Structure and Function of Amino Acid and Peptide Transport Proteins

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Fabio Casagrande
aus Lugano

Basel, 2008

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Andreas Engel, Prof. Dr. Manuel Palacin
and PD Dr. Dimitrios Fotiadis

Basel, den 24.06.08

Prof. Dr. Hans-Peter Hauri
Dekan

Ich erkläre, dass ich die Dissertation ”Structure and Function of Amino Acid and Peptide Transport Proteins” nur mit der darin angegebenen Hilfe verfasst und bei keiner anderen Universität und keiner anderen Fakultät der Universität Basel eingereicht habe

Fabio Casagrande

Table of Contents

1. General Introduction

All living cells are enclosed by biological membranes that separate the interior cytoplasm from the outer environment. Composed of a phospholipid bilayer with embedded proteins biological membranes act as insulators and filters. The transfer of selected substances and information across the membrane is controlled and mediated through membrane proteins. As a result, membrane proteins are central to almost all cellular processes. They play key roles in signalling between cells, in transport across cell membranes and in energy transduction processes. To accomplish these versatile cell functions, membrane proteins are available in an abundant diversity. Accordingly, it's not astounding that about 30% of all genes encoded in the human genome are membrane proteins. However, only 158 (Status: May 2008) unique membrane protein structures, mainly from bacterial membrane proteins, have been deposited thus far in the Protein Data Bank (PDB), compared to the more than 10'000 soluble protein structures. These numbers highlight the enormous structural work that remains to be done in the field of membrane proteins. From a biomedical perspective, membrane proteins constitute about 50% of possible targets for novel drugs. The three-dimensional (3D) structure of a protein is an essential starting point for the investigation of its molecular mechanisms of action, the basis for drug design. Therefore structural information of membrane proteins is of fundamental importance for human health.

High-resolution structure determination of membrane proteins is currently one of the greatest challenges in cell biology. Membrane proteins possess a hydrophobic belt that is required for their incorporation into lipid membranes. For the extraction and purification of membrane proteins detergents are used to keep the proteins in a solubilized state. The lack of structural information on membrane proteins is mainly related to their low expression levels, the instability in the detergent solution and their resistance to crystallization. The latter is a considerable limitation because X-ray crystallography, which at the present time is the most powerful technique for determining protein structures, requires highly ordered 3D protein crystals. Besides structure determination by X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, the method of electron crystallography on 2D crystals has become increasingly important in membrane protein research. The striking advantage of 2D crystals is that the membrane protein is analyzed in its native environment, the lipid bilayer. However, as with 3D crystals, the production of highly ordered 2D crystals is a major barrier.

Electron crystallography was first applied in 1975 by Henderson and Unwin to study the structure of bacteriorhodopsin in purple membranes (1, 2). Since then, substantial progress has been made in further development of electron crystallography, especially in sample preparation, cryo-transmission electron microscopy (cryo-TEM) imaging and data processing. Thus, improved cryo-TEM in combination with electron diffraction of 2D crystals is used to establish the 3D protein structure (3). The recent structure determination of the mammalian aquaporin-0 (AQP-0) at the remarkable resolution of 1.9 Å demonstrated impressively the potential of the cryo-TEM approach. So far seven atomic models (<4 Å) of membrane proteins have been determined by high-resolution electron crystallography (plant light harvesting complex 2 (4), AQP-1 (5, 6), nicotinic acetylcholine receptor (7, 8), AQP-0 (9, 10), AQP-4 (11), glutathione transferase 1 (2)). Thereby one has to notice that less than two dozen groups are pursuing electron crystallography, compared to the hundreds of groups in X-ray crystallography. Even more

important, 2D crystallization combined with TEM is not an all-or-nothing approach, meaning that also from poorly ordered crystals low resolution structures are obtainable. In conclusion, electron crystallography presents a highly favourable and successful method to explore the structures of membrane proteins (13).

The atomic force microscope (AFM) is a powerful tool to investigate the surface topography of membrane proteins embedded in lipid bilayers under near-physiological conditions, i.e., in buffer solution, at room temperature and under normal pressure. The high lateral (≥5 Å) and vertical resolution (~1 Å), and the high signal-to-noise ratio of the topographs acquired by AFM make this instrument unique to study surface structure and dynamics of single functional membrane proteins.

