Uncoupled substrate binding underlies the evolutionary switch between Na^+ and H^+-coupled prokaryotic aspartate transporters

Krishna D. Reddy[#1], Burha Rasool[1], Farideh Badichi Akher[1], Nemanja Kutlešić[1], Swati Pant[2], Olga Boudker[#1,3]

[1] Dept. of Physiology & Biophysics, Weill Cornell Medical College, 1300 York Ave, New York, NY 10021, USA

[2] Dept. of Biochemistry, Weill Cornell Medical College, 1300 York Ave, New York, NY 10021, USA

[3] Howard Hughes Medical Institute, Weill Cornell Medical College, 1300 York Ave, New York, NY 10021, USA

[#] Correspondence: kdr400219@gmail.com; olb2003@med.cornell.edu

Abstract

Secondary active membrane transporters use the electrochemical energy of ion gradients to concentrate their substrates. Transporters within the same family often evolve to use different ions, driven by physiological needs or bioavailability. How such functional differences arise despite similar three-dimensional protein structures is mostly unknown. We used phylogenetics and ancestral sequence reconstruction on prokaryotic glutamate transporters to recapitulate the evolutionary transition from Na^+-coupled (GT-Na) to H^+-coupled (GT-H) transport and discovered that it occurred via an intermediate clade (GT-Int). The reconstructed ancestral transporter AncGT-Int contains all residues required for Na^+ binding but switched from Na^+-coupled substrate binding, characteristic of GT-Na transporters, to uncoupled binding. The high-resolution cryo-EM structures of AncGT-Int show that it binds substrate without Na^+ ions in the same manner as GT-Na transporters, which instead require Na^+ ions. Unlike GT-Na transporters, remodeled by ions into a high-affinity substrate-binding configuration, apo AncGT-Int is already in this configuration. Our results show how allosteric changes eliminated Na^+ dependence of Na^+-coupled transporters before H^+ dependence arose, shedding light on ion coupling mechanisms and evolution in this family, and highlighting the power of phylogenetics and ancestral sequence reconstruction in the structure-function studies of membrane transporters.

Introduction

Transport of solutes against their concentration gradients is critical to maintaining cellular homeostasis. Membrane transporters catalyze the energetically unfavorable solute movement using the energy of ATP or the flow of ions down their electrochemical gradients; these mechanisms are termed primary and secondary active transport. Secondary active transporters evolved to use different ions depending on cellular environment and needs; this is particularly evident with proton (*pmf*) or sodium motive forces (*smf*). In contexts where membranes are H^+-leaky or Na^+ is abundant, cellular transporters prefer *smf*. Conversely, when Na^+ is scarce, and lipid membranes are tight, the transporters may instead use *pmf* (1,2). Accordingly, despite having a conserved fold, homologous transporters within the same family can couple the uptake of the substrate to Na^+ or H^+ gradients (3–6). Understanding the basis of such protein adaptability could inform how organisms and cells evolve to maintain homeostasis in the face of dramatically different environments and reveal the fundamental mechanistic principles by which transporters couple substrate uptake to diverse ion gradients.

Characterized prokaryotic homologs of the excitatory amino acid transporter (EAAT, also termed SLC1) family use exclusively Na^+ or H^+ gradients to drive concentrative uptake of amino acids or dicarboxylates (**Figure 1a**) (7,8). Initial understanding of their structure and mechanism arose from prototypical Na^+-coupled transporters (GT-Na), archaeal aspartate transporters Glt_{Ph} and Glt_{Tk} from *Pyrococcus horikoshii* and *Thermococcus kodakarensis*, respectively (9–11). These elevator-type transporters are homotrimers of independently functioning protomers, each comprised of a stationary scaffold and dynamic transport domain, moving between extracellular and intracellular positions (outward- and inward-facing states, OFS and IFS, **Supplementary Figure 1**) (12–17). Transported substrates and three coupled Na^+ bind within the transport domain, gated by a helix-turn-helix motif, termed helical hairpin 2 (HP2) (15,16,18–21). This overall mechanism is conserved in mammalian EAATs and related neutral amino acid transporters (18,19,22,23). Characterized H^+-coupled homologs (GT-H) include *Bacillus caldotenax* $GltT_{Bc}$, *Geobacillus stearothermophilus* $GltT_{Bs}$, *E. coli* $GltP_{Ec}$, and *Bacillus subtilis* $DctA_{Bs}$ (8,24,25). Though detailed structures and mechanisms of GT-H transporters have not been described, they show up to 35% sequence identity with GT-Na, homotrimeric architecture, and rigid scaffold domains during transport (26,27), suggesting that their overall transport mechanism is the same.

Na^+- and H^+-coupling mechanisms of GT-Na and GT-H are highly ion-selective when transporters are reconstituted into proteoliposomes. GT-Na shows no observable transport without Na^+, even if the H^+ gradient is present; conversely, GT-H does not transport in the presence of Na^+ gradients alone (7,8). How did these transporters evolve such different and strict ion coupling while relying on conserved architecture and mechanisms?

In GT-Na transporters, Na^+-coupling is linked to substrate binding. A higher Na^+ concentration on the extracellular side of the membrane ensures the transporter's high affinity for the substrate and a high binding probability, while in the cytoplasm, the affinity is lower and the substrate dissociation probability is higher (10,21,28,29). This Na^+-dependent substrate binding and dissociation provide directionality to the transport cycle, leading to concentrative uptake. The mechanism of H^+-coupling in GT-H is yet unknown. Sequence comparison and mutagenesis suggest that GT-H transporters have perturbed Na^+ sites (**Figure 1b-c**). In GT-Na transporters, Na^+ binding to Na1 and Na2 sites precedes substrate binding (19–21,30–32). In the last step before transport domain translocation, Na2 'locks' HP2 closed over the bound substrate (28,31,33–35). GT-H transporters have conserved substitutions in all three sites: T92[A/L/I/V/M], D405N, and M311L, respectively (**Figure 1b-c**). Introducing these and similar mutations in Glt_{Ph} can reduce Na^+ binding and diminish coupling in some cases (10,36–38). Though the proton carrier residues of GT-H are unknown, we previously suggested that a glutamate residue at the Glt_{Ph} S279 position at the tip of HP1, conserved in GT-H, is a candidate (36). These sequence changes do not explain how a high-affinity substrate-binding site forms in a Na^+-independent manner in GT-H and whether substrate binding to GT-H is H^+-dependent or H^+-coupling occurs through a different mechanism. Thus, the evolutionary mechanism of the *smf* and *pmf* switch remains unknown.

Phylogenetic analysis is a powerful approach to untangling evolutionary functional switches and identifying the linchpin sequence and structural features underlying protein mechanisms. The analysis can reveal

minimal subsets of stepwise, chronologically ordered changes required to alter functions, including the epistatic changes, which themselves are insufficient for the functional switch but are prerequisite for it (39). Subsequent ancestral sequence reconstruction (ASR) infers sequences along an evolutionary pathway, providing a platform to test evolutionary and mechanistic hypotheses (40). Previous structure-function ASR studies revealed the basis of drug specificity in kinases, fluorescence in coral proteins, hormone receptor ligand specificity, ribonuclease H folding, and thermostability of elongation factors (41–46).

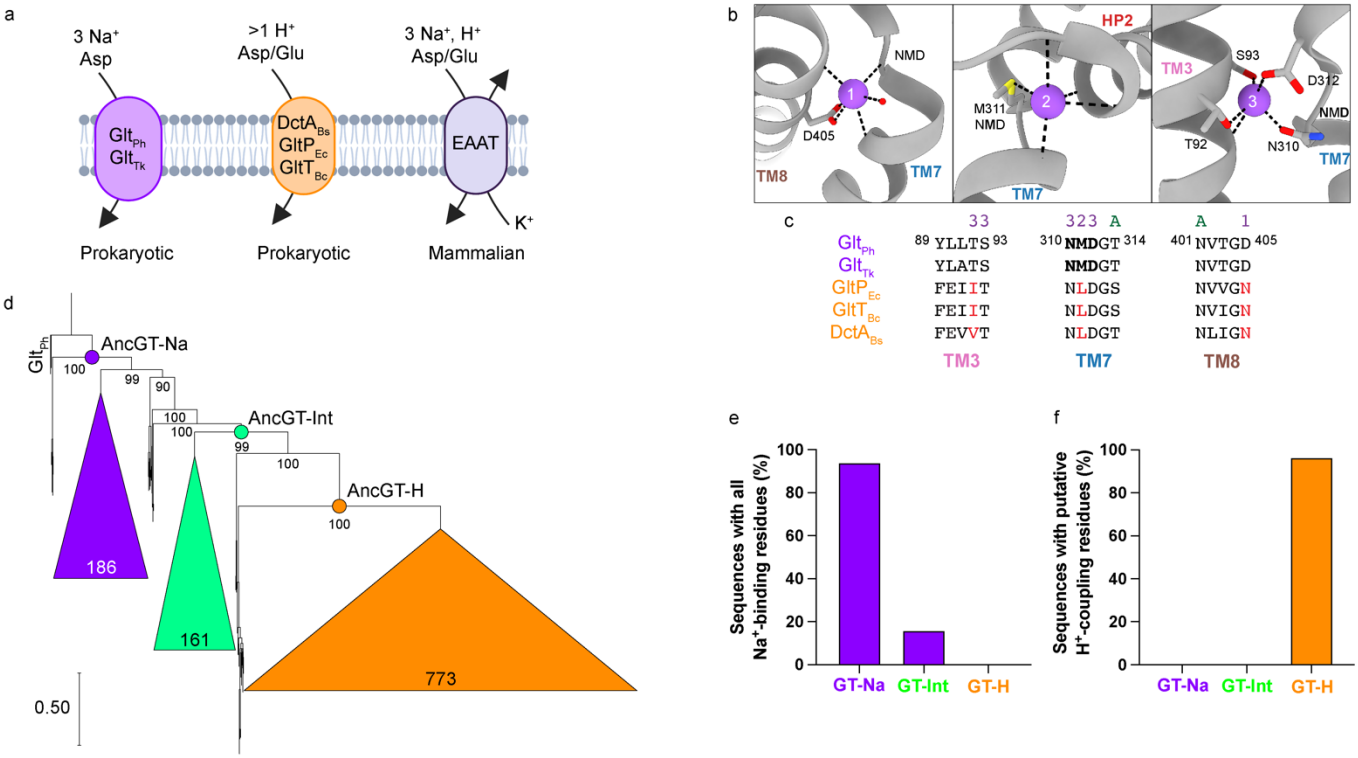

Figure 1: Evolution of ion-coupled glutamate transporters. (a) Ion-coupling stoichiometry of characterized concentrative transporters in the glutamate transporter family. (b) The architecture of Na^+-binding sites in Glt_{Ph} (PDB 6X15). Bound Na1-3 are shown as purple spheres. Dotted lines emphasize the coordinating interactions with sidechains, shown as sticks, and main chain carbonyl oxygens, shown as cartoons. (c) Sequence alignment of experimentally characterized prokaryotic homologs. The numbers above the alignment refer to the Na sites, and 'A' refers to sidechains that coordinate the substrate L-Asp. Sequence changes associated with *smf* to *pmf* transitions are shown in red. (d) Reduced maximum-likelihood phylogeny of glutamate transporters in *Firmicutes*. The collapsed GT-Na (purple), GT-H (orange), and GT-Int (green) clades are proportional to the number of the comprising sequences, which is shown within the clade. Numbers at nodes represent corresponding UFBoot2 values. Nodes corresponding to the reconstructed ancestral sequences are highlighted in the corresponding color and labeled AncGT-Na, -Int, and -H. (e) Percentage of sequences in each clade containing all known Na^+-binding site residues ([S/T]92-93, N310, [M/S]311, D312, D405). (f) Percentage of sequences in each clade containing putative signature H^+-coupling residues ([A/L/I/V/M]92, L311, N405).

Here, we implement ASR to uncover the molecular basis of ion coupling in membrane transporters for the first time. Unexpectedly, our phylogenetic tree (**Figure 1d**) suggests that the transition between GT-Na and GT-H transporter occurred via an evolutionary intermediate with unique features, GT-Int. To understand the mechanistic implications of this transition, we expressed, purified, and biochemically and structurally characterized the inferred ancestors - a first for membrane proteins. As expected, the inferred Na^+-coupled ancestor showed Na^+-dependent substrate binding and transport. In contrast, the inferred intermediate, AncGT-Int, featured substrate binding uncoupled from Na^+ ions, even though the amino acid changes between the ancestors do not alter the first coordination shell of the Na^+ sites. High-resolution cryo-EM structures of AncGT-Int show that it binds substrate in a manner identical to GT-Na and its Na^+-binding sites show conserved architecture, though they lack bound ions. Furthermore, even in the absence of Na^+ ions, apo AncGT-Int already features the high-affinity configuration of the substrate-binding site, which in GT-Na transporters requires Na^+ binding. We propose that the GT-Na to GT-Int transition is a prerequisite for the evolution of H^+-coupling and involves changes in the energy landscape of the transporter, eliminating the need to for Na^+ ions to reach the substrate-binding state. Our results suggest that allosteric regulatory mechanisms can control ion coupling in glutamate transporters, and, by analogy in other secondary active transporters. We describe a novel evolutionary pathway that allows diversification of ion coupling while maintaining strict ion specificity of individual clades. Our work showcases the power of the structure-function ASR in untangling the complex mechanisms of membrane transporters.

