured T_m is 65.5 ± 0.1 ºC. Data is mean ± SD, n = 5 replicates. **(F)** Cell based uptake assay of 50 µM L-[^{3}H]Arginine (54.5 Ci/mmol, 0.5 Ci/well) using different cell lines **(G)** Cell based competition assay of 50 µM L-[^{3}H]Arginine (54.5 Ci/mmol, 0.5 Ci/well) with varying concentrations of cystine using different cell lines. Data in **(F)** and **(G)** are mean ± SEM. n = 3 replicates. The results confirm the activity of the heterologously produced $b^{(0,+)}$AT1-rBAT complex.

Fig. S2.

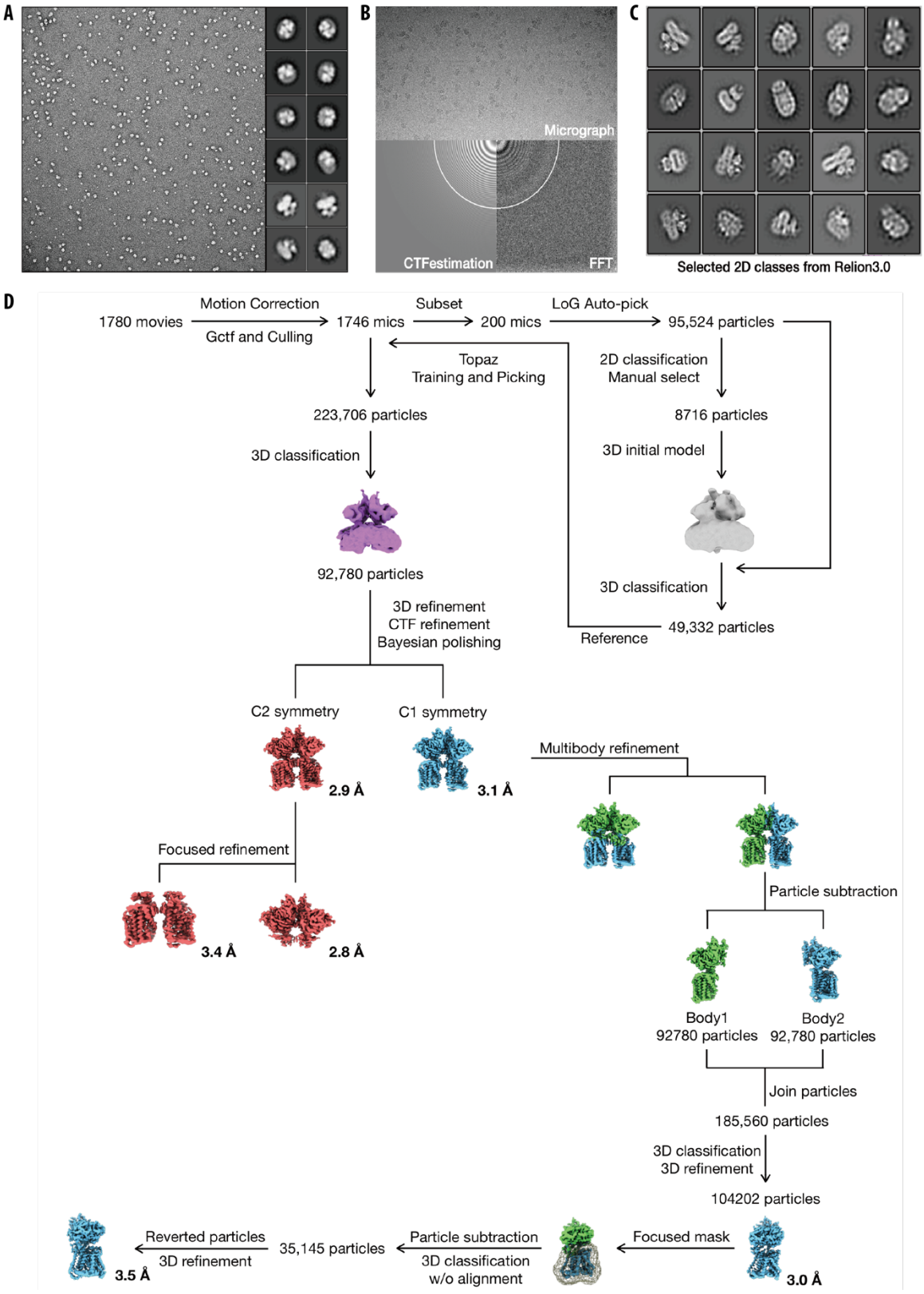
A
B
Micrograph
CTFestimation
FFT
C
Selected 2D classes from Relion3.0
D
1780 movies
Motion Correction
Gctf and Culling
1746 mics
Subset
200 mics
LoG Auto-pick
95,524 particles
Topaz
Training and Picking
2D classification
Manual select
223,706 particles
8716 particles
3D classification
3D initial model
92,780 particles
3D classification
49,332 particles
Reference
3D refinement
CTF refinement
Bayesian polishing
C2 symmetry
C1 symmetry
2.9 Å
3.1 Å
Multibody refinement
Focused refinement
3.4 Å
2.8 Å
Particle subtraction
Body1
92780 particles
Body2
92,780 particles
Join particles
185,560 particles
3D classification
3D refinement
104202 particles
Focused mask
3.0 Å
Particle subtraction
3D classification
w/o alignment
35,145 particles
Reverted particles
3D refinement
3.5 Å

Fig. S2. Cryo-EM data processing workflow. (A) Exemplary micrograph and 2D class averages from an initial negative staining EM experiment. **(B)** Exemplary micrograph and CTF estimation (Gctf) of the cryo-EM dataset. **(C)** 2D class averages from the cryo-EM dataset. **(D)** Schematic workflow of pre-processing, classification and refinement of cryo-EM data (see Supplementary Methods for details). Different processing approaches are indicated by color. Reported resolution values correspond to FSC = 0.143.

