5972 *Biochemistry* **2010,** *49,* 5972–5974
DOI: 10.1021/bi100918j

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Rapid Report

Evolution of Antiparallel Two-Domain Membrane Proteins. Swapping Domains in the Glutamate Transporter GltS[†]

Adam Dobrowolski and Juke S. Lolkema*

Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

Received June 8, 2010; Revised Manuscript Received June 17, 2010

ABSTRACT: Two-domain membrane proteins are believed to have evolved through duplication and fusion events. A set of evolutionary states of the Na^+-glutamate transporter of *Escherichia coli* was engineered. The two half-genes encoding the two domains were placed in a single operon in both orders ($GltS^{split}$), and the split genes were fused in the reverse order compared to the original protein ($GltS^{swap}$). The transporter halves were produced and shown to be active in Na^+-coupled glutamate transport. $GltS^{swap}$ was as active as the original transporter provided that the domains were connected by a linker of the same size that connected them in the original transporter.

Many membrane proteins consist of two domains that share a similar fold. A plausible model for the evolution of two-domain membrane proteins with an internal repeat involves duplication of a primordial gene followed by fusion, thus resulting in a single gene encoding a protein with two homologous domains (*1*, *2*) (Figure 1). Alternatively, the duplication and fusion steps proceed in a single step ("in gene" duplication) (*3*). The two domains have the same (parallel) or opposite (antiparallel) orientation in the membrane corresponding to an even or odd number of transmembrane segments (TMS) per domain, respectively. To account for the antiparallel orientation of the two domains, the ancestral membrane protein is hypothesized to be "dual topology"; i.e., it inserts with a random orientation into the membrane. Following duplication, the two dual topology proteins would adopt fixed but opposite orientations by genetic drift, resulting in the introduction of positively charged amino acid residues in cytoplasmic loops (positive-inside rule) (*4*). The first gene on the chromosome may encode one orientation or the other; at the protein level, this has no consequences for the antiparallel heterodimer that is formed. However, in the fused state, this results in two different proteins with the N-terminus either inside or outside the cell (Figure 1). Therefore, the evolutionary pathway has two outcomes representing proteins that differ in the order of the two domains in the primary structure. In the course of evolution, one of these outcomes will be selected at random or because of a selective advantage.

GltS of *Escherichia coli* transports glutamate in symport with Na^+ ions (*5*, *6*) and is a member of the bacterial ESS family [glutamate:Na^+ symporter, TC 2.A.27 (*7*)]. The transporter is active as a homodimer (*8*). Structural models of the monomeric subunit reveal an antiparallel two-domain structure (*9*, *10*). Each domain consists of five TMSs, and between the fourth and fifth TMS so-called reentrant loops that are believed to be crucial in the translocation mechanism (Figure S1 of the Supporting Information). The N-terminus of the first domain (N domain) is located in the periplasm. A set of *gltS* genetic constructs was engineered that correspond to different evolutionary states in the pathway presented in Figure 1 with the aim of redirecting the pathway to the alternative outcome. First, the protein was taken back one step in evolution by splitting the *gltS* gene in half and constructing two artificial operons containing the genes encoding the N and C domains in both orders [$GltS^{split}$ (Figure S1)]. Next, a set of genes encoding swapped GltS proteins ($GltS^{swap}$) was constructed by fusing the two half-genes in the reverse order.