Results

An evolutionary intermediate between Na^+- and H^+-coupled transporters

Biochemically characterized GT-H proteins are present in *E. coli* from the *Proteobacteria* phylum and *Geobacillus stearothermophilus* and *Bacillus subtilis* from the *Bacillota/Firmicutes* phylum. Putative GT-H genes with signature [T/S]92[A/L/I/V/M], D405N, and M311L substitutions compared to GT-Na genes are widespread throughout bacterial phylogeny. We first constructed a maximum parsimony tree of all bacterial glutamate transporter homologs and found that putative GT-H transporters cluster together regardless of phylum (**Supplementary Figure 2a**). Though we cannot exclude the existence of H^+-coupled transporters with completely different signature residues, we infer that the emergence of GT-H was a single gene duplication event, following which GT-H genes spread through horizontal gene transfer across bacteria.

We next selected sequences from *Firmicutes/Bacillota* with amino acid signatures of GT-H transporters and proximal sequences up until those with signature GT-Na residues (**Supplementary Figure 2a**). After clustering, alignment, and manual adjustment (**Supplementary Figure 3** and Methods), we constructed a maximum-likelihood (ML) tree. An ML tree optimizes the tree topology, branch lengths, and model of protein evolution in order to create the most probable tree needed to observe the given sequence data (47–49) (**Supplementary Figure 3**). We chose the outgroup-based approach of rooting the tree, using archaeal Glt_{Ph} as the root due to the evolutionary distance between bacteria and archaea (**Supplementary Figure 2a**). However, due to the difficulties of reconciling microbial gene trees to species trees, our rooting is subjective and might not reflect the true chronology of evolutionary history (50) (**Supplementary Figure 2b** and Discussion).

Our ML tree, presumably encapsulating the transition from Na^+ to H^+ coupling, clustered into three major clades (**Figure 1d**). Predictably, two of the clades displayed conserved signature sequences of Na^+ and H^+ coupling (**Supplementary Figure 4**). Surprisingly, we found an intermediate clade GT-Int along the evolutionary trajectory from GT-Na to GT-H (**Fig. 1d**) occurring with a UFBoot2 bootstrap value of 99, indicating high confidence. Within the GT-Na and GT-H clades, known Na^+-coordinating residues and signature residues of H^+-coupled transporters are conserved, with 94% and 97% of sequences containing all these residues, respectively (**Figure 1e-f**). In contrast, only 16% of GT-Int sequences contain all Na^+-coordinating residues, though many sequences have several of these residues (**Figure 1e**, **Supplementary Figure 4**). No GT-Int sequences contained all H^+-coupling signature residues (**Figure 1f**).

To further understand this evolutionary transition, we reconstructed ancestral sequences. The process considers the optimized ML parameters to determine the posterior probability (PP) of each amino acid occurring at each position (site) in the multiple sequence alignment. The final ML amino acid sequence for each node is constructed using the most probable amino acid at every site. We examined ancestral sequences corresponding to the closest ancestor to the root (AncGT-Na), the ancestor of GT-Int (AncGT-Int), and the first ancestor with robustly reconstructed GT-H signature sequences (AncGT-H) (**Figure 1d**). The ancestral sequences were reconstructed with an average PP per site ranging from 0.82-0.87 (**Supplementary Figure 5d**), which is sufficiently high for ASR studies (51,52). As expected, reconstructed AncGT-Na and AncGT-H sequences showed amino acids comprising Na^+ sites typical of the characterized GT-Na and GT-H transporters, respectively (**Supplementary Figure 5c**). Surprisingly, AncGT-Int featured the full complement of Na^+ site residues (**Figure 2a**, **Supplementary Figure 5c**), in contrast to the poor conservation of these sites in GT-Int extant transporters. As this is the last ancestor before the transition to H^+-coupling, we sought to characterize its functional properties.

Design of ancestral transporters for biochemical analysis

We next designed ancestral constructs for protein expression and functional/structural characterization. Since poorly aligned loops and tails are excluded from ASR analysis, we designed 'Frankenstein' constructs (**Supplementary Figure 6**), where we inserted loops and tails from an extant member of the phylogeny with the highest sequence identity to AncGT-Int. Most of these were single residue insertions, except for the C-terminus and the loop between TMs 3 and 4 (3L4). These highly variable regions are distant from the substrate- and ion-binding sites and are unlikely to affect coupling. Indeed, the C-terminus

has been deleted in Glt_{Ph} for structural and functional studies (9), and cleavage of 3L4 affected transport kinetics but not ion-coupling (53). For consistency, we used the same loops and tails for all subsequent ancestral constructs. We expressed ancestors in *E. coli* either with or without a C-terminal GFP optimized for membrane proteins (54), followed by Twin-Strep and 10x-His affinity purification tags. Despite extensive optimization of loop/tail sequences and expression/purification conditions, the ancestral sequences showed low expression yields of ~20-100 µg of purified proteins per liter of *E. coli* culture, compared to ~10-fold higher yields of extant bacterial and archaeal homologs. Nevertheless, we purified them using Streptactin 4Flow XT resin and size-exclusion chromatography to near homogeneity (**Figure 2b,c**).

Uncoupled substrate and cation binding to the evolutionary intermediate

We first expressed, purified, and reconstituted AncGT-Na into proteoliposomes, expecting it to function as a Na^+- and aspartate/glutamate symporter because it is 79% identical to Glt_{Ph} in the transport domain (**Supplementary Figures 5a and 6**). Indeed, in the presence of a negative membrane potential and an inward Na^+-gradient, we observed a time-dependent accumulation of ^{3}H-L-Asp in proteoliposomes, while no ^{3}H-L-Asp was retained in the absence of Na^+ (**Figure 2d**). These results suggest that AncGT-Na is properly folded and recapitulates Na^+-coupling of extant GT-Na transporters. When we repeated the experiment with AncGT-Int, we observed lower ^{3}H-Asp retention by the liposomes; however, this was unaffected by removal of Na^+, and this signal could be further reduced in the presence of competitive inhibitor DL-TBOA (**Figure 2d**). As we did not observe any ^{3}H-L-Asp retention in exchange conditions, with 500 µM unlabeled L-Asp inside the vesicles and 1 µM ^{3}H-L-Asp outside (not shown), we suggest that AncGT-Int in our proteoliposomes can bind but not transport L-Asp independently of Na^+, in a manner that can be inhibited by DL-TBOA.

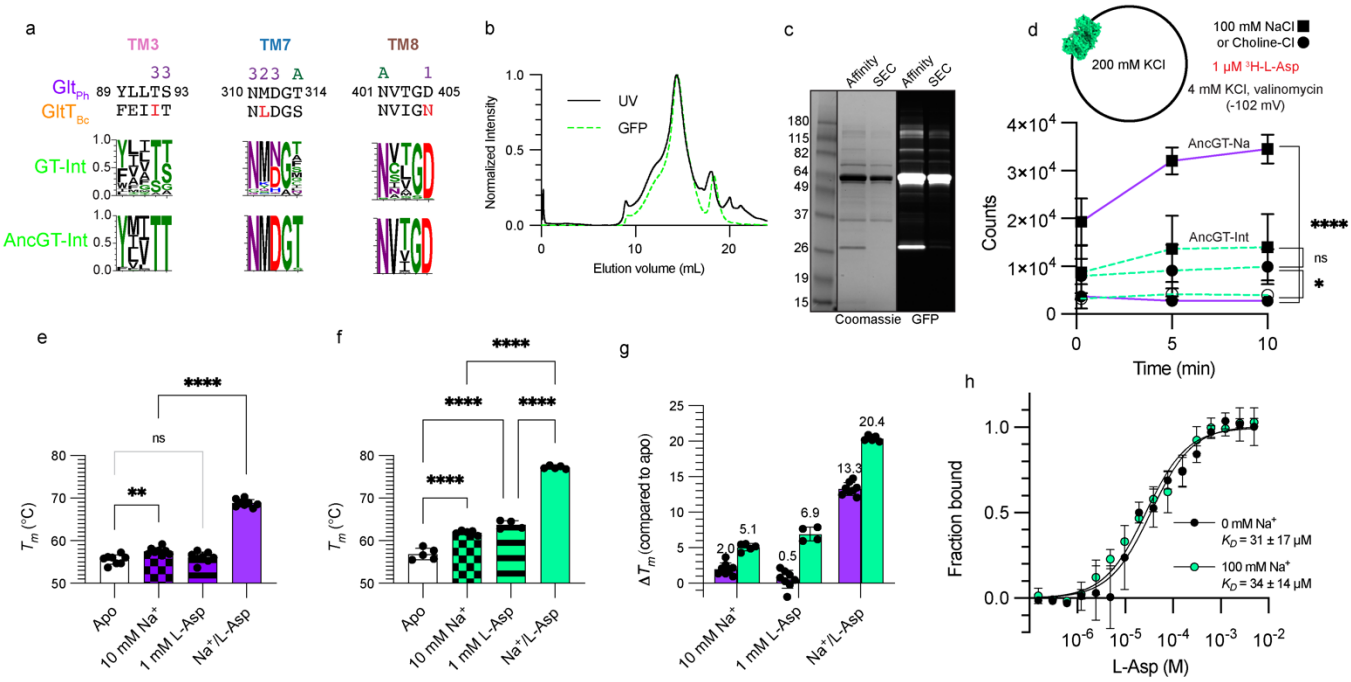

Figure 2: Characterization of a cation-independent ancestral transporter. (a) Comparison of the signature Na^+-coupling residues in GT-Int extant transporters and ancestral AncGT-Int. GT-Int WebLogo represents alignment of extant transporters in GT-Int clade. AncGT-Int WebLogo is a visual representation of the posterior probability per site in the ancestral reconstruction; the top-most residue at each site represents the 'maximum-likelihood' residue. (b) Representative size-exclusion chromatography profiles of GFP-tagged AncGT-Int purified in apo conditions (50 mM KH_2PO_4/K_2HPO_4, 1 mM DDM). Black and dotted green lines correspond to relative signals at 280 nm absorption and 485/510 nm fluorescence, respectively. (c) Representative gels of GFP-tagged AncGT-Int, stained with Coomassie (left) or imaged by in-gel GFP fluorescence (right). "Affinity" refers to pooled and concentrated elution fractions following streptactin affinity purification. "SEC" refers to pooled and concentrated peak fractions eluting at 14.5 mL following size-exclusion chromatography. (d) ^{3}H-L-Asp uptake in the presence (filled squares) or absence (filled circles) of an Na^+ gradient. A simplified proteoliposome reaction setup is above the graph (see Methods). All experiments were performed in the presence of a negative membrane voltage, generated by valinomycin-mediated K^+ efflux. Lines represent AncGT-Na (solid purple) or AncGT-Int (dashed green). Open circles represent AncGT-Int in the absence of an Na^+ gradient and in the presence of 1 mM DL-TBOA. Statistical significance was measured with unpaired t-tests. Results were averaged from three independent experiments. (e-f) Melting temperatures (T_m) of AncGT-Na (e) or AncGT-Int (f) in buffer alone (Apo), 10 mM NaCl, 1 mM L-Asp, or both NaCl and L-Asp. Replicates are from at least two independent experiments; statistical significance was measured with one-way ANOVAs. (g) Change in T_m compared to apo protein for AncGT-Na (purple) or AncGT-Int (green). (h) L-Asp binding by AncGT-Int in 100 mM NaCl (green circles) or 100 mM choline chloride (black circles). K_D in NaCl and choline chloride were 31 ± 17 µM and 34 ± 14 µM, respectively. Results are from three independent experiments.

To further probe the binding properties of AncGT-Na and AncGT-Int, we used nano-differential scanning fluorimetry (nanoDSF) to measure ligand-dependent protein melting temperatures (T_m) (55,56). In AncGT-Na, we did not observe a significant T_m increase when adding 1mM L-Asp compared to the apo transporter. Also, adding 10 mM NaCl alone to apo AncGT-Na resulted in only 2.0ºC ± 0.9 T_m increase. In contrast, we observed a dramatically improved stability of 13.3ºC ± 0.9 when we added L-Asp and NaCl together (**Figure 2e,g; Supplementary Figure 6a**). Thus, L-Asp and Na^+ only bind tightly together, suggesting that they bind cooperatively, as has been reported for Glt_{Ph} and Glt_{Tk} (7,10,21,28,29,57,58). In contrast, we observed that adding L-Asp and NaCl separately to AncGT-Int increased T_m by 6.9ºC ± 1.0 and 5.1ºC ± 0.5, respectively, showing that each can bind to AncGT-Int without the other (**Figure 2f,g; Supplementary Figure 7a**). We next used microscale thermophoresis (MST) to measure substrate binding to AncGT-Int and observed that AncGT-Int binds aspartate with similar micromolar affinities in the presence and absence of 100 mM Na^+ (**Figure 2h**). Notably, the MST signal-to-noise values upon L-Asp binding ranged from 8.6-19.5, reflecting the high confidence of the obtained binding isotherms (**Supplementary Figure 7b**). Together, these results show that aspartate binding to AncGT-Int ancestral protein is independent of Na^+ binding and suggest that the GT-Int clade and its ancestors functionally diverged from GT-Na by uncoupling substrate from Na^+-binding.