Fig. S3.

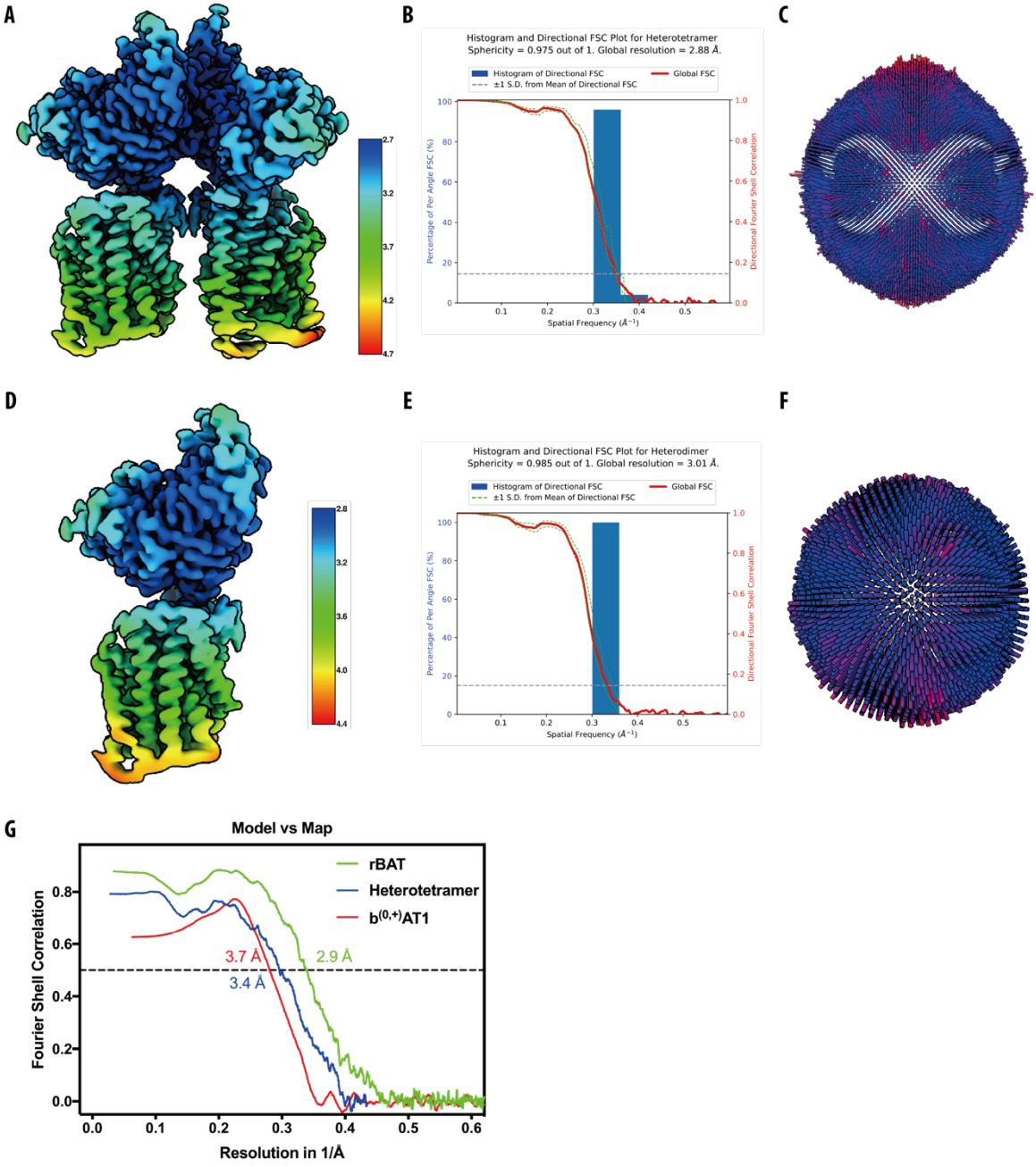

Fig. S3. Map quality analysis of the $b^{(0,+)}$AT1-rBAT structure. (A) Local resolution map of the heterotetrameric $b^{(0,+)}$AT1-rBAT complex with corresponding **(B)** 3DFSC plot, and **(C)** Euler angle distributions. **(D)** Local resolution map of the heterodimeric $b^{(0,+)}$AT1-rBAT complex with corresponding **(E)** 3DFSC plot, and **(F)** Euler angle distributions. **(G)** Map-to-model correlation curve of maps from three different focused refinement jobs.

Fig. S4.