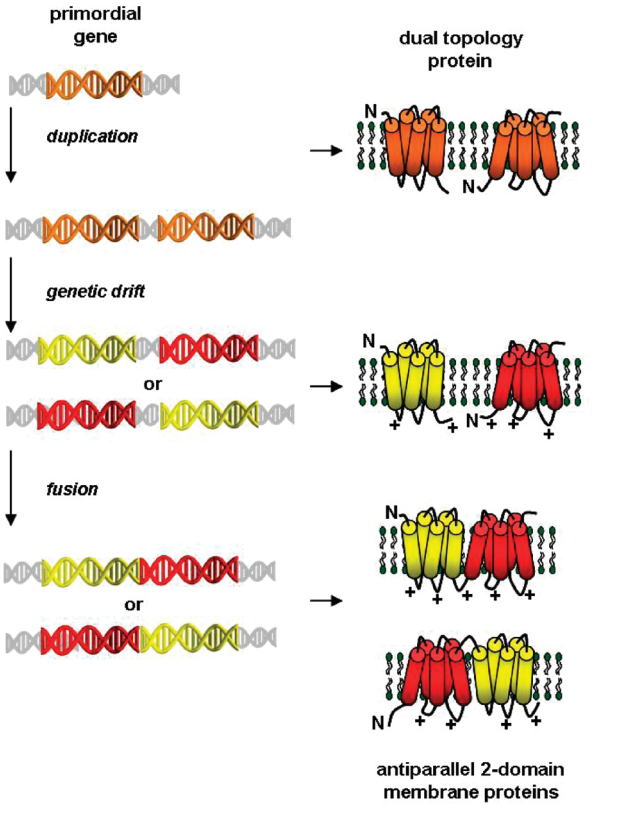

FIGURE 1: Model for the evolution of antiparallel two-domain membrane proteins. The left column shows genetic states. Genes are color-coded as follows: brown, gene encoding a dual-topology protein; yellow, gene encoding a protein with a fixed orientation in the membrane; red, same as yellow but with the opposite orientation. The right column shows encoded proteins embedded in the membrane. Cylinders represent transmembrane segments. N denotes the N-terminus. A plus indicates the position of positively charged amino acid residues in the protein sequence.

[†]The research was funded by the Dutch Organization for Scientific Research (NWO-ALW).

*To whom correspondence should be addressed. Phone: 31503632155. Fax: 31503632154. E-mail: j.s.lolkema@rug.nl.

 Published on Web 06/17/2010

Presumably, GltS^{swap} has the N-terminus of the first domain (C domain) in the cytoplasm.

The half-proteins of GltS^{split} corresponding to the N and C domains were engineered with N-terminal His tags allowing purification by Ni^{2+}-NTA affinity chromatography. Expression of the proteins in *E. coli* DH5α followed by purification from the membrane fraction showed two bands on SDS−PAGE with very similar mobilities (Figure 2A, left). The apparent molecular masses were 17−18 kDa, strongly suggesting that they corresponded to the N domain (201 residues) and C domain (200 residues). Mass spectrometry analysis of the bands confirmed that the upper band corresponded to the N domain and the lower band to the C domain. The order of the two half-genes in the artificial operons did not seem to affect the expression levels significantly.

The split versions of GltS were tested for their ability to accumulate L-[^{14}C]glutamate in right-side-out (RSO) membrane vesicles prepared from the same *E. coli* cells. Glutamate uptake was assessed in the presence of a proton motive force that was generated using the artificial ascorbate/PMS electron donor system. Membrane vesicles prepared from the host cells contain a basal level of glutamate transport activity due to endogenous glutamate transporters encoded on the chromosome [Figure 2B (●)]. RSO membranes containing wild-type GltS exhibited an activity that was approximately 5 times higher than the background activity (◆). Irrespective of the order of the two half-genes in the split *gltS* constructs, RSO membranes prepared from cells coexpressing the two domains retained ∼40% uptake activity of membranes containing the wild-type version of GltS (▲ and ○). It follows that GltS^{split}, expressed from an artificial operon encoding the two domains as separate proteins, is stably expressed and actively accumulates glutamate in RSO membranes.

Three versions of GltS^{swap} (Figure S1) were constructed that differed in the length of the linker that connects the C-terminus of the C domain and the N-terminus of the N domain (Table T1 of the Supporting Information). The linkers consisted of just two serine residues (SS), 12 residues (SSGSGSGSGSGS), or 19 residues (SSGSGSGSGSGSGSGGSGS). All three versions of GltS^{swap} were extended with an N-terminal His tag.

Wild-type GltS purified by Ni^{2+}-NTA affinity chromatography migrates as a single band with an apparent molecular mass of 35 kDa on SDS−PAGE (Figure 2A, right) (*8*). Following purification using the same protocol, no protein product of the expected size was observed for the GltS^{swap} version with the shortest linker of two Ser residues (not shown). In agreement, RSO membrane vesicles derived from cells harboring the construct showed glutamate uptake activity corresponding to the background level (not shown). It follows that this version of GltS^{swap} is not stably assembled in the membrane. In contrast, the GltS^{swap} versions with 12- and 19-residue linkers resulted in clear protein bands on SDS−PAGE with apparent molecular masses slightly larger than that of wild-type GltS, in line with the additional mass introduced by the linkers (Figure 2A, right). Remarkably, only the version with the 19-residue linker was active in Na^{+}-glutamate symport activity in RSO membranes [Figure 2B (■)], while the version with the 12-residue linker resulted in uptake levels similar to those observed for the control cells (□). The activity of the version with the 19-residue linker was comparable to the activity observed for wild-type GltS (◆).