All amino acid sequence changes between AncGT-Na and AncGT-Int lie outside the ion-binding sites, suggesting that allosteric mechanisms underlie uncoupling of Na^+ and L-Asp binding in AncGT-Int and GT-Int clade. It is common to use sequence comparison to deduce the amino acid changes responsible for the functional change. However, if the changes occur outside of a small number of residues in the ligand binding sites, this approach is hampered by many sequence differences between the evolutionarily distant proteins irrelevant to their functional differences. For example, the extant Na^+-coupled Glt_{Ph} and H^+-coupled $GltT_{Bc}$ share only 37% sequence identity, corresponding to 140 differences in the 222 amino acid transport domain (**Supplementary Figure 5a**). By contrast, AncGT-Na and AncGT-Int are 78% identical, and AncGT-Int and AncGT-H are 70% identical. Thus, only 26 and 40 of amino acids with PPs above 0.9 in at least one of the ML sequences differ between AncGT-Na/AncGT-Int and AncGT-Int/AncGT-H transport domains, respectively (**Supplementary Figure 5b**). Therefore, using the ASR approach significantly reduces the number of amino acid changes potentially underlying altered ion coupling. Notably, we did not find sequence changes in AncGT-Na and AncGT-Int that could be immediately recognized as responsible for uncoupling aspartate and Na^+ binding in AncGT-Int. Instead, the changes were mostly subtle, motivating us to use structural analyses to reveal the basis of ion/substrate uncoupling.

The high-affinity substrate binding site forms in the absence of cations

We hypothesized that uncoupling between Na^+ and aspartate binding in AncGT-Int occurs because Na^+ is no longer necessary to form the high-affinity substrate binding site. We purified AncGT-Int, reconstituted it into MSP1E3 lipid nanodiscs (**Supplementary Figure 8**), exchanged it into a buffer containing no alkali cations (20 mM HEPES pH 7.4, 100 mM NMDG-Cl), and froze cryo-EM grids either before or after adding 1 mM L-Asp, i.e., apo and bound conditions, respectively. For AncGT-Int, we observed three outward-facing states without substrate, which we termed Apo_{OFS}/NMD-out, Apo_{OFS}/NMD-in, and Apo_{OFS}/NMD-in-2; all refined to 3.0 Å resolution (**Supplementary Figure 9, 12-13, Supplementary Table 1**). We also observed a single outward-facing state in the presence of substrate, Asp_{OFS}, which we refined to 3.4 Å overall resolution (**Supplementary Figure 10, 12-13, Supplementary Table 1**). In Apo_{OFS}/NMD-out, the NMD motif in TM7 was flipped out, as previously observed in apo Glt_{Ph} and EAATs (**Figure 3c**). In both Apo_{OFS}/NMD-in states, NMD motifs were flipped in, as typically observed in Na^+- and Na^+/substrate-bound archaeal and eukaryotic transporters (**Figure 3c,e-f**). We also found inward-facing structural classes in both apo and substrate-bound AncGT-Int (**Supplementary Figure 9-10, 13**). However, despite extensive processing, the cytoplasmic halves of the transport domains remained poorly resolved, suggesting significant structural heterogeneity (**Supplementary Figure 13**); we did not pursue these further.

As a control, we reconstituted Glt_{Ph} into MSP1E3 lipid nanodiscs and froze grids under apo conditions, i.e., without L-Asp and alkali ions. Cryo-EM imaging revealed the presence of an Apo_{OFS} and an Apo_{iOFS} with final maps at 2.7 and 3.1 Å resolution (**Supplementary Figure 11-13, Supplementary Table 1**). iOFS is an intermediate state in which the transport domain moves inward toward IFS by about a quarter of the

271 way that was observed in both archaeal and human homologs (16,32,59,60). As expected, all resolved
272 maps showed the flipped-out NMD motifs (**Figure 3d, Supplementary Figure 11**).
273

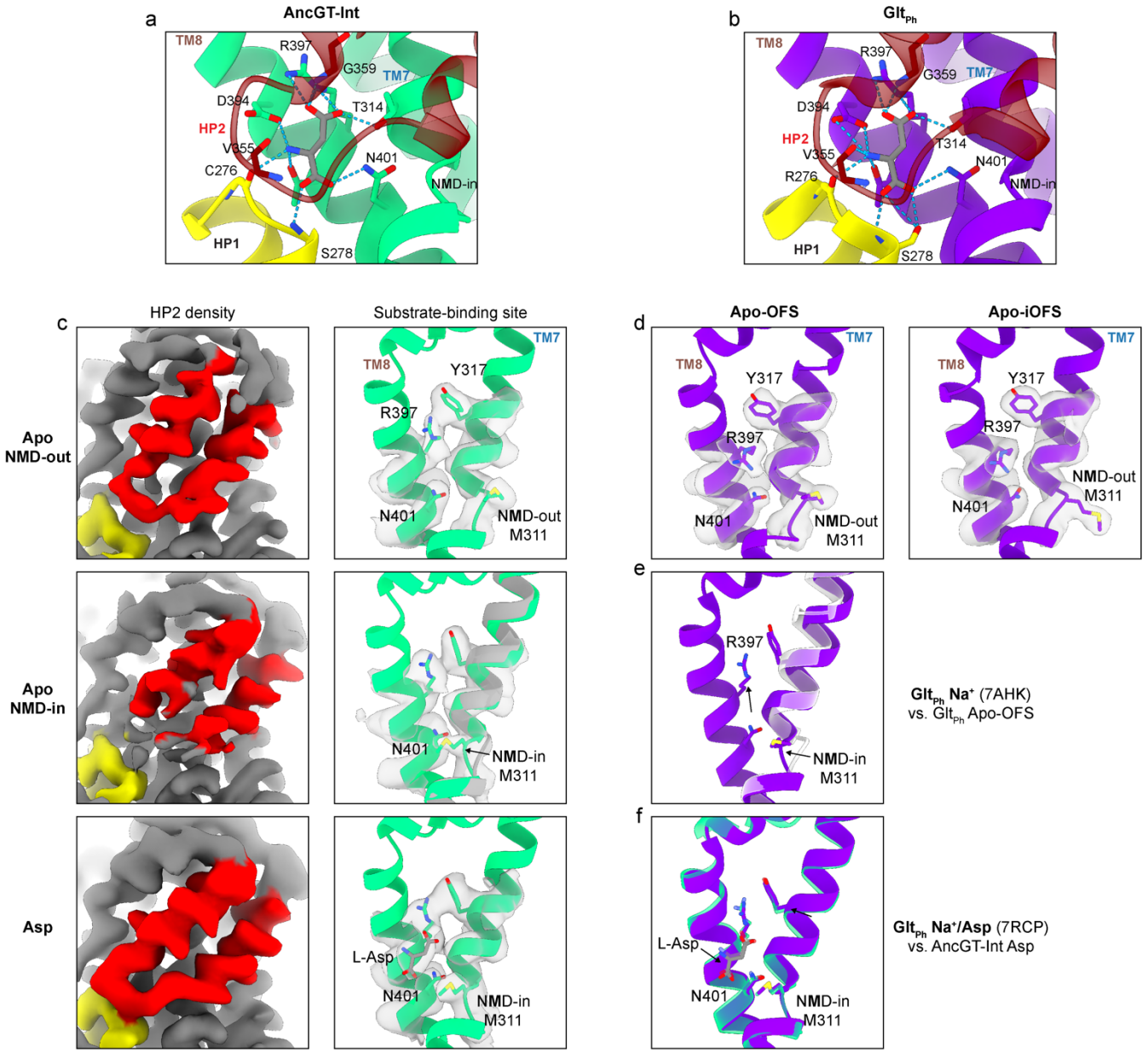

Figure 3: Substrate-binding architecture and rearrangements in AncGT-Int, compared to Glt$_{Ph}$. All numbered residues use Glt$_{Ph}$ numbering. All structures are aligned to each other, using the rigid HP1/TM7a region (259-309 in Glt$_{Ph}$) as a template. All densities are contoured at 10σ. (a-b) Coordination of L-Asp in AncGT-Int (a) and Glt$_{Ph}$ (b), determined using default parameters of 'hbonds' in ChimeraX. (c) Cryo-EM densities of HP2 (left) and the substrate binding site (rigt) in AncGT-Int Apo$_{OFS}$/NMD-out (top), Apo$_{OFS}$/NMD-in (middle) and Asp$_{OFS}$ (bottom). HP2 (red) and HP1 (yellow) densities are unsharpened. Substrate-binding densities are displayed using 'volume zone' on a sharpened map. In Apo$_{OFS}$/NMD-in, the molecular model is Apo$_{OFS}$/NMD-out (gray) demonstrates the difference in the NMD loop. (d) Apo Glt$_{Ph}$ in Apo$_{OFS}$ (left) and Apo$_{iOFS}$ (right) conformations; densities are sharpened. (e) Comparison of Glt$_{Ph}$ Apo$_{OFS}$ structure (translucent) to OFS Na$^+$-bound Glt$_{Ph}$ (purple; 7AHK) demonstrates the restructuring of the NMD loop. (f) Comparison of AncGT-Int Asp$_{OFS}$ (green, translucent) to Na$^+$/L-Asp bound Glt$_{Ph}$ (purple; 7RCP).

274
275 In AncGT-Int Asp$_{OFS}$, we observed well-structured closed HP2 and the substrate binding site architecture
276 remarkably like Na$^+$/Asp-bound Glt$_{Ph}$ (**Figure 3a-b, Figure 3f**). A strong non-protein density in the site,
277 which we modeled as L-Asp, suggested that AncGT-Int binds the amino acid without alkali cations in the
278 same rotameric state as Glt$_{Ph}$ in the presence of Na$^+$ (**Figure 3c**). Side chains of conserved T314$_{292}$,
279 D394$_{366}$, R397$_{369}$, and N401$_{373}$ (hereafter, normal text corresponds to Glt$_{Ph}$ numbering, subscripts
280 represent AncGT-Int numbering) coordinate L-Asp (**Figure 3a**). R397 is the principal residue interacting
281 with the side chain carboxylate of L-Asp, and D394$_{366}$ coordinates the amino group; these residues
282 determine the specificity of the transporter for the acidic amino acids over neutral amino acids and

dicarboxylates, respectively (61–63). The amide nitrogen of $N401_{373}$ coordinates the main chain carboxylate of L-Asp, and $N401_{373}$ is one of the most conserved amino acids throughout the evolution of glutamate transporters. The M311 residue points toward L-Asp, forming van der Waals interactions.

In AncGT-Int Apo_{OFS}/NMD-out, HP2 is open with M311 lodging between the NMD motif and the hairpin; the binding pocket is empty and hydrated (**Figure 3c**). In Apo_{OFS}/NMD-in-2, HP2 is closed, and we observe excess non-proteinous density in the pocket (**Supplementary Figure 14**). The density is too large for L-Asp and might correspond to a spuriously bound buffer component; it has been suggested that this site is generally attractive to anions (20). Consistently, the overall structure resembled Asp_{OFS}, except HP2 is distorted around the region where the Na2 site would be. HP2 showed similar overall conformation in Apo_{OFS}/NMD-in, but its tip was disordered, and the substrate-binding pocket was empty and hydrated, while M311 was in the correct position for substrate binding (**Figure 3c**).

Several structural rearrangements in Na^+-coupled transporters convert the low-affinity substrate-binding site in apo transporters into the high-affinity site upon Na^+ binding (10,20,21,31). First, the non-helical NMD loop motif, which breaks TM7 into TM7a and TM7b, restructures from a 'flipped-out' to a 'flipped-in' configuration, repositioning the $N401_{373}$ side chain (**Figure 3e**), whose amide oxygen hydrogen-bonds to the main chain nitrogens at the start of TM7b (**Figure 3b**). The flip also brings M311 proximal to the substrate. Interestingly, mutating it to alanine, but not leucine, in Glt_{Ph} reduces substrate affinity and Na^+ coupling (31,33,34). Concomitantly, $R397_{369}$ in TM8 moves from occupying the substrate-binding site to form a cation-π interaction with $Y317_{295}$ in TM7b (**Figure 3e**), positioning its guanidium group to coordinate the substrate side chain carboxylate. The two structural rearrangements are thought to be allosterically coupled, though the mechanism remains unknown. As expected, we observed the low-affinity configuration in Glt_{Ph} Apo_{OFS} and Apo_{iOFS} with $R397_{369}$ occupying the L-Asp-binding site, and the NMD motif flipped out (**Figure 3d**). The two structures differ in the position of the transport domain and the conformation of HP2, which is open in Apo_{OFS} and closed in Apo_{iOFS} (**Supplementary Figure 13**). M311 either props the open HP2 in Apo_{OFS} or points out into the lipid from under the closed HP2 of Apo_{iOFS}.