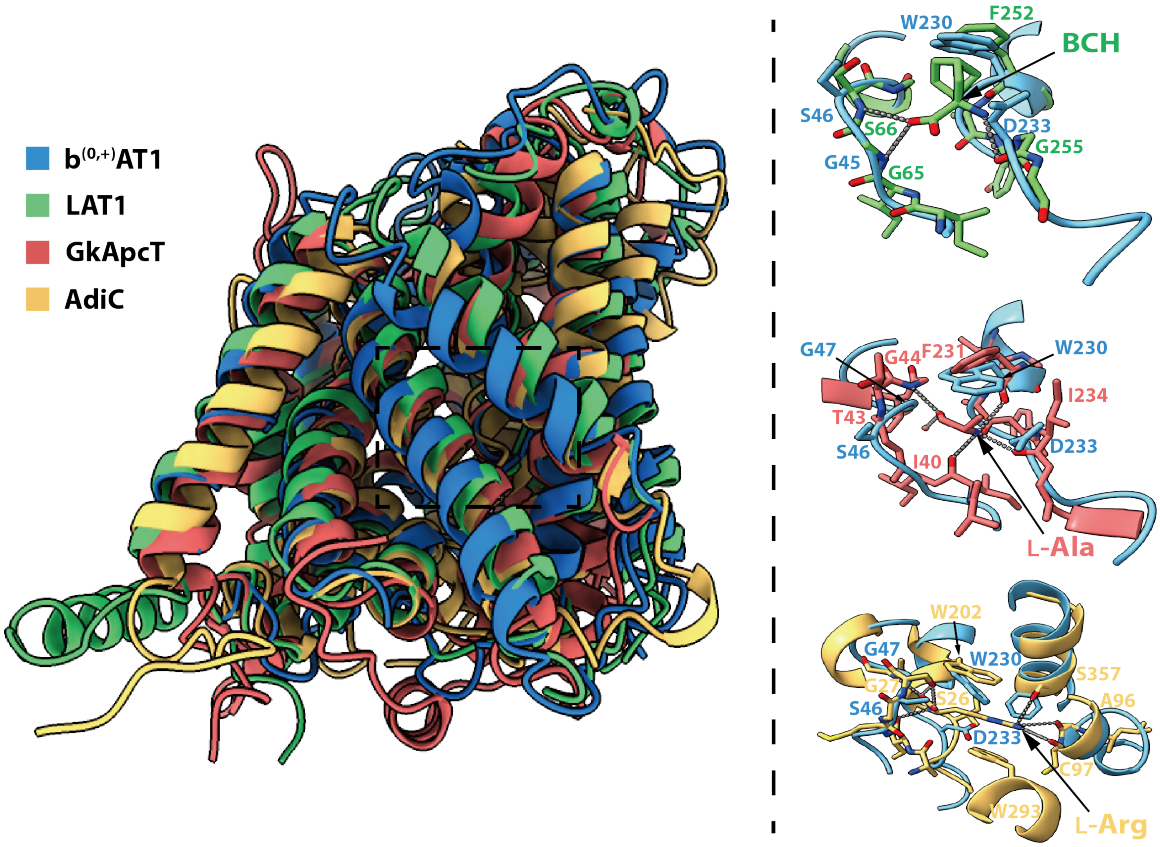

Fig. S4. Comparison of overall conformation and substrate-binding sites between $b^{(0,+)}$AT1, LAT1, GkApcT and AdiC. The overall structure of $b^{(0,+)}$AT1 was captured in an inward-facing conformation. Its overall fold resembles that of LAT1, GkApcT and AdiC**.** The putative binding pocket of $b^{(0,+)}$AT1 is characterized by a local environment which is common among SLC7 family amino acid transporters. LAT1 (PDB 6IRT), GkApcT (PDB 5OQT) and AdiC (PDB 3L1L) are colored green, red and yellow, respectively.

Fig. S5.

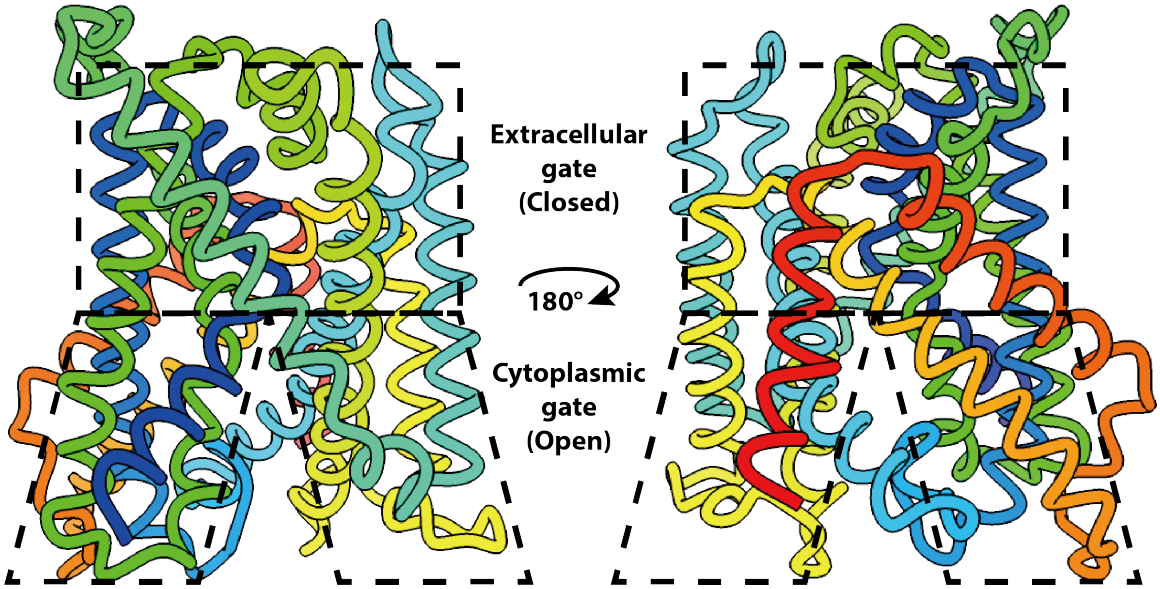

Fig. S5. The inward-facing conformation of $b^{(0,+)}$AT1. Schematic illustration of extracellular and cytoplasmic barriers of $b^{(0,+)}$AT1 in the inward-facing conformation.