GltS was shown previously to be a dimeric protein (*8*). Treatment of the purified GltS^{swap} variants with the 12- and 19-residue linkers with the unspecific cross-linker glutaraldehyde

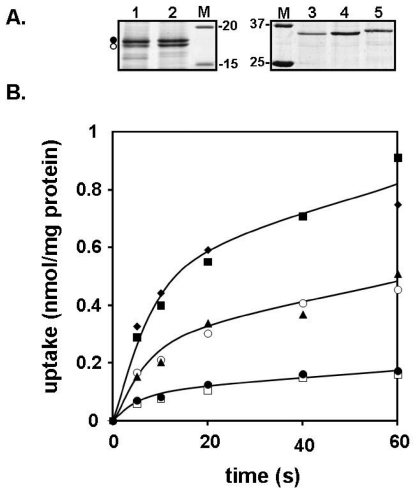

FIGURE 2: Expression (A) and activity (B) of reconstructed evolutionary states of GltS. (A) SDS−PAGE of purified GltS^{split} (left) produced from plasmids with the two half-genes in both orders (lanes 1 and 2) and the full-length GltS variants (right): wild type (lane 3), form swapped with the 12-residue linker (lane 4), and form swapped with the 19-residue linker (lane 5). White and black circles indicate the positions of the N and C domains, respectively. Lanes marked M contained molecular mass standards with the masses in kilodaltons as indicated. The gel was stained with Coomassie Brilliant Blue. (B) Glutamate uptake activity of GltS variants in RSO membranes. L-[^{14}C]Glutamate uptake was assessed in RSO membrane vesicles containing wild-type GltS (◆), the two versions of GltS^{split} (○ and ▲), the two versions of GltS^{swap} with the 12-residue (□) and 19-residue (■) linkers, and control membranes (●).

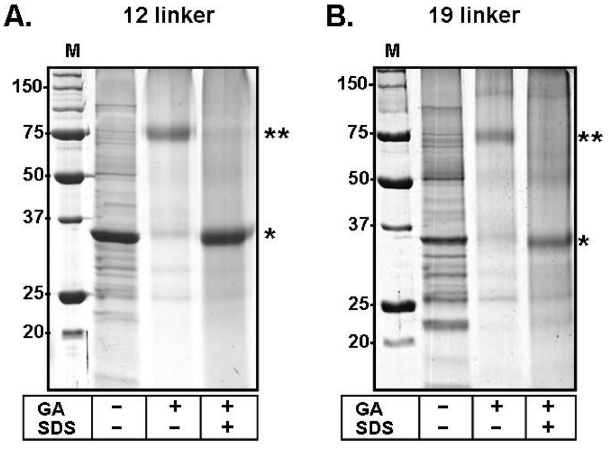

FIGURE 3: Dimeric state of GltS^{swap}. Purified GltS^{swap} variants with 12-residue (A) and 19-residue (B) linkers were treated with glutaraldehyde (GA) at a concentration of 2.5 mM for 20 min in the presence and absence of 0.1% SDS as indicated, followed by SDS−PAGE. One asterisk denotes monomeric GltS, and two asterisks denote dimeric GltS. Lane M contained molecular mass standards as indicated on the left in kilodaltons.

followed by analysis by SDS−PAGE resulted in the disappearance of the 35 and 36 kDa bands, respectively, and at the same time, a new, somewhat fuzzy, band appeared running at approximately double the mass (Figure 3A,B). A small fraction of the protein did not enter the gel to any significant extent, suggesting some aggregation in the protein preparation. Treatment with glutaraldehyde in the presence of SDS prevented cross-linking of the proteins, showing that cross-linking was the result of formation of a complex rather than random collisions. The same fuzziness of the monomeric protein band suggested that it is due to random labeling of the GltS protein molecules with

glutaraldehyde. The results indicate that even though $GltS^{swap}$ with the 12-residue linker was not active in Na^+-coupled glutamate transport, the protein still formed a dimeric complex.