In stark contrast, all AncGT-Int apo conformations featured $R397_{369}$ interacting with $Y317_{295}$ and poised to coordinate the substrate L-Asp (**Figure 3c**). In Apo_{OFS}/NMD-in, but not Apo_{OFS}/NMD-out, conformations also display the flipped-in NMD loop and the proper positions of M311 and N401 (**Figure 3c**). Thus, the $R397_{369}$ position in AncGT-Int is uncoupled from ion binding and the NMD motif movements, and the latter can also occur spontaneously without Na^+ ions.

Spontaneous cation-independent rearrangements of ion-binding sites

In Na^+-coupled transporters, the NMD motif flips in upon Na^+ binding to the Na3 site, where it is coordinated by the side chains of N310 and D312 and T92 and [T/S]93 in TM3 (**Figure 4a,b**). Flipped-in NMD positions N310 carbonyl oxygen to coordinate the Na^+ ion in Na1 together with D405 in TM8 and M311 sulfur to coordinate Na2 together with carbonyl oxygens at the ends of TM7a and HP2a. These concerted rearrangements, accompanied by the high-affinity substrate-binding site formation, are thought to underlie the cooperative binding of 3 Na^+ ions and substrate. To examine the structure of unoccupied Na^+-binding sites in AncGT-Int substrate-bound conditions, we took special care when preparing samples for cryo-EM grids. We first removed all alkali cations and substrate by size-exclusion chromatography in alkali- and L-Asp-free buffer; we then added L-Asp back to the apo transporter, preventing the possibility of Na^+ ions carrying over during sample preparation. Strikingly, the architectures of all three Na^+ binding sites of alkali-free AncGT-Int Asp_{OFS} were identical to Na^+/L-Asp bound Glt_{Ph} as seen at the current resolution (**Figure 4a,e**). We also observed identical architectures of Na1 and Na3 sites in the AncGT-Int Apo_{OFS}/NMD-in maps, in which the side chains are well-resolved and configured as would be expected for Na^+-occupied sites (**Figure 4c**). At the Na2 site, although TM7 residues are correctly positioned, HP2 is distorted in Apo_{OFS}/NMD-in maps. Though the former Na^+-binding sites are likely occupied by water molecules in AncGT-Int, the resolution of the maps is insufficient to model waters with confidence. In contrast to AncGT-Int Apo_{OFS}/NMD-in, most of the Na1 and Na3 sites of Apo_{OFS}/NMD-out are more similar to the distorted sites of the apo Glt_{Ph} (**Figure 4c**), with the notable exception of TM3, where T92 and [T/S]93 seem to be already properly positioned.

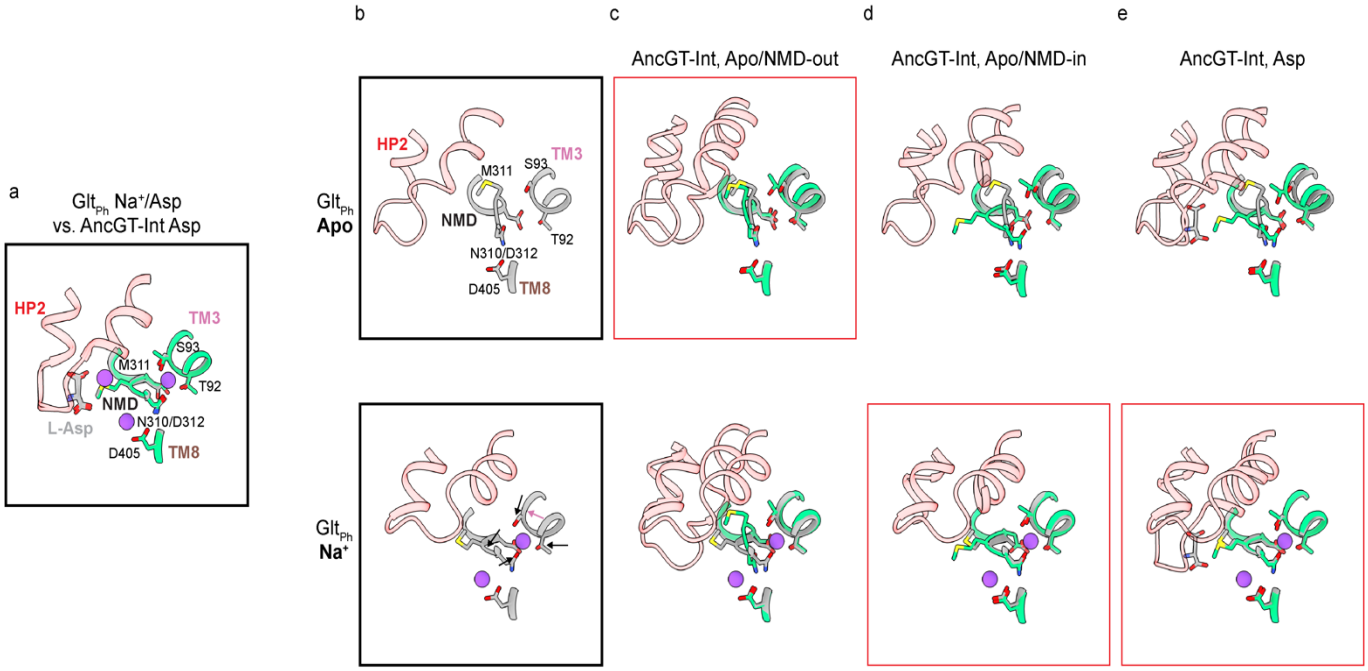

Figure 4. AncGT-Int does not require Na^+ to access high-affinity binding conformations. (a) Similarity of Na^+-binding sites between Glt_{Ph} (gray; PDB 6X15) and AncGT-Int (green) in Na^+/L-Asp and L-Asp conditions, respectively. Na^+ ions in Glt_{Ph} are purple; L-Asp is colored by atom type with carbon in gray. Note that the protein structures are nearly identical. (b) Structural changes upon Na^+ binding in GT-Na transporters. In the Apo_{OFS} state (top), R397 is in the substrate binding pocket (not shown) and M311 points out of the pocket. Upon Na^+-binding (bottom), the NMD loop (containing N310, M311, and D312) and M311 side chain flip into the pocket and TM3 (containing S92, T93) moves closer to the NMD loop to form the Na3 site, which results in Na1 binding and formation of the Na2/L-Asp binding sites. (c-e) Comparison of structural changes in ion-binding sites between Glt_{Ph} and AncGT-Int Apo_{OFS}/NMD-out (c), Apo_{OFS}/NMD-in (d), and Asp_{OFS} (e). Structures were aligned to Glt_{Ph} 225-309 (TM6-TM7a) and 364-425 (TM8), excluding TM3 and NMD loop/TM7b to properly visualize their movements. Red boxes indicate which Glt_{Ph} structure is most similar to a given AncGT-Int structure.

Together with our binding data, the structures suggest that ancestral AncGT-Int can spontaneously adopt high-affinity substrate and Na^+-binding sites configurations without Na^+ ions and bind Na^+ and L-Asp independently of each other. Comparing the overall transport domain structures of Glt_{Ph} and AncGT-Int in low- and high-affinity configurations, we noted that the transition involves not only the local changes around the NMD motif and the substrate-binding site but also the overall repacking of the extracellular half of the transport domain, which was much more prominent in Glt_{Ph} than AncGT-Int. Specifically, TM7b, which starts at the end of the NMD motif and runs through the transport domain center, shows a greater tilt difference between low- and high-affinity states of Glt_{Ph}, which translates to pronounced repacking of all TMs that encircle TM7b, including 3, 6, 8 and HP2 (**Supplementary Figure 15a,b**). Packing differences in this region of the transport domain have been implicated in setting Glt_{Ph} affinity for L-Asp (36) and, therefore, might be allosterically coupled to the low-/high-affinity transition. Strikingly, the region of TM7b at the core of this helical bundle contains one of the family signature sequence motifs – F-[I/V]-A-[N/Q] (64), which contacts all other involved helices. ASR reconstructions showed that AncGT-Na has a near 1 posterior probability of containing $N327_{305}$, while AncGT-Int has a similar probability for Q, suggesting that $N327_{305}Q$ mutation might contribute to the evolutionary transition from Na^+-coupled to independent function. Indeed, $Q327_{305}$ in AncGT-Int hydrogen bonds to the top of TM3 and the conserved PG motif at the start of the 3L4 loop (**Supplementary Figures 6, 15c**), perhaps stabilizing TM7-TM3 interactions and minimizing the packing differences in the low- to high-affinity transition. In contrast, $N327_{305}$ is too short to reach TM3 and only interacts with PG in the loop.

Discussion

In this report, we applied phylogenetics and ASR to gain insights into ion-coupling mechanisms in glutamate transporters. We were puzzled by how related glutamate transporters use either *smf* or *pmf* while relying on shared protein architecture. Sequence comparison of the extant GT-Na and GT-H transporters revealed characteristic signature residues. However, we previously failed to convert GT-Na into GT-H by mutating these residues (36), suggesting the transition requires other unidentified changes. Pinpointing these is challenging because extant sequences typically contain hundreds of differences. To

367 overcome this, we used ASR to reconstruct plausible AncGT-Na and AncGT-H sequences with
368 significantly reduced divergence. To capture the evolutionary transition from GT-Na to GT-H, we selected
369 bacterial sequences that were proximal to each other in the evolutionary tree and covered the juncture
370 between GT-Na and GT-H clades. Strikingly, the evolutionary analysis revealed the existence of an
371 intermediate clade GT-Int between GT-Na and GT-H, which does not contain GT-H signature residues;
372 while overall GT-Na residues are somewhat conserved, most individual proteins within the clade do not
373 have the entire complement of GT-Na residues (**Figure 2a**).
374
375 Biochemical and structural data suggest that GT-Int can bind the substrate without Na^+ ions, hinting they
376 might be cation-independent transporters. Specifically, AncGT-Int showed robust L-Asp binding in the
377 absence of Na^+ ions in four orthogonal experiments, including radiolabeled ligand binding to AncGT-Int-
378 containing proteoliposomes, MST, ligand-induced thermal protein stabilization, and cryo-EM. However, we
379 did not observe substrate transport by AncGT-Int. Notably, cryo-EM imaging showed that AncGT-Int can
380 access OFS and IFS in the presence of L-Asp or under apo conditions, suggesting that it can undergo
381 conformational transitions necessary for transport. It is possible that the dynamics of the transitions are
382 too slow for measurable uptake. Examining alternate ASR proteins might help resolve the question in
383 future studies.
384
385 *Evolutionary remodeling of ion-coupling in prokaryotic glutamate transporters*
386
387 In our phylogeny, Na^+-coupled transporters appeared most ancient, which is plausible given that
388 characterized archaeal transporters are Na^+-coupled (**Supplementary Figure 2a**). If so, our results
389 suggest that uncoupling of ion and substrate binding was an intermediate evolutionary event during the
390 transition from Na^+ to H^+ coupling. While we think it less likely, an alternate root could place uncoupled
391 ancestors as most ancient, diverging into Na^+ and H^+-coupled branches (**Supplementary Figure 2b**). In
392 either scenario, AncGT-Int precedes H^+-coupling. We find it unlikely that H^+-coupled transporters are the
393 hypothetical root, as only one classified archaeon from the newly proposed archaeal phylum
394 *Woesearchaeota* contains a sequence with GT-H signature residues (65).
395
396 Because substrate and ion binding occur within the transport domain, we focused on this domain when
397 considering the origin of ion coupling changes. Most striking to us was the small number of high-probability
398 amino acid differences between AncGT-Na and AncGT-Int, which occur throughout the transport domain
399 and do not include any residues previously implicated in ion binding or coupling (**Supplementary Figure**
400 **5**). AncGT-Int contains the full complement of Na^+ and aspartate binding residues, though they are not
401 highly conserved in the extant GT-Int clade (**Figure 2a**). Lack of conservation suggests the lack of
402 evolutionary pressure on maintaining Na^+ binding sites, consistent with cation-independent transport.
403 Indeed, we observed that AncGT-Int could bind substrate and Na^+ ions independently of each other and
404 could no longer allosterically couple them.
405
406 The hypothesis that GT-Int are no longer Na^+-coupled is consistent with the ecology of the bacteria that
407 contain them (2). GT-Na sequences mostly originate from bacteria living in high-salinity environments. In
408 contrast, both GT-H and GT-Int sequences more often originate from bacteria living in low-salinity
409 environments such as soil and intestinal tracks, suggesting they no longer use Na^+ gradients. While the
410 functional studies of extant GT-Int transporters will need to be carried out, we expect that they do not
411 couple to either Na^+ or H^+ gradients because they lack or show no conservation of the relevant residues.
412 If so, the question arises – what might be their physiological function if they no longer perform ion-driven
413 active transport? One possibility is that these transporters work in concert with other proteins. Examining
414 the gene neighbors of GT-Int transporters in their bacterial genomes revealed that many are adjacent to
415 enzymes involved in amino acid metabolism and amino acid racemases involved in cell wall biosynthesis.
416 Thus, it is possible that GT-Int mediate amino acid uptake coupled to their enzymatic utilization by
417 downstream enzyme(s). Eliminating the need to maintain Na^+ coupling perhaps allowed uncoupling to
418 evolve as a prerequisite to the H^+-coupling evolution. Notably, we observed no such preferences in the
419 neighbors of GT-Na or GT-H transporters, suggesting that they work independently as *bona fide* secondary
420 active transporters.
421

Further evolutionary transition from GT-Int to GT-H might have occurred through increasing the local hydrophobicity around the former Na3 site (**Supplementary Figure 5g**). This is reminiscent of the observation that a subtle increase in the local hydrophobicity leads to the specificity shift from Na^+ to H^+ in the ion-binding sites of prokaryotic F-type rotary ATPases (3,6) and might suggest that the site is remodeled into an H^+-binding site in GT-H. Perhaps, protonation of the conserved Asp in the NMD motif is necessary for L-Asp binding and occlusion. It is also possible that AncGT-Int provided a scaffold for cation-independent substrate binding, and H^+-coupling evolved due to H^+-dependent transport domain translocation, consistent with the appearance of the putative conserved protonation site in HP1 (**Supplementary Figure 5g**). Interestingly, a previous study showed that the uptake by an H^+-coupled $GltT_{Bs}$ could be further stimulated by Na^+ gradients in native *Geobacillus stearothermophilus* membrane vesicles but not in *E. coli* membrane vesicles or proteoliposomes (8,25). Perhaps coupling through the preserved Na3 could change depending on the lipid composition, inducing changes in protein packing or energy balance between the low- and high-affinity states.