Fig. S6.

Fig. S6. Sequence alignment of human SLC7 family members. Red stars indicate the conserved G(S/A)G motif and residues involved in formation of substrate binding sites.

Fig. S7.

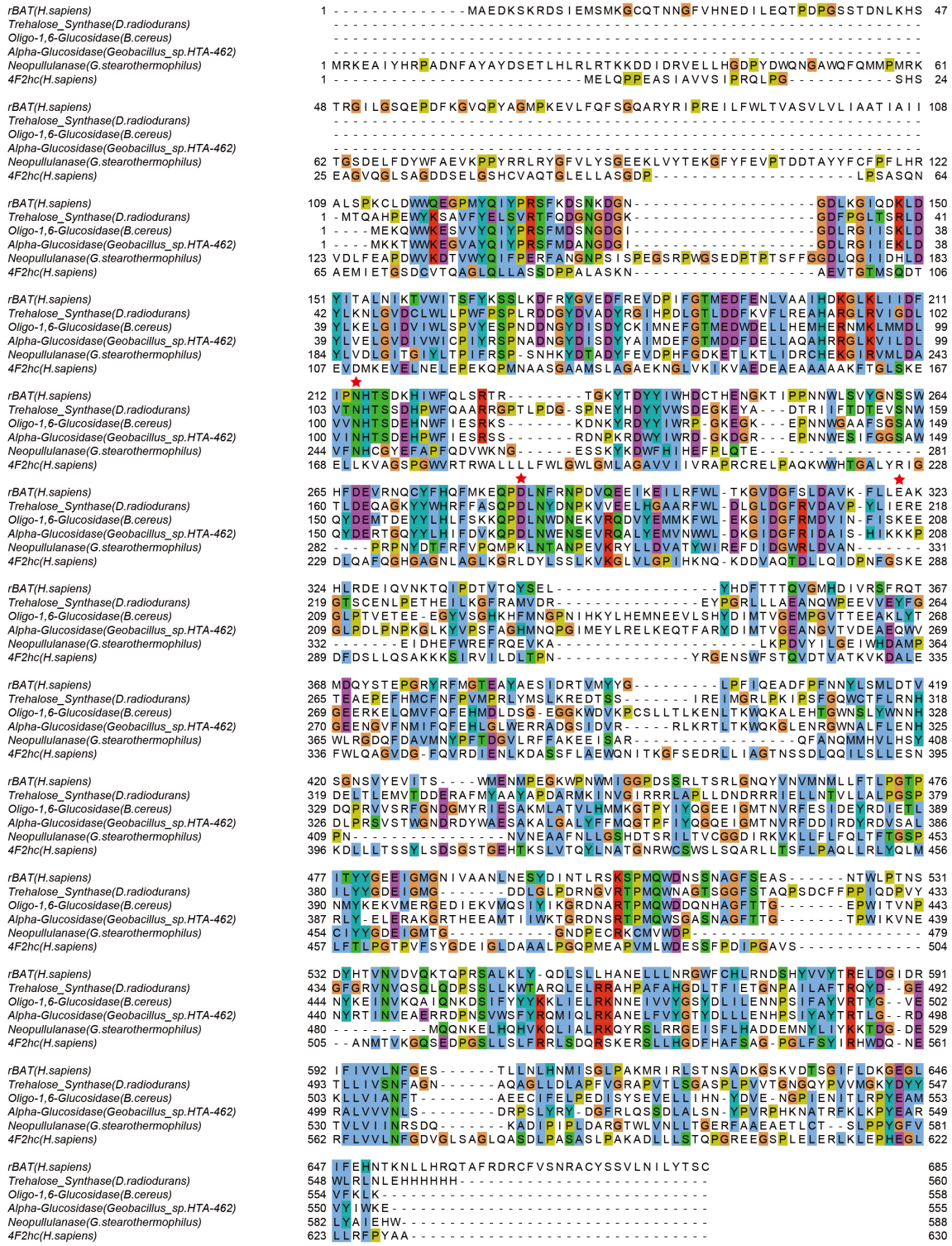

Fig. S7. Full length sequence alignment of human rBAT, 4F2hc and GH13 α-amylase family members. Red stars indicate conserved residues participating in Ca^{2+} binding and coordination.

Fig. S8.