Besides high-resolution surface imaging, structural and mechanical properties of single membrane proteins can be studied by single molecule force spectroscopy (SMFS), an AFM-related technique. In a typical SMFS experiment the cantilever tip approaches the membrane protein, pushes on it and then retracts. During this approach-retraction cycle the force acting on the molecule is measured and plotted as a function of the tip-surface distance: the so-called 'force curve' is thus obtained. Such force curves reveal details about inter- and intra-molecular interactions, unfolding barriers and energy landscapes in membrane proteins. Because these measurements take place in solution at physiological conditions, the binding of ligands and the subsequent alteration within the protein may be detected and visualized. This offers the unique possibility to directly monitor structural changes related to biological processes.

The structural studies of amino acid and peptide transporters by TEM and AFM reported in this thesis became possible thank to an European initiative called Eugindat (European genomics initiative on disorders of plasma membrane amino acid transporters, http://www.ub.es/eugindat/). In Eugindat biological and medical scientists with various backgrounds were united in one consortium to concentrate the research on amino acid and peptide transporters.

The transport of amino acids into cells is a crucial process for all living species from bacteria to humans. Defects on proteins involved in this transport lead to strong disturbances in the amino acid metabolism of the organism. Humans strongly rely on amino acids in their diet, since nine essential amino acids cannot be synthesized from other precursors. Consequently, it's all the more important that systems for the uptake, distribution and reabsorption of amino acids work properly. Thereby, the proximal tubule plays a central role by reabsorbing over 95% of the filtered amino acid load. In the case where elevated levels (>5%) are detected in urine, the term aminoaciduria is applied. Primary Inherited Aminoacidurias (PIA) is a group of rare diseases arisen from genetic defects in amino acid transporters expressed in the plasma membrane of renal epithelial cells. PIA members are classified by the target amino acid or acids involved. The group includes Cystinuria, Lysinuric Protein Intolerance (LPI), Dicarboxylic Aminoaciduria (DA), Hartnup Disorder (HDis), Iminoglycinuria (IG) and unlabeled aminoacidurias. Cystinuria and LPI are the best studied PIAs. It was demonstrated that members of the heteromeric amino acid transporter family (HAT family) are the molecular base of cystinuria and LPI. HATs are composed of a heavy subunit (HSHAT) and the corresponding light subunit (LSHAT) that are liked together by a disulfide bridge. HSHATs are N-glycosylated type II membrane glycoproteins, whereas LSHATs are nonglycosylated polytopic membrane proteins with twelve putative transmembrane segments (TMS) (14-16). Two genes rBAT (HSHAT) and $b^{0,+}$AT (LSHAT) could be identified responsible for cystinuria (17, 18) while mutations in the system 4F2hc (HSHAT) and y^+LAT1 (LSHAT) lead to LPI (19, 20).

To acquire a thorough knowledge of the structure of relevant transporters for PIA and

renal reabsorption of amino acids, 2D and 3D crystallization of membrane and soluble proteins for structure determination was addressed within Eugindat. Recently, Eugindat members reported the X-ray structure of 4F2hc protein at 2.1 Å (monoclinic) and 2.8 Å (orthorhombic) resolutions (21). In contrast, there are no structures from eukaryotic or human amino acid transporters available. So far, only two structures of bacterial transport proteins with homology to their eukaryotic counterparts were successfully solved: the atomic structures of a bacterial leucine transporter (22) and of a bacterial glutamate transporter (23). Despite the awarded effort, these two transporters do not correspond to the class of LSHAT proteins. However, these examples indicate that important information to understand structure-function relationships of amino acid transporters is gained from prokaryotic homologues. In pursuit of our aim to reveal first structural information of PIA related amino acid transporters, we searched and studied prokaryotic homologues of LSHAT.

The LSHAT membrane proteins $b^{0,+}$AT and y^{+}LAT1 are members of the L-type amino acid transporter family (LAT) that is a subfamily of the amino acid/polyamine/organocation (APC) transporter superfamily. The APC superfamily counts nearby 250 members in prokaryotes and eukaryotes that function as solute cation symporters and solute cation antiporters (17). Most APC members are predicted to possess twelve α-helical transmembrane segments (TMS) with cytosolic located N- and C-termini (24, 25). According to the high sequence identity and homology to eukaryotic and human APC transporters, we selected after an exhaustive search for structural (TEM/AFM) and functional studies the two prokaryotic amino acid transporters AdiC and SteT. As shown in the presented thesis, the L-arginine/agmatine antiporter AdiC and the threonine/serine exchange transporter SteT represent excellent proteins to elucidate the molecular architecture of transporters from the APC superfamily.