Energetics of ion-coupling in secondary active transport

Regardless of the physiological function of the extant GT-Int, reconstructed AncGT-Int, which binds L-Asp in a cation-independent manner, provides insights into the mechanism of allosteric ion coupling. Our most striking structural observation is that apo AncGT-Int features high-affinity substrate-binding site configuration in the absence of Na^+ ions, alleviating the need for ion binding for tight substrate binding (**Figure 5**). Thus, we suggest that the key determinant of coupling is the relative free energies of the low- and high-affinity states. When Na^+ and substrate binding is coupled, as in Na^+-coupled glutamate transporters, the free energy of the low-affinity conformation is low compared to the high-affinity conformation. Thus, the substrate binding energy alone is insufficient to overcome the penalty of the low- to high-affinity conformational transition. In this case, Na^+ binding pays the energetic price, stabilizing the high-affinity state. In contrast, in cation-independent AncGT-Int, the low-affinity state is already energetically unfavorable, and the high-affinity state is predominantly populated in the absence of ions. While Na^+ ions can still bind to the transporter, they do not provide additional benefits and are unnecessary for substrate binding. Thus, the transporter's conformational ensemble is modulated allosterically; such a mechanism is proposed to be a key part of evolutionary diversification of proteins (66–68).

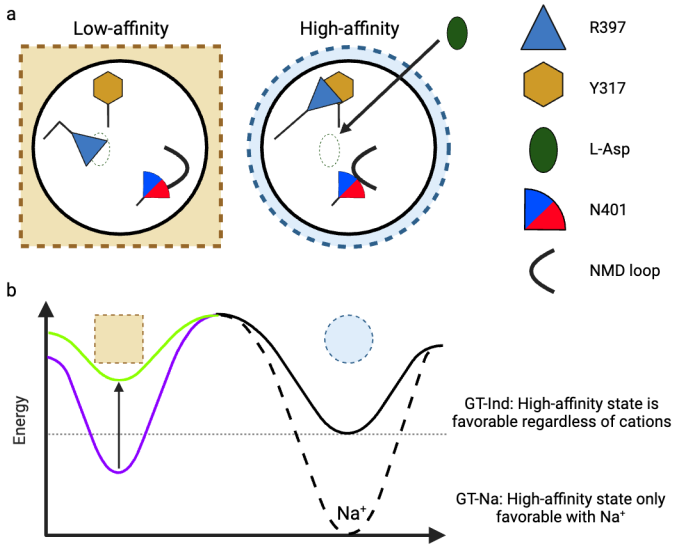

Figure 5. Evolution of energetics in ion-coupled glutamate transporters. (a) The low-affinity (yellow) and high-affinity (blue) binding configurations are visualized using changes in R397, N401, and NMD loop conformations. The dashed lines represent structural changes in the transport domain outside of the binding sites allosterically coupled to the conformations of R397, N401, and NMD loop. (b) The proposed energy landscapes of coupled vs. uncoupled glutamate transporter homologs describing substrate and Na^+ binding. Solid lines represent the relative energies of the low- and high-affinity states in the absence of Na^+; the dashed line represents the changes upon Na^+ binding. Relative to the high-affinity state (black), the energy of the low-affinity states in GT-Na (purple) is lower than in GT-Int (green) transporters. In GT-Na transporters (including Glt_{Ph}, Glt_{Tk}, and AncGT-Na), the high-affinity substrate binding is energetically favored and populated only after Na^+ binding. In GT-Int transporters, the uncoupling mutations, represented by the black arrow, increase the energy of the low-affinity state so that it becomes unfavored even in the absence of Na^+ ions. Created in BioRender.

The immediate structural distinction between the low- and high-affinity states of Glt_{Ph} and AncGT-Int is the position of $R397_{369}$ (**Figures 3 and 5**). In the low-affinity states of Glt_{Ph} – Apo_{OFS} and Apo_{iOFS} – $R397_{369}$ occupies the substrate-binding site. In the high-affinity states of Na^+-bound Glt_{Ph} and apo AncGT-Int, $R397_{369}$ is out of the site poised to coordinate the side chain carboxylate of L-Asp. While this movement is clearly critical for L-Asp binding, it must be a consequence of other rearrangements and not the basis of ion-coupling because $R397_{369}$ is only present in the family members transporting acidic amino acid and dicarboxylates, but not in neutral amino acid transporters. Indeed, R397C switches substrate specificity of Glt_{Ph} to serine/alanine, but does not eliminate Na^+-coupling (63). Instead, we think that the repositioning of the non-helical NMD motif and the structurally adjacent, universally conserved $N401_{373}$ residue, coordinating the main chain carboxylate of the substrate, are at the heart of the low- to high-affinity transition. In the low-affinity state, the non-helical NMD motif, breaking TM7, is flipped out, and in the high-affinity state it flips in (**Figures 3 and 5**). This movement properly positions $N401_{373}$ for substrate binding and leads to the formation of the three Na^+-binding sites. Thus, the high-affinity state in AncGT-Int likely has a high affinity for both the substrate and Na^+ ions. We were surprised to observe in Na^+-free AncGT-Int bound to L-Asp the architectures of the Na1, Na2 and Na3 sites that were identical - within the current resolution of our structures - to Na^+-occupied sites of Glt_{Ph}. Also, in Apo_{OFS}/NMD-in states, Na1 and Na3 sites show proper geometry. Further computational studies will be needed to understand how the structured water molecules and side chain interactions stabilize the nearly perfectly formed Na^+ binding sites. Notably, mutating M311 in Na2 site of GT-Na to leucine found in GT-H does not significantly alter coupling, transport, or HP2 closure upon substrate binding (31,33–35), consistent with the idea that uncoupling form the Na^+ ion in GT-H can occur through mechanisms other than disrupting the site.

The necessary and sufficient mutations explaining why the low- to high-affinity transition is energetically costly in Glt_{Ph} but free or favored in AncGT-Int must be among the 26 high-probability amino acid changes between AncGT-Na and AncGT-Int transport domains. Notably, most of these mutations are subtle, and none stand out as a clear single culprit (**Supplementary Figure 5e,f**). Structural comparisons between the low-affinity apo and high-affinity Na^+-bound Glt_{Ph} structures and AncGT-Int Apo_{OFS}/NMD-out and Apo_{OFS}/NMD-in structures suggest that the energetics of low- to high-affinity transition might be determined by the packing of the extracellular half of the transport domain coupled to the NMD motif through the core TM7b helix. Mutations in HP2 of Glt_{Ph}, affecting the packing, alter substrate affinity (69,70). Moreover, slow alternative repacking of the region has been implicated in long-lasting binding modes exhibiting different substrate affinities (36). Several of the high-probability changes between Na^+ coupled and independent ancestors map to this region and one of them, $N327_{305}Q$ might be responsible for the stabilization of TM7b/TM3 packing, perhaps favoring the high-affinity configuration (**Supplementary Figure 15**). It is possible and likely that multiple mutations, combining in complex ways, are responsible for the phenotypic switch (71). Further disentangling of these changes will lead to a richer understanding of how ion-coupling works, is regulated, and can be rationally modulated.

508 **Methods**
509
510 *Sequence retrieval, curation, and alignment*
511
512 We designed a workflow to construct maximum likelihood phylogenies capturing the GT-Na to GT-H
513 evolutionary transition based on previous ASR studies (52) (**Supplementary Figure 3**). Aiming to have a
514 sequence set with a size amenable to ASR, we retrieved 20,000 non-redundant *Bacillota/Firmicutes*
515 phylum GT sequences via NCBI PSI-BLAST, using WP_011876993.1 as the original query (72,73). After
516 clustering by 99% using CD-HIT (74), removal of sequences with ambiguous amino acid assignment, and
517 alignment by MUSCLE v5 (75), a maximum parsimony tree was generated by MP-boot (76). Subsets of
518 sequences were selected based on proximity to sequences with signatures of H^+-coupling (as compared
519 to Glt_{Ph}: [T/S]92[A/L/I/V/M], M311L, D405N) and nearby sequences up until those with signature GT-Na
520 residues. The remaining sequences were clustered to 90% using CD-HIT, and the resulting alignment was
521 manually curated to exclude sequences with major gaps or insertions. Archaeal Glt_{Ph} (UniProt ID O59010)
522 was included to be the tree root. The final sequence set (1,145 sequences) was aligned by MAFFT-DASH,
523 which incorporates information from PDB structures into the alignment matrix (77). Finally, variable loops
524 and tails were removed, and the resulting alignment was manually adjusted to account for poorly aligned
525 regions.
526
527 *Phylogenetic analysis, ancestral sequence reconstruction, and construct design*
528
529 Maximum-likelihood (ML) trees and ancestral sequences were generated using IQ-TREE 2.1.4-beta (78).
530 Evaluation of all potential evolution models and rate heterogeneity parameters using ModelFinder revealed
531 that LG+F+R was the best protein evolution model (79–81). FreeRate (F) heterogeneity considers site-
532 specific differences in evolutionary rate, relaxing the constraint of equal site probabilities imposed by
533 discrete Gamma rate heterogeneity. The number of possible rate categories that could be assigned to a
534 site (R) was empirically determined for each run using ModelFinder to query a range of 10-20.
535
536 As with other ASR studies of microbial proteins, our tree topology is not reconciled to the microbial species
537 tree, as this has poor confidence especially at lower taxonomic ranks (50). We chose the outgroup-based
538 approach of rooting the tree, using archaeal Glt_{Ph} as the root due to the evolutionary distance between
539 bacteria and archaea. After performing over 60 replicate tree runs, we found that the best *Firmicutes* tree
540 yielded robust UFBoot2 values, ranging between 91-100% for the GT-H node (**Figure 1d**). When we
541 repeated the analysis for the *Proteobacteria* phylum, the best tree had 5-87% UFBoot2 values for the
542 equivalent nodes, reflecting non-reproducible trees. Thus, we used the *Firmicutes* tree for all subsequent
543 analyses.
544
545 The run yielding the best tree, as determined by log-likelihood, was inspected for long branches, defined
546 as > 0.7 substitutions/site. These were removed; in the second round of tree-searching, the analysis was
547 repeated with 254 independent tree search runs. The final tree had the best-fit model of LG+F+R14,
548 chosen according to the Bayesian Information Criterion. Bootstrap values were determined by UFBoot2
549 using 1000 replicates (82). Ancestral sequences were reconstructed from the best tree using the IQ-TREE
550 implementation (78) based on Yang's marginal reconstruction method (47–49). The ML sequences were
551 generated using custom Python scripts, either written manually or generated using ChatGPT. The final ML
552 amino acid sequence for each node is constructed using the amino acid with the highest posterior
553 probability (PP) at every site. All tree figures were generated using TreeViewer (83). Visual representations
554 of sequence alignments were made using WebLogos (84).
555
556 *Protein expression and purification*
557
558 Since poorly aligned loops and tails are not reconstructed in the ASR analysis, we had to graft them in for
559 expression constructs. We took comparable loops and tails from an extant sequence closest to AncGT-Int
560 (KJS87745.1); the 3L4 has two lysines that were mutated to histidines (K105H, K109H) to prevent potential
561 proteolysis. All protein constructs were codon-optimized for *E. coli*, synthesized (GenScript), and cloned
562 into a modified pBAD vector containing a C-terminal thrombin cleavage site followed by either an ALFA
563 peptide tag (85) or a C-terminal thermostable GFP which is specialized for membrane proteins (54). An

additional Twin-Strep tag and a 10x-HIS tag were added to the C-terminus of these constructs, all separated by glycine-serine (GSSS) linkers. The Glt_{Ph} construct used was the CAT7 construct, with seven histidines introduced for stability (9). Vectors were transformed into *E. coli* BL21(DE3) or DH10B cells, cells were grown in 2xYT media, and expression was induced by adding 0.1% arabinose at 0.8-1.0 OD_{600}. Proteins were expressed for 2-3 hours at 37°C. Membranes were purified as previously described for Glt_{Ph}, with all steps being performed at 4°C (36). Briefly, cells were broken on a high-pressure cell disruptor (Avestin) for three passes with cooled water continuously applied. The resulting lysate was spun at 17,400 x g for 15 min, and the supernatant was spun at 186,000 x g for 1 hour. The resulting membrane pellet was homogenized in 20 mM HEPES/NaOH pH 7.4, 10 mM EDTA, and 10% Sucrose; finally, the slow and fast spins were repeated, and the resulting pellet was frozen at -80°C.