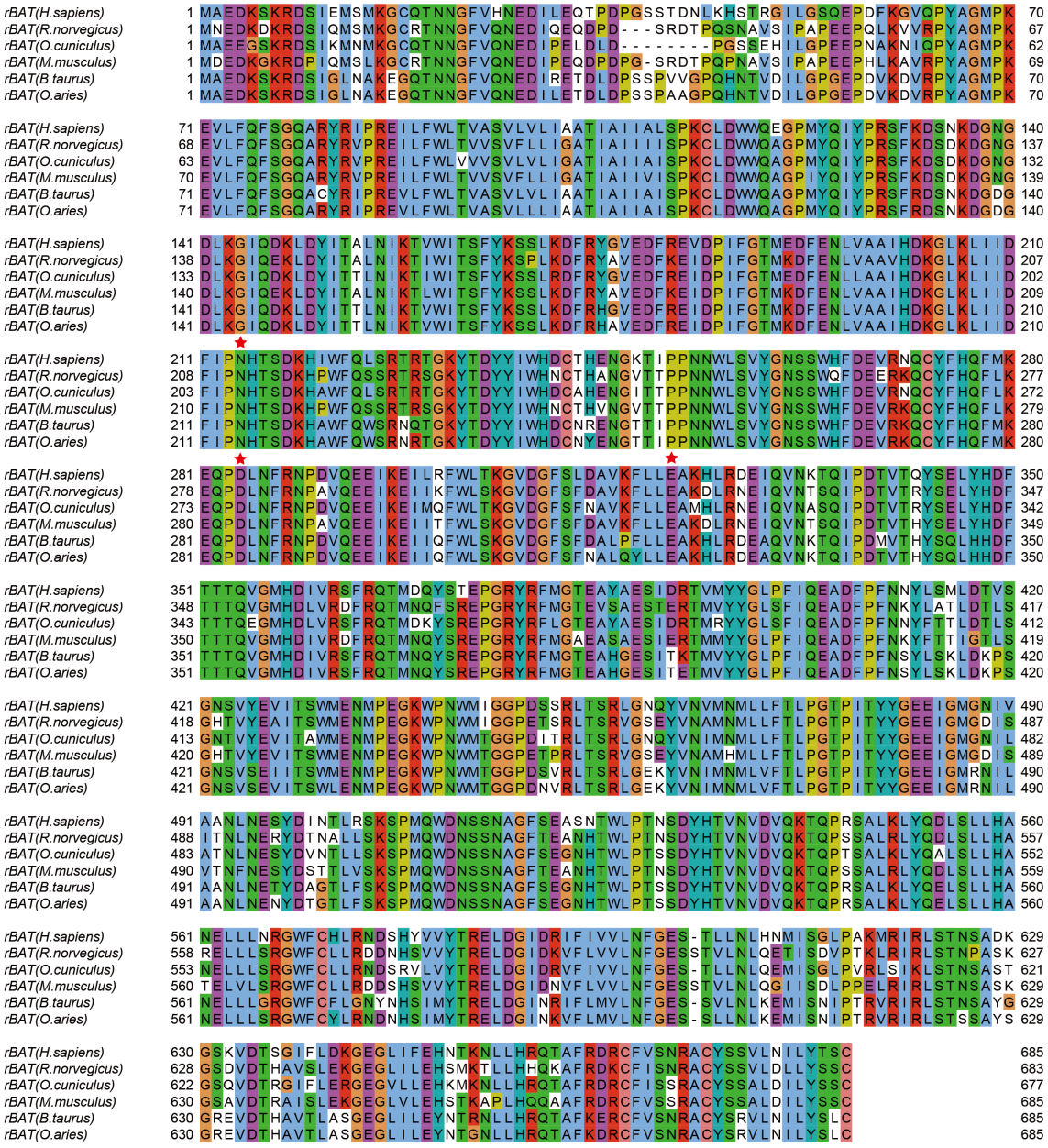

Fig. S8. Full length sequence alignment of mammalian rBAT homologs. Red stars indicate conserved residues participating in Ca^{2+} binding and coordination.