Within Eugindat, structural and functional studies were extended to cover other important transporters involved in human health, i.e. peptide transporters. Members of this second class of transport proteins were extensively studied in the past and belong to the peptide transporter (PTR) family (26, 27). Peptide transporters are integral membrane proteins that mediate the cellular uptake of di- and tripeptides. Similar to the amino acid transport, peptide transport is of fundamental importance in all species. In human peptide transport at the brush border membranes of small intestine, kidney and lung is handled by two members from the PTR family designated as PEPT1 and PEPT2. PEPT1 is considered as the major route by which protein digestion products enter the body. Previous studies demonstrated that peptide transporters have broad substrate specificity transporting essentially all 400 possible dipetides, 8000 possible tripeptides and a large spectrum of peptidomimetics into the cell (i.e. pharmacologically active compounds) (28-31). Peptide transporters are therefore potent drug delivery systems. Substrate translocation is coupled to the proton movement down an electrochemical proton gradient with the membrane potential as the main driving force.

In contrast to the wealth of functional information, no structural information on peptide transporters is available. For our structural and functional studies within Eugindat, we selected from several possible candidates the bacterial PTR family members YbgH, YdgR (TppB, DtpA) and YhiP (DtpB) from *Escherichia coli*. The structural work on these bacterial peptide transporters presented in this thesis represents the first published structural information of these important class of transport proteins.