All protein purification steps were performed at 4-8°C. Crude membranes were homogenized in 20 mM HEPES/NaOH, pH 7.4, 200 mM NaCl, 1 mM monopotassium L-aspartate (L-Asp); solubilization was initiated by the addition of 40 mM n-dodecyl-β-D-maltopyranoside (DDM; Anatrace), and the solubilizing membranes were gently rocked for two hours. Over the course of our experiments, we noticed that tags were being non-specifically proteolyzed upon solubilization; to reduce this, we added 1 mM PMSF and 1 mM EDTA before adding the DDM. After solubilization, samples were clarified by high-speed ultracentrifugation at 186,000 x g for 1 hour.

For ancestral constructs, the supernatant was applied to a prepacked 1 mL Streptactin XT 4Flow high-capacity column (IBA Biosciences) at a constant flow rate of 0.5-1 mL/min. The resin was washed until UV_{280} absorbance was flat, corresponding to fourteen column volumes of 20 mM HEPES/NaOH, pH 7.4, 200 mM NaCl, 1 mM L-Asp, and 1 mM DDM. Subsequently, the resin was eluted with the same buffer containing 50 mM biotin (IBA Biosciences), and proteins were concentrated.

For Glt_{Ph}, the supernatant was applied to a 1 mL HisTrap HP (Cytiva) at a constant flow rate of 1 mL/min. The resin was washed until UV_{280} absorbance was flat, corresponding to thirty column volumes of 20 mM HEPES/NaOH, pH 7.4, 200 mM NaCl, 1 mM L-Asp, 40 mM imidazole, and 1 mM DDM. Subsequently, the resin was eluted with the same buffer containing increased imidazole (250 mM), and proteins were concentrated.

Reconstitution into proteoliposomes

Unilamellar liposomes were prepared essentially as previously described (36). Briefly, a 3:1 (w:w) ratio of *E. coli* polar lipids and egg phosphatidylcholine (Avanti Polar Lipids) was dried, hydrated in 50 mM HEPES/Tris, pH 7.4, 200 mM KCl at a final concentration of 5 mg/mL, flash frozen, and stored at -80°C. On the day of the reconstitution, liposomes were thawed, extruded through a 400 nm filter, and destabilized with Triton X-100 using a 1:2 (w:w) Triton X-100 to lipid ratio. For the counterflow and Na^+-coupled uptake experiments, GFP-free proteins were purified by SEC using Superdex 200 Increase 10/300 column (Cytiva) in 20 mM HEPES/NaOH, pH 7.4, 200 mM NaCl, 1 mM L-Asp, and 1 mM DDM; proteins were subsequently concentrated using a 100 kDa cutoff concentrator. To form proteoliposomes, the purified protein was mixed with destabilized liposomes (1:100, w:w) for 30 minutes, then detergent was removed with BioBeads SM-2 (Bio-Rad) at 80 mg/ml, which were applied twice at 22°C for two hours, once overnight at 4°C, and once more at 22°C for two hours.

Uptake and counterflow experiments

After detergent removal with BioBeads, proteoliposomes underwent the same general procedure for buffer exchange and concentration, regardless of experiment. Proteoliposomes were exchanged into the appropriate buffer using a PD-10 column (Cytiva), flash frozen in liquid N_2, thawed, concentrated via ultracentrifugation (100,000 x g for one hour), resuspended to a final concentration of 25-50 mg/mL, flash frozen, and stored at -80°C. On the day of the experiment, the proteoliposomes were thawed and extruded through a 400 nm filter 13 times. To minimize sample dilution in the extruder, the proteoliposome volume was no less than 100 µL. All subsequent experiments were performed at 30□C.

619 For uptake experiments, liposomes were prepared as described above to a final concentration of 50 mg/mL
620 in 20 mM HEPES/Tris pH 7.4, 200 mM KCl, and 100 mM choline chloride (ChoCl). To initiate Na^+-driven
621 uptake in the presence of a negative membrane potential (-102 mV), liposomes were diluted 1:100 in 20
622 mM HEPES/Tris pH 7.4, 2 mM KCl, 198 mM choline chloride, 100 mM NaCl, 0.9 μM valinomycin, and 1
623 μM ^{3}H-L-Asp (PerkinElmer). For negative controls without Na^+-gradient, the NaCl was replaced with
624 choline chloride.
625
626 To stop the reactions, 200 μL of reaction mixture (100 μg of liposomes) was diluted in 2 mL ice-cold
627 quenching buffer (20 mM HEPES/Tris pH 7.4, 200 mM lithium chloride). The quenched reaction was
628 immediately filtered through a 0.22-μm filter membrane (Millipore Sigma) and washed three times with 2
629 ml quenching buffer. Washed membranes were inserted into scintillation vials, and the membranes were
630 soaked overnight in 5 ml Econo-Safe Counting Cocktail. Radioactivity in liposomes was measured using
631 an LS6500 scintillation counter (Beckman Coulter). Technical replicates from an individual reconstitution
632 were taken in duplicate and averaged, and at least two separate reconstitutions were taken per point.
633
634 *Nano Differential Scanning Fluorimetry (nanoDSF)*
635
636 All steps were performed on the same day of the protein solubilization and purification, since we observed
637 variance in results depending on the freshness of the protein. We opted to use GFP-tagged ancestral
638 proteins because in the GFP-free protein, exact T_ms could not be precisely determined in certain
639 conditions. Furthermore, even in the presence of detergent, the T_m of the fused GFP is greater than 90ºC
640 (54), which does not interfere the T_ms observed in our experiments. After affinity purification, the proteins
641 with C-terminal GFP were concentrated using a 100 kDa cutoff concentrator (MilliporeSigma). To remove
642 ligands, the protein was dialyzed using a PD MiniTrap G-25 (Cytiva). For AncGT-Na, this buffer was 20
643 mM HEPES, 200 mM NMDG-Cl, pH 7.4, and 1 mM DDM; for AncGT-Int, this buffer was 50 mM
644 KH_2PO_4/K_2HPO_4 buffer, pH 7, and 1 mM DDM. Different buffers were used for optimal first derivative
645 curves, and we observed identical T_ms regardless of dialysis buffer composition. The dialyzed, apo protein
646 was concentrated and diluted into the same respective buffers at a final concentration of ~1 mg/mL; for
647 AncGT-Int, the buffer contained 10 mM decyl-maltoside (DM) instead of 1 mM DDM. This was necessary
648 to observe melting of AncGT-Int, presumably by the shorter chain detergent destabilizing the overall
649 protein. Thermal stability in the presence and absence of various potential ligands was measured using a
650 Tycho NT.6 instrument (Nanotemper).
651
652 *Microscale thermophoresis (MST)*
653
654 After affinity purification, the proteins with C-terminal GFP were concentrated using a 100 kDa cutoff
655 concentrator, then SEC-purified on Superose 6 Increase 10/300 column (Cytiva) in 50 mM
656 KH_2PO_4/K_2HPO_4 buffer pH 7 and 1 mM DDM. For the MST experiment, monopotassium L-aspartate was
657 serially diluted 2-fold in water for a total of 16 concentrations. Each aspartate dilution was mixed with in a
658 1:1 (v:v) ratio of protein samples, so that the final protein concentration was 100-300 nM, the appropriate
659 salt was 100 mM, the DDM was 0.4 mM, and the KH_2PO_4/K_2HPO_4 was 10-20 mM. Measurements were
660 taken using Monolith NT.115 Blue/Red MST (Nanotemper), High MST power, Blue detector (for GFP), and
661 premium capillaries (Nanotemper).
662
663 For data analysis, obvious outliers caused by aggregates (as determined by instrument software) were
664 discarded. The ideal response time for each individual experiment was determined by instrument software,
665 and this was always between 1.5s-5s. Individual binding titrations were normalized to fraction bound by K_D
666 fits in the instrument software. To obtain the final binding affinity value, the experiments were repeated
667 three independent times, the data points were averaged, and the K_D was calculated from the average
668 values.
669
670 *Cryo-EM sample preparation*
671
672 Membrane Scaffold Protein (MSP1E3) was expressed and purified as previously described (86) in Na^+/L-
673 Asp free conditions. A 3:1 (w:w) ratio of *E. coli* polar lipids and egg phosphatidylcholine (Avanti Polar

674 Lipids) was dried, hydrated in 50 mM HEPES/Tris, pH 7.4 and 200 mM KCl, and solubilized in 72 mM DDM
675 at final lipid concentration of 18 mM as previously described (15,86).
676
677 For AncGT-Int, GFP-free transporter constructs were affinity purified, concentrated, and mixed with
678 MSP1E3 and solubilized lipids (both containing no Na^+ or L-aspartate) at a molar ratio of 1.5:1:50, aiming
679 for a final concentration of lipid greater than 5 mM. This mixture was incubated at 22°C for 30 minutes,
680 then 350 mg BioBeads SM-2 (BioRad) per mL mixture (w:v) was added for two hours at 22°C with gentle
681 shaking, and another round of BioBeads was applied overnight at 4°C with gentle shaking. The resulting
682 MSP1E3/protein preparations were cleared by ultracentrifugation at 100,000 x g for 30 minutes. The
683 supernatant was applied to a Superose 6 Increase 10/300 column in 20 mM HEPES pH 7.4, 100 mM
684 NMDG-Cl, and the peak fractions were concentrated to 7-9 mg/mL, as measured by A280 absorbance.
685 Reconstitution efficiency was verified using SDS-PAGE.
686 Glt_{Ph} was reconstituted into MSP1E3 as previously described (15). Affinity purified protein was further
687 purified via size-exclusion chromatography (Superdex 200 Increase, 10/300). Peak fractions were
688 concentrated, mixed with MSP1E3 and solubilized lipids at a molar ratio of 0.75:1:50, aiming for a final
689 concentration of lipid greater than 5 mM. This mixture was incubated and treated with BioBeads SM-2 in
690 the exact same fashion as ancestral protein. The resulting MSP1E3/protein preparations were exchanged
691 in Na^+-free buffer by three rounds of 10-fold dilution/concentration, and cleared by ultracentrifugation at
692 100,000 x g for 30 minutes. The supernatant was applied to a Superose 6 Increase 10/300 column in 20
693 mM HEPES/Tris pH 7.4 and 50-mM choline chloride. The peak fractions were concentrated to 8 mg/mL,
694 as measured by A280 absorbance. Reconstitution efficiency was verified using SDS-PAGE.
695
696 For the AncGT-Int substrate-bound dataset, 1 mM L-Asp was added to the sample just before freezing. To
697 improve particle distribution, fluorinated Fos-Choline-8 (Anatrace) dissolved in water was added to the
698 sample to a final concentration of 1.5 mM. UltrAuFoil R1.2/1.3 300-mesh gold grids (Quantifoil) were glow-
699 discharged at 25 mA for 80 s. 3 µL of protein was applied to each grid, incubated for 20 s under 100%
700 humidity at 4°C, blotted for 2.5 s under 0 blot force, and plunge frozen in liquid ethane using a Vitrobot
701 Mark IV (Thermo Fisher Scientific).
702
703 *Cryo-EM data acquisition and processing*
704
705 Imaging was performed on a Titan Krios (FEI) equipped with a K3 camera (Gatan) using Leginon software.
706 Specifics of each data collection and each individual map and model are provided in **Supplementary**
707 **Table 1.** Visual summaries of processing workflows and EM validation are provided in **Supplementary**
708 **Figures 9-12**.
709
710 <u>Initial processing pipeline for AncGT-Int</u>
711
712 For all datasets, raw movies were imported into cryoSPARC 4.0-4.4 and subjected to Patch Motion
713 Correction and Patch CTF Estimation, where the movies were binned 2x (87). Low-quality movies were
714 removed, as determined by CTF fits, ice thickness, intensity, etc. Particles were manually picked to train
715 Topaz (88), where for each data set, the optimal picking model and estimated number of particles per
716 micrograph were empirically determined using the Topaz Cross Validation tool in cryoSPARC. After picking
717 and extraction with 4x binning, an *ab initio* 3D model was constructed, and high-quality particles were
718 sorted using multiple heterogeneous refinement jobs, where 'good' particles sort into the *ab initio* model
719 class, and 'junk' particles sort into three pure noise 'decoy' volume classes (created by performing one
720 iteration of *ab initio*). When greater than 95% of particles converged into the protein volume, the remaining
721 particles were re-extracted to full box size and subjected to 3D Classification without alignment, and Non-
722 Uniform (NU) Refinement (89). To maximize the number of particles obtained, the Topaz model was
723 retrained on the clean particle stack after rounds of 2D classification, and heterogeneous refinement was
724 repeated with the new particle stack and the high-resolution volume from NU-Refinement (replacing the
725 *ab initio* volume). Remaining particles were re-extracted to full box size and subjected to NU-Refinement.
726
727
728
729