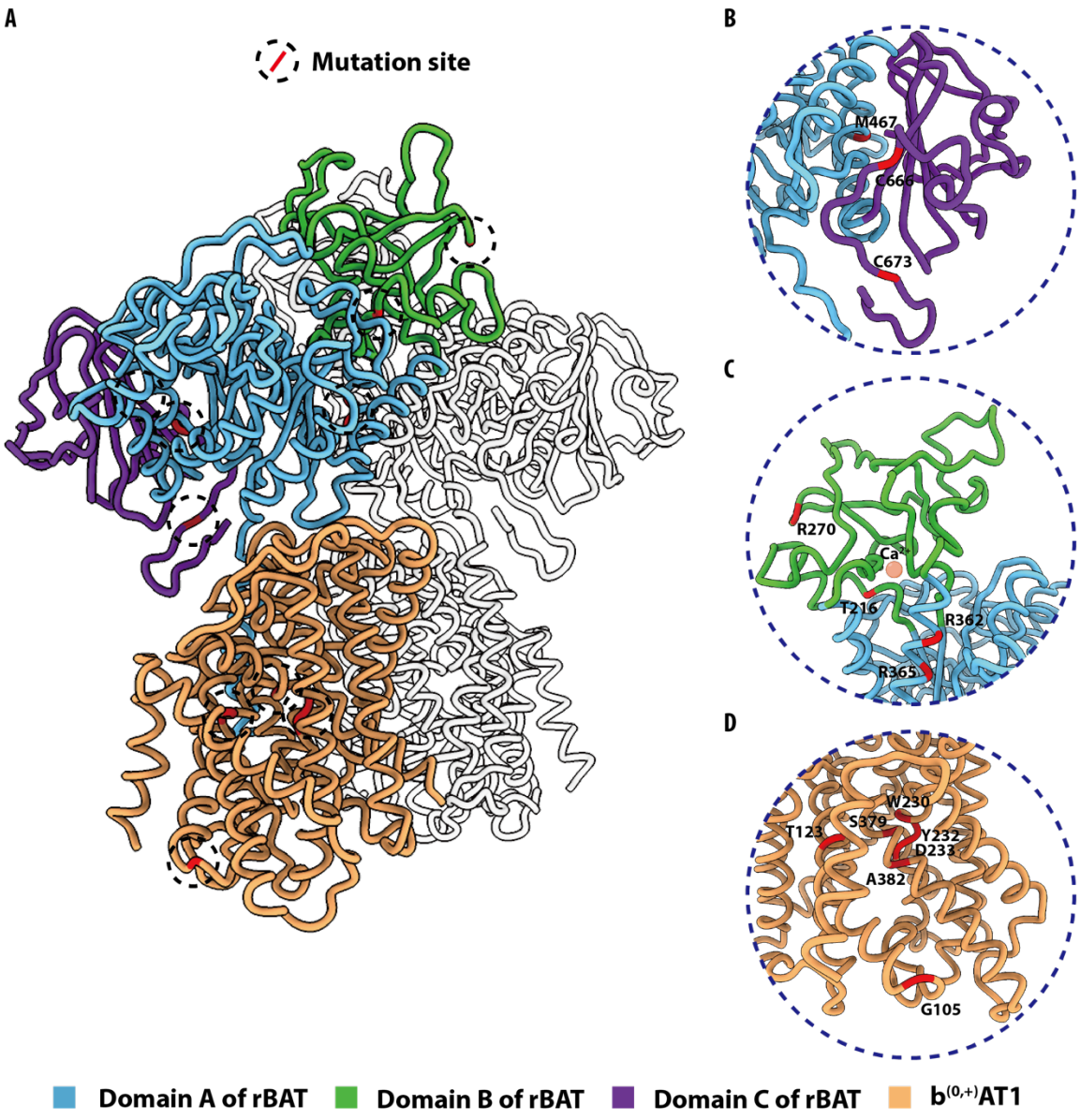

Fig. S9. Cystinuria causative mutations of the $b^{(0,+)}$AT1-rBAT complex. (**A**) Overall structure of the heterodimeric $b^{(0,+)}$AT1-rBAT complex. Locations of cystinuria-related mutations are indicated in red. (**B**) Closeup view of domain C mutations. The M467 mutation is found at the interface between domains B and C. Residues C666 and C673 are involved in formation of disulfide bonds within the C-loop of rBAT. (**C**) Closeup view of domain A and B mutations. T216 is in immediate vicinity of the Ca^{2+}-bind site. **(D)** Closeup view of $b^{(0,+)}$AT1 mutations. W230 completes the substrate binding site of $b^{(0,+)}$AT1and acts as a gating residue. G105 locates in the IL1 at a close position to

TM3. A detailed summary of the respective mutations and their physiological effects is given in the Table. S2.

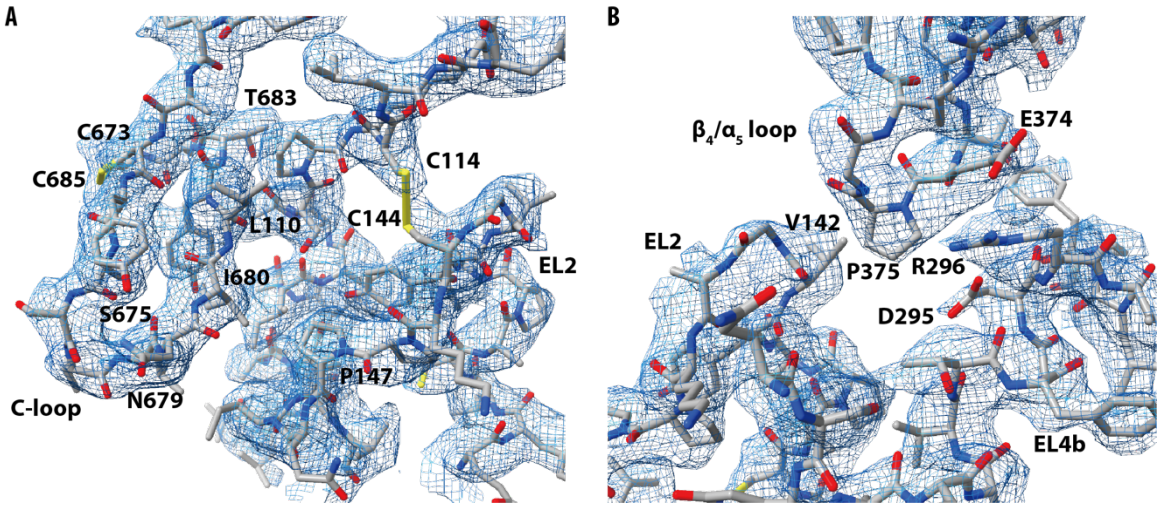

Fig. S10. Interaction interfaces of rBAT and $b^{(0,+)}$AT1. (A) Atomic model and density map of the C-terminal loop of rBAT and EL2 of $b^{(0,+)}$AT1. **(B)** Atomic model and density map of the β_4/α_5 loop of rBAT and EL2 and EL4b of $b^{(0,+)}$AT1. EL = extracellular loop.