Apparently, there is a threshold distance between the C-terminus of the first domain and the N-terminus of the second domain above which stable assembly in the membrane occurs. A linker of two residues resulted in complete degradation of the polypeptide chain, suggesting conformational stress and misfolding. Twelve residues in the linker would be above the threshold and resulted in a stable dimeric complex in the membrane (Figure 3). However, the complex was not active. Activity was observed when the linker was increased in length to 19 residues. The latter suggests that turnover of the complex involves intersubunit movement and that the 12-linker version is locked in one catalytic state of the protein.

GltS of *E. coli* is a member of the ESS family [glutamate:Na^+ symporter, TC 2.A.27 (*7*)] that is believed to share the antiparallel two-domain structure with 32 other families in structural class ST[3] of the MemGen classification (*11*−*13*). Sequence analysis showed that in all but one family, [st312]NhaC (TC 2.A.35), the N-terminus of the N domain is at the external face of the membrane and the N-terminus of the C domain in the cytoplasm (*11*). The high frequency of one particular order of the domains suggests an evolutionary advantage for this organization, or alternatively, the families originate all from a single ancient duplication event before diversification into the different gene families occurred. Arguing against a single duplication event, an analysis of the DUF606 family, a family of membrane proteins in which different evolutionary states depicted in Figure 1 are found in current members, demonstrated that duplication events within one family are quite frequent (*3*). In this study, we demonstrate that changing the order of the domains in an antiparallel two-domain membrane protein that is at the end of its evolution with respect to the orientation in the membrane does not significantly affect its biogenesis and function. $GltS^{swap}$ is produced by the cells, inserted into the membrane, and is equally active in ion-coupled solute transport. The only condition appears to be that the linker connecting the two domains be of approximately the same length as the linker in the original protein. Evolution toward the same domain order in almost all families sharing this structural organization must be driven by less pronounced features that nevertheless are important on an evolutionary time scale. Possibly, the cellular disposition of the connecting loop in the cytoplasm or exposed to the hazardous external medium may have played a role. At any rate, this study demonstrates that genetic engineering allows for the reconstruction of evolutionary pathways and that evolutionary pathways can be manipulated.

ACKNOWLEDGMENT

We thank Fabrizia Fusetti of The Netherlands Proteomics Centre/Membrane Enzymology group of the University of Groningen for analyzing samples by mass spectrometry.

SUPPORTING INFORMATION AVAILABLE

Details of the molecular cloning, transport assays, and other experimental procedures, Figure S1, and Tables T1 and T2. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Rapp, M., Seppälä, S., Granseth, E., and von Heijne, G. (2007) *Science 315*, 1282–1284.
2. Poolman, B., Geertsma, E., and Slotboom, D.-J. (2007) *Science 315*, 1229–1231.
3. Lolkema, J. S., Dobrowolski, A., and Slotboom, D. J. (2008) *J. Mol. Biol. 378*, 596–606.
4. von Heijne, G. (1986) *EMBO J. 5*, 3021–3027.
5. Kalman, M., Gentry, D. R., and Cashel, M. (1991) *Mol. Gen. Genet. 225*, 379–386.
6. Deguchi, Y., Yamato, I., and Anraku, Y. (1990) *J. Biol. Chem. 265*, 21704–21708.
7. Saier, M. H. (2000) *Microbiol. Mol. Rev. 64*, 354–411.
8. Mościcka, K. B., Krupnik, T., Boekema, E. J., and Lolkema, J. S. (2009) *Biochemistry 48*, 6618–6623.
9. Dobrowolski, A., Sobczak-Elbourne, I., and Lolkema, J. S. (2007) *Biochemistry 46*, 2326–2332.
10. Dobrowolski, A., and Lolkema, J. S. (2009) *Biochemistry 48*, 7448–7456.
11. Lolkema, J. S., and Slotboom, D. J. (2003) *J. Mol. Biol. 327*, 901–909.
12. Lolkema, J. S., and Slotboom, D. J. (2005) *Mol. Membr. Biol. 22*, 177–189.
13. Lolkema, J. S., and Slotboom, D. J. (2008) *Mol. Membr. Biol. 25*, 567–570.