730 AncGT-Int apo data set processing
731
732 Our ancestral protein datasets were highly heterogeneous and poorly symmetric; throughout the
733 processing we realized that they were reminiscent of other seemingly mobile membrane protein structures
734 (90,91). To overcome this, we attempted to enrich particles with symmetric features so that we could
735 classify individual protomers. In the apo dataset, we converted particle coordinates to Relion format using
736 PyEM (92). Particles were re-extracted in Relion using movies motion corrected and binned 2x using
737 Relion's implementation of MotionCor2 (93). These particles were re-imported into cryoSPARC, refined
738 using NU-Refinement (C1), and the resulting particle images were converted to Relion format with PyEM.
739 These particles were then subjected to two rounds of Bayesian polishing (94), classification with
740 heterogeneous refinement, and classification with NU-Refinement (C1). To remove 'disordered' particles,
741 we used 3D classification without alignment (C1) in cryoSPARC. We further optimized symmetric particles
742 by performing one round of *ab initio* with C3 symmetry imposed. Using the best class, we achieved a
743 resolution of 2.8 Å in C3, which allowed us to resolve individual protomers using symmetry expansion and
744 focused classification using a mask on the transport domain. Each class underwent local refinement
745 masked on the entire protomer (including transport/scaffold domains) using pose/shift Gaussian priors
746 during alignment (standard deviations of 3º, 2 Å).
747
748 AncGT-Int substrate-bound processing
749
750 As with the apo dataset, the substrate-bound dataset was similarly heterogeneous and asymmetric. Using
751 the previous approach, we encountered trouble imposing C3 symmetry, preventing focused classification
752 of individual protomers. During our processing, we realized that we could more efficiently remove
753 'disordered' particles using outputs from 3D Variability, where the first variability component reflects this
754 'order-disorder' transition between rigid, easily aligned particles and presumably mobile, difficult to align
755 particles. We then took the first frame of the 3DVA movie (disordered, 'bad' volume), the last frame of the
756 movie (ordered, 'good' volume), and two decoy noise volumes as inputs for heterogeneous refinement.
757 Though a fundamentally similar approach to 3D classification used in the apo dataset, the new approach
758 allowed us to impose C3 symmetry with good resolutions (3.4 Å). We then symmetry expanded the
759 particles, performed 3D classification without alignment on the masked transport domain, and locally
760 refined the protomers using pose/shift Gaussian priors during alignment (standard deviations of 3º, 2 Å),
761 as with the apo dataset.
762
763 Glt_{Ph} apo processing
764
765 Raw movies were imported into cryoSPARC 4.4 and subjected to Patch Motion Correction and Patch CTF
766 Estimation, where the movies were binned 2x (87). Low-quality movies were removed, as determined by
767 CTF fits, ice thickness, intensity, etc. Particles were manually picked to train Topaz (88), where the optimal
768 picking model and estimated number of particles per micrograph were empirically determined using the
769 Topaz Cross Validation tool in cryoSPARC. After picking and extraction, an *ab initio* 3D model was
770 constructed. After one round of heterogeneous refinement with *ab initio* volumes, we performed NU-
771 Refinement (C3) on the remaining particles, symmetry expanded the particles, performed 3D classification
772 without alignment on the masked transport domain, and locally refined the protomers using pose/shift
773 Gaussian priors during alignment (standard deviations of 3º, 2 Å).
774
775 *Model building and refinement*
776
777 For model building, maps were sharpened using DeepEMhancer (95). We generated a homology model
778 of AncGT-Int in SWISS-MODEL, using an outward-facing state of Glt_{Ph} (PDB ID 2NWX) as a template
779 (9,96). The templates for Glt_{Ph} were either 6UWF (OFS) or 6UWL (iOFS). The model was adjusted into the
780 density using ISOLDE (97), and the final model was real-spaced refined and validated in Phenix using the
781 unsharpened density and default parameters (98). All structural figures were made with ChimeraX (99).
782
783 **Data availability**
784
785 Bioinformatic analyses, structures, and EM datasets will be made available immediately upon publication.

Author contributions

K.D.R. and O.B. conceived the project and wrote the manuscript. K.D.R. performed all evolutionary analyses. K.D.R. and B.R. performed the functional experiments and prepared the figures. K.D.R., B.R., and F.B.A. prepared grids and processed cryo-EM data. K.D.R., B.R., F.B.A., and O.B. built and refined the molecular models. N.K. and S.P. performed foundational pilot experiments, including expression/purification optimization, and preliminary cryo-EM data collection and processing.

Acknowledgments

We thank Joseph Thornton and members of the Thornton lab, Michael Harms, Luce Skrabanek, and Eva Fortea for helpful discussions regarding bioinformatics and phylogenetics. We thank Will Eng for assistance with protein expression, GenScript for prompt and patient support with gene synthesis, and Vanessa Braband, Bridget Milorey, and Dinorah Leyva from NanoTemper for helping with troubleshooting MST parameters. We thank members of the Boudker lab, especially Amanda Scopelliti, Xiaoyu Wang, Biao Qiu, and Yessenia Lopez, for helpful discussions regarding biochemistry and cryo-EM. The work was supported by National Institutes of Health (NIH) grants F32 NS102325 (to K.D.R.) and NIH grants R01NS064357 and R37NS085318 (to O.B.). O.B. is an investigator of the Howard Hughes Medical Institute.

Phylogenetic analyses were performed on the Weill Cornell HPC cluster maintained by the Scientific Computing Unit. All cryo-EM grids were initially screened at the Weill Cornell Cryo-EM Core Facility with help from Carl Fluck. Apo AncGT-Int was collected at the NYU cryo-EM facility with help from Bing Wang. Substrate-bound AncGT-Int was collected at the Columbia University Medical Center Cryo-EM core with assistance from Robert Grassucci. An earlier substrate-bound AncGT-Int dataset (with assistance from Kasahun Nehelu) and the apo-Glt_{Ph} dataset (with assistance from Aaron Owji and Aygul Ishemgulova) were collected at the Simons Electron Microscopy Center and National Resource for Automated Molecular Microscopy located at the New York Structural Biology Center, supported by grants from the Simons Foundation (349247), NYSTAR, and the NIH National Institute of General Medical Sciences (GM103310) with additional support from Agouron Institute (F00316) and NIH S10 OD019994-01.

References

1. Lane N, Martin WF. The Origin of Membrane Bioenergetics. Cell. 2012 Dec 21;151(7):1406–16.
2. Paulsen IT, Sliwinski MK, Saier MH. Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificities11Edited by G. Von Heijne. Journal of Molecular Biology. 1998 Apr 6;277(3):573–92.
3. Leone V, Pogoryelov D, Meier T, Faraldo-Gómez JD. On the principle of ion selectivity in Na+/H+-coupled membrane proteins: Experimental and theoretical studies of an ATP synthase rotor. Proceedings of the National Academy of Sciences. 2015 Mar 10;112(10):E1057–66.
4. Forrest LR, Krämer R, Ziegler C. The structural basis of secondary active transport mechanisms. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 2011 Feb 1;1807(2):167–88.
5. Jin Y, Nair A, Veen HW van. Multidrug Transport Protein NorM from Vibrio cholerae Simultaneously Couples to Sodium- and Proton-Motive Force *. Journal of Biological Chemistry. 2014 May 23;289(21):14624–32.
6. Schlegel K, Leone V, Faraldo-Gómez JD, Müller V. Promiscuous archaeal ATP synthase concurrently coupled to Na+ and H+ translocation. Proceedings of the National Academy of Sciences. 2012 Jan 17;109(3):947–52.
7. Ryan RM, Compton ELR, Mindell JA. Functional characterization of a Na+-dependent aspartate transporter from Pyrococcus horikoshii. J Biol Chem. 2009 Jun 26;284(26):17540–8.
8. Gaillard I, Slotboom DJ, Knol J, Lolkema JS, Konings WN. Purification and reconstitution of the glutamate carrier GltT of the thermophilic bacterium Bacillus stearothermophilus. Biochemistry. 1996 May 14;35(19):6150–6.
9. Yernool D, Boudker O, Jin Y, Gouaux E. Structure of a glutamate transporter homologue from Pyrococcus horikoshii. Nature. 2004 Oct 14;431(7010):811–8.

10. Boudker O, Ryan RM, Yernool D, Shimamoto K, Gouaux E. Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. Nature. 2007 Jan 25;445(7126):387–93.
11. Jensen S, Guskov A, Rempel S, Hänelt I, Slotboom DJ. Crystal structure of a substrate-free aspartate transporter. Nat Struct Mol Biol. 2013 Oct;20(10):1224–6.
12. Ruan Y, Miyagi A, Wang X, Chami M, Boudker O, Scheuring S. Direct visualization of glutamate transporter elevator mechanism by high-speed AFM. Proc Natl Acad Sci U S A. 2017 Feb 14;114(7):1584–8.
13. Crisman TJ, Qu S, Kanner BI, Forrest LR. Inward-facing conformation of glutamate transporters as revealed by their inverted-topology structural repeats. Proc Natl Acad Sci U S A. 2009 Dec 8;106(49):20752–7.
14. Reyes N, Ginter C, Boudker O. Transport mechanism of a bacterial homologue of glutamate transporters. Nature. 2009 Dec 17;462(7275):880–5.
15. Wang X, Boudker O. Large domain movements through the lipid bilayer mediate substrate release and inhibition of glutamate transporters. Elife. 2020 Nov 6;9:e58417.
16. Arkhipova V, Guskov A, Slotboom DJ. Structural ensemble of a glutamate transporter homologue in lipid nanodisc environment. Nat Commun. 2020 Feb 21;11(1):998.
17. Gendreau S, Voswinkel S, Torres-Salazar D, Lang N, Heidtmann H, Detro-Dassen S, et al. A Trimeric Quaternary Structure Is Conserved in Bacterial and Human Glutamate Transporters *. Journal of Biological Chemistry. 2004 Sep 17;279(38):39505–12.
18. Garaeva AA, Guskov A, Slotboom DJ, Paulino C. A one-gate elevator mechanism for the human neutral amino acid transporter ASCT2. Nat Commun. 2019 Jul 31;10(1):3427.
19. Qiu B, Matthies D, Fortea E, Yu Z, Boudker O. Cryo-EM structures of excitatory amino acid transporter 3 visualize coupled substrate, sodium, and proton binding and transport. Sci Adv. 2021 Mar;7(10):eabf5814.
20. Alleva C, Kovalev K, Astashkin R, Berndt MI, Baeken C, Balandin T, et al. Na+-dependent gate dynamics and electrostatic attraction ensure substrate coupling in glutamate transporters. Science Advances. 2020 Nov 18;6(47):eaba9854.
21. Guskov A, Jensen S, Faustino I, Marrink SJ, Slotboom DJ. Coupled binding mechanism of three sodium ions and aspartate in the glutamate transporter homologue GltTk. Nat Commun. 2016 Nov 10;7:13420.
22. Garaeva AA, Slotboom DJ. Elevator-type mechanisms of membrane transport. Biochemical Society Transactions. 2020 May 5;48(3):1227–41.
23. Yu X, Plotnikova O, Bonin PD, Subashi TA, McLellan TJ, Dumlao D, et al. Cryo-EM structures of the human glutamine transporter SLC1A5 (ASCT2) in the outward-facing conformation. Boudker O, Forrest L, Vandenberg RJ, Meyerson J, editors. eLife. 2019 Oct 3;8:e48120.
24. Groeneveld M, Weme RGJDO, Duurkens RH, Slotboom DJ. Biochemical characterization of the C4-dicarboxylate transporter DctA from Bacillus subtilis. J Bacteriol. 2010 Jun;192(11):2900–7.
25. Tolner B, Ubbink-Kok T, Poolmann B, Konings WN. Cation-selectivity of the l-glutamate transporters of Escherichia coli, Bacillus stearothermophilus and Bacillus caldotenax: dependence on the environment in which the proteins are expressed. Molecular Microbiology. 1995;18(1):123–33.
26. Groeneveld M, Slotboom DJ. Rigidity of the Subunit Interfaces of the Trimeric Glutamate Transporter GltT During Translocation. Journal of Molecular Biology. 2007 Sep 21;372(3):565–70.
27. Yernool D, Boudker O, Folta-Stogniew E, Gouaux E. Trimeric subunit stoichiometry of the glutamate transporters from Bacillus caldotenax and Bacillus stearothermophilus. Biochemistry. 2003 Nov 11;42(44):12981–8.
28. Oh S, Boudker O. Kinetic mechanism of coupled binding in sodium-aspartate symporter GltPh. Elife. 2018 Sep 26;7:e37291.
29. Reyes N, Oh S, Boudker O. Binding thermodynamics of a glutamate transporter homolog. Nat Struct Mol Biol. 2013 May;20(5):634–40.
30. Bastug T, Heinzelmann G, Kuyucak S, Salim M, Vandenberg RJ, Ryan RM. Position of the third Na+ site in the aspartate transporter GltPh and the human glutamate transporter, EAAT1. PLoS One. 2012;7(3):e33058.
31. Verdon G, Oh S, Serio RN, Boudker O. Coupled ion binding and structural transitions along the transport cycle of glutamate transporters. Elife. 2014 May 19;3:e02283.