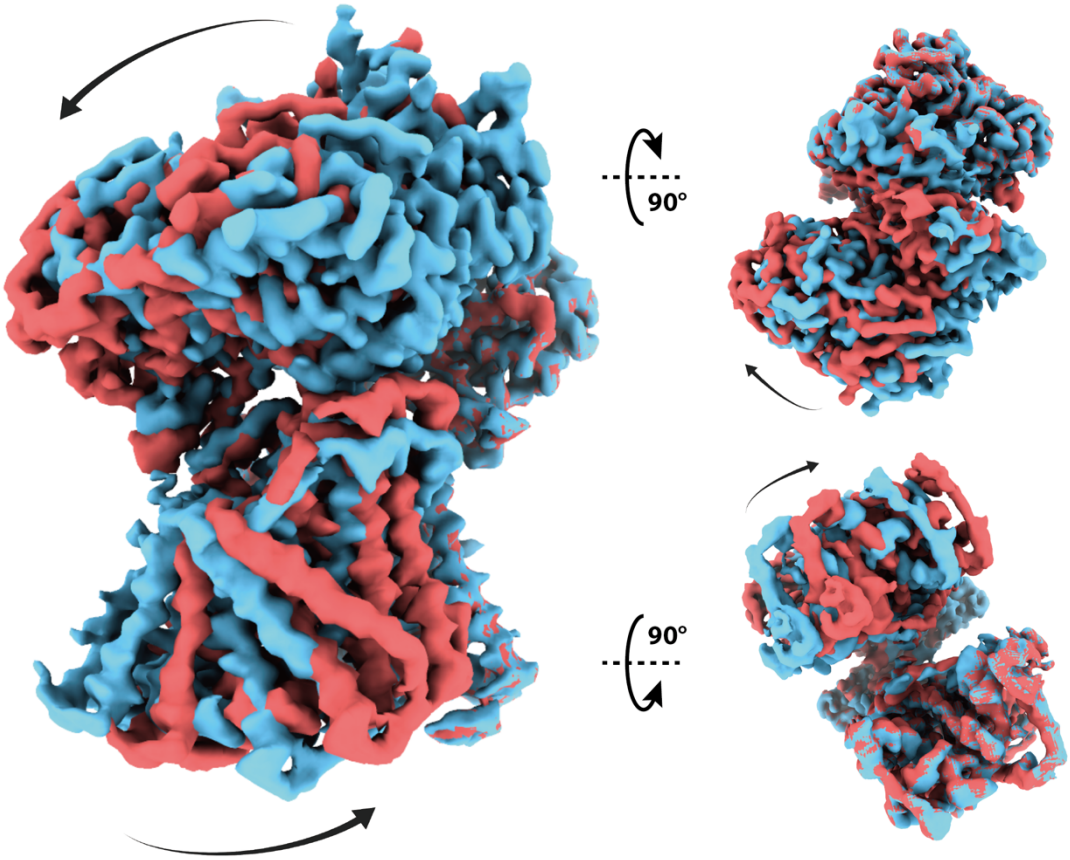

Fig. S11. Conformational heterogeneity of the $b^{(0,+)}$AT1-rBAT complex revealed by principle component analysis. (**A**) The most dominant component of the conformational heterogeneity is shown. We observe a relative rotational movement between the two HAT units along the membrane axis. The rotation center is found to be at the interaction interface between of the two extracellular rBAT domains. For a more comprehensive illustration of the movement see movies S1 and S2. The initial conformational is colored in blue. The highest rotational sampling angle is colored in red.

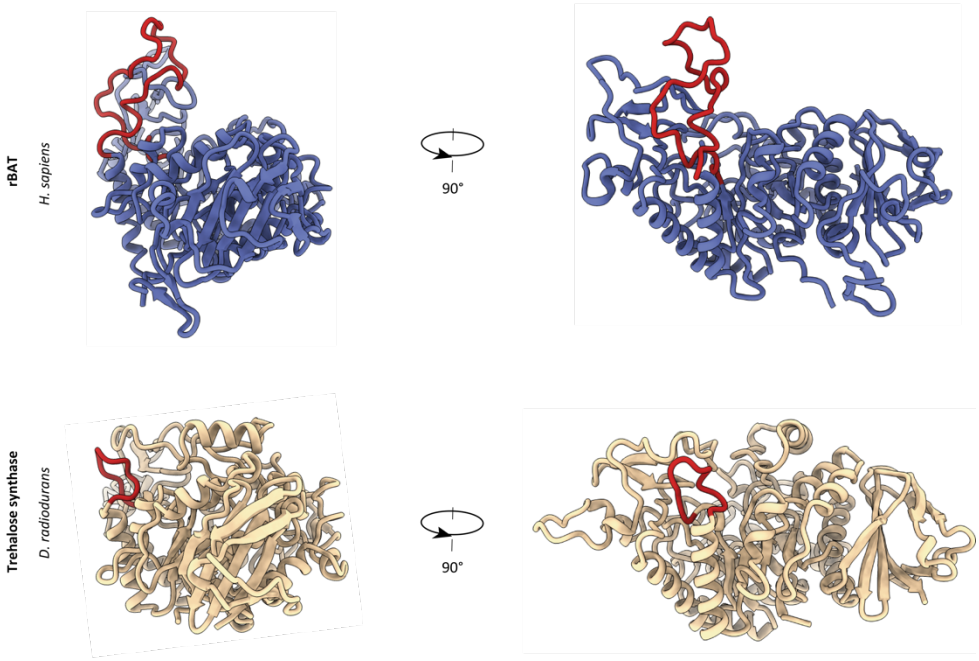

Fig. S12. Structural comparison of the β4/α4 extension between human rBAT and the bacterial trehalose synthase. Ribbon representation of the extracellular domain of human rBAT and the trehalose synthase from *D. radiodurans* (PDB 4TVU). The β4/α4 extensions of the two structures highlighted in red indicate that the rBAT domain possesses a larger loop segment than the rehalose synthase. Despite the structural resemblance between rBAT and the trehalose synthase, the bacterial member of the glycoside hydrolase (GH13) family is lacking critical residue required for dimer formation.