32. Qiu B, Boudker O. Symport and antiport mechanisms of human glutamate transporters. Nat Commun. 2023 May 4;14(1):2579.
33. Zhou W, Trinco G, Slotboom DJ, Forrest LR, Faraldo-Gómez JD. On the Role of a Conserved Methionine in the Na+-Coupling Mechanism of a Neurotransmitter Transporter Homolog. Neurochem Res. 2022 Jan;47(1):163–75.
34. Riederer EA, Valiyaveetil FI. Investigation of the allosteric coupling mechanism in a glutamate transporter homolog via unnatural amino acid mutagenesis. Proceedings of the National Academy of Sciences. 2019 Aug 6;116(32):15939–46.
35. Setiadi J, Kuyucak S. Elucidation of the Role of a Conserved Methionine in Glutamate Transporters and Its Implication for Force Fields. J Phys Chem B. 2017 Oct 19;121(41):9526–31.
36. Reddy KD, Ciftci D, Scopelliti AJ, Boudker O. The archaeal glutamate transporter homologue GltPh shows heterogeneous substrate binding. Journal of General Physiology. 2022 May 2;154(5):e202213131.
37. Tao Z, Zhang Z, Grewer C. Neutralization of the Aspartic Acid Residue Asp-367, but Not Asp-454, Inhibits Binding of Na+ to the Glutamate-free Form and Cycling of the Glutamate Transporter EAAC1*. Journal of Biological Chemistry. 2006 Apr 14;281(15):10263–72.
38. Riederer EA, Moënne-Loccoz P, Valiyaveetil FI. Distinct roles of the Na+ binding sites in the allosteric coupling mechanism of the glutamate transporter homolog, GltPh. Proceedings of the National Academy of Sciences. 2022 May 10;119(19):e2121653119.
39. Starr TN, Thornton JW. Epistasis in protein evolution. Protein Science. 2016;25(7):1204–18.
40. Hochberg GKA, Thornton JW. Reconstructing Ancient Proteins to Understand the Causes of Structure and Function. Annu Rev Biophys. 2017 May 22;46(1):247–69.
41. Wilson C, Agafonov RV, Hoemberger M, Kutter S, Zorba A, Halpin J, et al. Kinase dynamics. Using ancient protein kinases to unravel a modern cancer drug's mechanism. Science. 2015 Feb 20;347(6224):882–6.
42. Ugalde JA, Chang BSW, Matz MV. Evolution of Coral Pigments Recreated. Science. 2004 Sep 3;305(5689):1433–1433.
43. Harms MJ, Thornton JW. Historical contingency and its biophysical basis in glucocorticoid receptor evolution. Nature. 2014 Aug;512(7513):203–7.
44. Gaucher EA, Govindarajan S, Ganesh OK. Palaeotemperature trend for Precambrian life inferred from resurrected proteins. Nature. 2008 Feb;451(7179):704–7.
45. Pillai AS, Chandler SA, Liu Y, Signore AV, Cortez-Romero CR, Benesch JLP, et al. Origin of complexity in haemoglobin evolution. Nature. 2020 May;581(7809):480–5.
46. Lim SA, Hart KM, Harms MJ, Marqusee S. Evolutionary trend toward kinetic stability in the folding trajectory of RNases H. Proceedings of the National Academy of Sciences. 2016 Nov 15;113(46):13045–50.
47. Yang Z, Kumar S, Nei M. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics. 1995 Dec;141(4):1641–50.
48. Hanson-Smith V, Kolaczkowski B, Thornton JW. Robustness of ancestral sequence reconstruction to phylogenetic uncertainty. Mol Biol Evol. 2010 Sep;27(9):1988–99.
49. Koshi JM, Goldstein RA. Probabilistic reconstruction of ancestral protein sequences. J Mol Evol. 1996 Feb;42(2):313–20.
50. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 2018 Nov;36(10):996–1004.
51. Eick GN, Bridgham JT, Anderson DP, Harms MJ, Thornton JW. Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty. Mol Biol Evol. 2017 Feb 1;34(2):247–61.
52. Orlandi KN, Phillips SR, Sailer ZR, Harman JL, Harms MJ. Topiary: Pruning the manual labor from ancestral sequence reconstruction. Protein Sci. 2023 Feb;32(2):e4551.
53. Mulligan C, Mindell JA. Mechanism of transport modulation by an extracellular loop in an archaeal excitatory amino acid transporter (EAAT) homolog. J Biol Chem. 2013 Dec 6;288(49):35266–76.
54. Cai H, Yao H, Li T, Hutter CAJ, Li Y, Tang Y, et al. An improved fluorescent tag and its nanobodies for membrane protein expression, stability assay, and purification. Commun Biol. 2020 Dec 10;3(1):753.

55. Majd H, King MS, Palmer SM, Smith AC, Elbourne LD, Paulsen IT, et al. Screening of candidate substrates and coupling ions of transporters by thermostability shift assays. Elife. 2018 Oct 15;7:e38821.
56. Nji E, Chatzikyriakidou Y, Landreh M, Drew D. An engineered thermal-shift screen reveals specific lipid preferences of eukaryotic and prokaryotic membrane proteins. Nat Commun. 2018 Oct 12;9(1):4253.
57. Ewers D, Becher T, Machtens JP, Weyand I, Fahlke C. Induced fit substrate binding to an archeal glutamate transporter homologue. Proceedings of the National Academy of Sciences. 2013 Jul 23;110(30):12486–91.
58. Hänelt I, Jensen S, Wunnicke D, Slotboom DJ. Low Affinity and Slow Na+ Binding Precedes High Affinity Aspartate Binding in the Secondary-active Transporter GltPh*. Journal of Biological Chemistry. 2015 Jun 26;290(26):15962–72.
59. Verdon G, Boudker O. Crystal structure of an asymmetric trimer of a bacterial glutamate transporter homolog. Nat Struct Mol Biol. 2012 Feb 12;19(3):355–7.
60. Huang Y, Wang X, Lv G, Razavi AM, Huysmans GHM, Weinstein H, et al. Use of paramagnetic 19F NMR to monitor domain movement in a glutamate transporter homolog. Nat Chem Biol. 2020 Sep;16(9):1006–12.
61. Bendahan A, Armon A, Madani N, Kavanaugh MP, Kanner BI. Arginine 447 Plays a Pivotal Role in Substrate Interactions in a Neuronal Glutamate Transporter *. Journal of Biological Chemistry. 2000 Dec 1;275(48):37436–42.
62. Scopelliti AJ, Ryan RM, Vandenberg RJ. Molecular Determinants for Functional Differences between Alanine-Serine-Cysteine Transporter 1 and Other Glutamate Transporter Family Members *. Journal of Biological Chemistry. 2013 Mar 22;288(12):8250–7.
63. Scopelliti AJ, Font J, Vandenberg RJ, Boudker O, Ryan RM. Structural characterisation reveals insights into substrate recognition by the glutamine transporter ASCT2/SLC1A5. Nat Commun. 2018 Jan 2;9(1):38.
64. Seal RP, Amara SG. A Reentrant Loop Domain in the Glutamate Carrier EAAT1 Participates in Substrate Binding and Translocation. Neuron. 1998 Dec 1;21(6):1487–98.
65. Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ, et al. Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. Curr Biol. 2015 Mar 16;25(6):690–701.
66. Motlagh HN, Wrabl JO, Li J, Hilser VJ. The ensemble nature of allostery. Nature. 2014 Apr;508(7496):331–9.
67. Saavedra HG, Wrabl JO, Anderson JA, Li J, Hilser VJ. Dynamic allostery can drive cold adaptation in enzymes. Nature. 2018 Jun;558(7709):324–8.
68. Sailer ZR, Harms MJ. Molecular ensembles make evolution unpredictable. Proceedings of the National Academy of Sciences. 2017 Nov 7;114(45):11938–43.
69. Huysmans GHM, Ciftci D, Wang X, Blanchard SC, Boudker O. The high-energy transition state of the glutamate transporter homologue GltPh. The EMBO Journal. 2021 Jan 4;40(1):e105415.
70. Ciftci D, Martens C, Ghani VG, Blanchard SC, Politis A, Huysmans GHM, et al. Linking function to global and local dynamics in an elevator-type transporter. Proc Natl Acad Sci U S A. 2021 Dec 7;118(49):e2025520118.
71. Morrison AJ, Wonderlick DR, Harms MJ. Ensemble epistasis: thermodynamic origins of nonadditivity between mutations. Genetics. 2021 Sep 1;219(1):iyab105.
72. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research. 1997 Sep 1;25(17):3389–402.
73. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic Acids Research. 2008 Jul 1;36(suppl_2):W5–9.
74. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012 Dec 1;28(23):3150–2.
75. Edgar RC. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat Commun. 2022 Nov 15;13(1):6968.
76. Hoang DT, Vinh LS, Flouri T, Stamatakis A, von Haeseler A, Minh BQ. MPBoot: fast phylogenetic maximum parsimony tree inference and bootstrap approximation. BMC Evolutionary Biology. 2018 Feb 2;18(1):11.

77. Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res. 2019 Jul 2;47(W1):W5–10.
78. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020 May 1;37(5):1530–4.
79. Soubrier J, Steel M, Lee MSY, Der Sarkissian C, Guindon S, Ho SYW, et al. The influence of rate heterogeneity among sites on the time dependence of molecular rates. Mol Biol Evol. 2012 Nov;29(11):3345–58.
80. Le SQ, Gascuel O. An improved general amino acid replacement matrix. Mol Biol Evol. 2008 Jul;25(7):1307–20.
81. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017 Jun;14(6):587–9.
82. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol. 2018 Feb 1;35(2):518–22.
83. Bianchini G, Sánchez-Baracaldo P. TreeViewer Version 2.1.0 [Internet]. Zenodo; 2023 [cited 2023 Aug 7]. Available from: https://zenodo.org/record/7768344
84. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A Sequence Logo Generator. Genome Res. 2004 Jun 1;14(6):1188–90.
85. Götzke H, Kilisch M, Martínez-Carranza M, Sograte-Idrissi S, Rajavel A, Schlichthaerle T, et al. The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. Nat Commun. 2019 Sep 27;10(1):4403.
86. Ritchie TK, Grinkova YV, Bayburt TH, Denisov IG, Zolnerciks JK, Atkins WM, et al. Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. Methods Enzymol. 2009;464:211–31.
87. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017 Mar;14(3):290–6.
88. Bepler T, Morin A, Rapp M, Brasch J, Shapiro L, Noble AJ, et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nat Methods. 2019 Nov;16(11):1153–60.
89. Punjani A, Zhang H, Fleet DJ. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. Nat Methods. 2020 Dec;17(12):1214–21.
90. Vallese F, Kim K, Yen LY, Johnston JD, Noble AJ, Calì T, et al. Architecture of the human erythrocyte ankyrin-1 complex. Nat Struct Mol Biol. 2022 Jul;29(7):706–18.
91. Paknejad N, Sapuru V, Hite RK. Structural titration reveals Ca2+-dependent conformational landscape of the IP3 receptor. Nat Commun. 2023 Oct 28;14(1):6897.
92. Asarnow D, Palovcak E, Cheng Y. asarnow/pyem: UCSF pyem v0.5 [Internet]. Zenodo; 2019 [cited 2023 Jul 23]. Available from: https://zenodo.org/record/3576630
93. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods. 2017 Apr;14(4):331–2.
94. Zivanov J, Nakane T, Scheres SHW. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. IUCrJ. 2019 Jan 1;6(Pt 1):5–17.
95. Sanchez-Garcia R, Gomez-Blanco J, Cuervo A, Carazo JM, Sorzano COS, Vargas J. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. Commun Biol. 2021 Jul 15;4(1):1–8.
96. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Research. 2018 Jul 2;46(W1):W296–303.
97. Croll TI. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr D Struct Biol. 2018 Jun 1;74(Pt 6):519–30.
98. Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, et al. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol. 2018 Jun 1;74(Pt 6):531–44.
99. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 2018 Jan;27(1):14–25.