Table S1. Cryo-EM data collection, refinement and validation statistics

	rBAT (homodimer)	$b^{(0,+)}AT1$	$b^{(0,+)}AT1$-rBAT (heterotetramer)
Data collection			
Accession number	EMD-10936	EMD-10940	EMD-10933
Magnification	105k	105k	105k
Voltage / kV	300	300	300
Dose / $e^{-}Å^{-2}$	40	40	40
Pixel size / Å	0.837	0.837	0.837
Defocus range / μm	-1.1 to -2.2	-1.1 to -2.2	-1.1 to -2.2
Recorded movies	1780	1780	1780
Initial particle images	223706	223706	223706
Final particle images	92789	185560	92789
Camera	Gatan K3	Gatan K3	Gatan K3
Microscope	Titan Krios G3i	Titan Krios G3i	Titan Krios G3i
Image processing			
Symmetry imposed	C2	C1	C2
Resolution ($FSC_{0.143}$) / Å	2.82	3.43	2.88
Applied B-factor / Å	-30	-10	-30
Map resolution range	2.7-3.7	3.2-4.4	2.7-4.7
Model refinement			
PDB accession	6YUZ	6YV1	6YUP
Validation			
$FSC^{map\text{-}to\text{-}model}{}_{(0.5)}$ / Å	2.9	3.7	3.4
MolProbity score	1.95	1.63	2
Composition	-	-	
Atoms	9782^{+}	2496^{+}	14774
Protein residues	1188	328	1844
Ligands	8 NAG, 2 CA		8 NAG, 2 CA
Bonds (R.M.S.D.)			
Length (Å)	0.005	0.002	0.003
Angles (°)	0.54	0.58	0.909
B-factors (min/max/mean)			
Protein	39.2/138.4/60.7	15.63/79.6/40.11	15..63/138.2/53.96
Ligand	36.71/82.83/77.06		36.71/82.83/77.06
Clashscore	7.9	9.39	
Ramachandran plot (%)	-	-	
Favored	95.68	99.04	96.84
Allowed	3.9	0.96	2.88
Outliers	0.4	0	0.28
Rotamer outliers (%)	1.95	1.48	1.8

Table S2. Representatives of cystinuria-associated mutations in $b^{(0,+)}$AT1 and rBAT

Mutation	Location	Structural comments	Mutation defect summary
$b^{(0,+)}$AT1			
Gly105Arg Gly105Glu	IL1	In the IL1 region and is close to TM3	Mutated charge and/or larger side chains may interfere with the conformation switches
Thr123Met	TM3	In the middle of TM3 and near the binding pocket zone	Mutated hydrophobic side chain may form van der Waals contacts with residues from neighbour helix and affect its behaviour even the overall conformation
Trp230Arg	TM6 (binding pocket)	Act as occlusion gate of the binding pocket on the extracellular side	Loss of hydrophobic interaction and steric hindrance would prevent the successful transport of substrates
Tyr232Cys	TM6	Locates right in the binding pocket zone	Mutated smaller side chain may result in loose packing and disrupt the substrate interaction
Asp233Glu	TM6 (binding pocket)	Play a crucial dual role as part of the binding pocket	A slightly larger residue may affect the tight packing pattern of the binding pocket
Ser379Arg	TM10	In the middle of TM10 and near the binding pocket	Mutated charge and/or larger side chain may change the size of the binding pocket and disrupt substrate interaction
Ala382Thr	TM10	In the middle of TM10 and near the binding pocket	Mutated polar and/or larger side chain may change the size of the binding pocket and disrupt the substrate interaction
rBAT			
Thr216Met	Domain B	Adjacency of Asn214 and forms hydrogen bond with Tyr237	Loss of hydrogen bond and mutated longer residue may disrupt the stabilization of the ion binding site
Arg270X Arg270Leu	Domain B	Forms hydrogen bond network along with Thr234, Gln272 and Asp241	Arg270X: Truncated ectodomain Arg270Leu: Loss of hydrogen bond network

Arg362Cys Arg362His	Domain B	Forms charge interaction with Asp359 from another protomer	Arg362Cys: Loss of charge interaction and may destabilize the dimerization Arg362His: Shorter residue leads to weaker charge interaction and may destabilize the dimerization
Arg365Trp Arg365Gln Arg365Pro Arg365Leu	TM6	TIM α4 in domain A	
Met467Thr Met467Lys	At the interface between domain A and C	Forms hydrophobic interaction network as described in main text	Met467Thr/Lys: Mutated polar and charge side chain may disrupt hydrophobic interactions network
Cys666Trp	C-terminus loop	Forms disulfide bond with Cys571	Loss of disulfide bond and may destabilize the c-terminus which also participates in the interaction with $b^{(0,+)}$AT1
Cys673Trp Cys673Arg	C-terminus loop	Forms disulfide bond with Cys685	Loss of disulfide bond and may destabilize the c-terminus which also participates in the interaction with $b^{(0,+)}$AT1

Movie S1 (separate file). Principle components of multibody analyses (part I)

Repositioning of the reconstructed body densities along three major motion for the heterotetrameric complex reveals the flexibility of the TMH domain over the extracellular domain. Soft masks for extracellular domain and TMH domain were applied for conducting this multibody refinement.

Movie S2 (separate file). Principle components of multibody analyses (part II)

Repositioning of the reconstructed body densities along three major motion for the heterotetrameric complex reveals the flexibility between two heterodimeric subunits. Soft masks for heterodimeric bodies were applied for conducting this multibody refinement.