1.1 REFERENCES

1. Unwin PN. and Henderson R. (1975) *J Mol Biol* **94** 425-40.
2. Henderson R. and Unwin PN. (1975) *Nature* **257** 28-32.
3. Stahlberg H., Fotiadis D., Scheuring S., Rémigy H., Braun T., Mitsuoka K., Fujiyoshi Y., Engel A. (2001) *FEBS Lett* **504** 166-72
4. Kuhlbrandt W., Wang DN. and Fujiyoshi Y. (1994) *Nature* **367** 614-621
5. Murata K., Mitsuoka K., Hirai T., Walz T., Agre P., Heymann JB., Engel A. and Fujiyoshi Y. (2000) *Nature* **407** 599-605
6. Ren G., Reddy VS., Cheng A., Melnyk P. and Mitra AK. (2001) *Proc Natl Acad Sci USA* **98** 1398-1403
7. Miyazawa A., Fujiyoshi Y. and Unwin N. (2003) *Nature* **423** 949-955
8. Unwin N. (2005) *J Mol Biol* **346** 967-989
9. Gonen T., Sliz P., Kistler J., Cheng Y. and Walz T. (2004) *Nature* **429** 193-197
10. Gonen T., Cheng Y., Sliz P., Hiroaki Y., Fujiyoshi Y., Harrison SC. and Walz T. (2005) Nature **438** 633-638
11. Hiroaki Y., Kazutoshi T., Kamegawa A., Gyobo N., Nishikawa K., Suzuki H., Walz T., Sasaki S., Mitsuoka K. and Kimura K. et al. (2006) *J Mol Biol* **355** 628-639.
12. Holm PJ., Bhakat P., Jegerschold C., Gyobu N., Mitsuoka K., Fujiyoshi Y., Morgenstern R. and Hebert H. (2006) *J Mol Biol* **360** 934-945
13. Hite RK., Raunser S. and Walz T. (2007) Curr Opin Struct Biol **17** 389-95
14. Palacin M. and Kanai Y. (2004) *Pflugers Arch* **447** 490-494
15. Verrey F., Closs EI., Wagner CA., Palacin M., Endou H. and Kanai Y. (2004) *Pflugers Arch* **447** 532-542
16. Palacin M., Nunes V., Font-Llitjos M., Jimenez-Vidal M., Fort J., Gasol E., Pineda M., Feliubadalo L., Chillaron J. and Zorzano A. (2005) *Physiology (Bethesda)* **20** 112-124
17. Calonge MJ., Gasparini P., Chillaron J., Chillon M., Gallucci M., Rousaud F., Zelante L., Testar X., Dallapiccola B., Di Silverio F., Barcelo P., Estivill X., Zorzano A., Nunes V. and Palacin M. (1994) *Nat Genet* **6** 420-425
18. Feliubadalo L., Font M., Purroy J., Rousaud F., Estivill X., Nunes V., Golomb E., Centola M., Aksentijevich I., Kreiss Y., Goldman B., Pras M., Kastner DL., Pras E., Gasparini P., Bisceglia L., Beccia E., Gallucci M. and *et al.* (1999) *Nat Genet* **23** 52-57
19. Torrents D., Mykkanen J., Pineda M., Feliubadalo L., Estevez R., de Cid R., Sanjurjo P., Zorzano A., Nunes V., Huoponen K., Reinikainen A., Simell O., Savontaus ML., Aula P. and Palacin M. (1999) *Nat Genet* **21** 293-296;
20. Borsani G., Bassi MT., Sperandeo MP., De Grandi A., Buoninconti A., Riboni M., Manzoni M., Incerti B., Pepe A., Andria G., Ballabio A., and Sebastio G. (1999) *Nat Genet* **21** 297-301
21. Fort J., de la Ballina LR., Burghardt HE., Ferrer-Costa C., Turnay J., Ferrer-Orta C., Usón I., Zorzano A., Fernández-Recio J., Orozco M., Lizarbe MA., Fita I. and Palacín M. (2007) *J Biol Chem* **282** 31444-52
22. Yamashita A., Singh SK., Kawate T., Jin Y. and Gouaux E. (2005) *Nature* **437** 203-5
23. Yernool D., Boudker O., Jin Y. and Gouaux E. (2004) *Nature* **431** 811-8
24. Jack DL., Paulsen IT. and Saier MH. (2000) *Microbiology* **146** 1797-814
25. Gasol E., Jimenez-Vidal M., Chillaron J., Zorzano A. and Palacin M. (2004) *J Biol Chem* **279** 31228-36
26. Paulsen IT. and Skurray RA. (1994) *Trends Biochem Sci* **18** 404
27. Steiner HY., Naider F. and Becker JM. (1995) *Mol Microbiol* **16** 825-34
28. Brandsch M., Knutter I. and Leibach FH. (2004) *Eur J Pharm Sci* **21** 53-60
29. Daniel H. (2004) *Annu Rev Physiol* **66** 361-384
30. Daniel H. and Kottra G. (2004) *Pflugers Arch* **447** 610-618
31. Terada T. and Inui K. (2004) *Curr Drug Metab* **5** 85-94

2. Projection Structure of a Member of the Amino Acid/Polyamine/Organocation Transporter Superfamily

Fabio Casagrande[1], Merce Ratera[2], Mohamed Chami[1], Andreas Schenk[1], Eva Valencia[2], Jesus Maria Lopez[3], David Torrents[3], Andreas Engel[1], Manuel Palacin[2] and Dimitrios Fotiadis[4]

[1] M. E. Müller Institute for Structural Biology, Biozentrum of the University of Basel, CH-4056 Basel, Switzerland

[2] Institute for Research in Biomedicine, Barcelona Science Park, Department of Biochemistry and Molecular Biology, Faculty of Biology, University of Barcelona and Centro de Investigacion Biomedica en Red de Enfermedades Raras, E-08028 Barcelona, Spain

[3] Barcelona Supercomputing Center, Life Science Program-Institucio Catalana de Recerca i Estudis Avancats, E-08034 Barcelona, Spain

[4] Institute for Biochemistry and Molecular Medicine of the University of Berne, CH-3012 Berne, Switzerland

2.1 ABBREVIATIONS

APC, amino acid/ polyamine/ organocation; **BN-PAGE**, blue native-polyacrylamide gel electrophoresis; **DDM**, n-dodecyl-β-D-maltoside; **gi**, GenInfo; **ITC**, isothermal titration calorimetry; **MFS**, major facilitator superfamily; **SDS-PAGE**, sodium dodecylsulphate-polyacrylamide gel electrophoresis; **TEM**, transmission electron microscopy

2.2 ABSTRACT

The L-arginine/agmatine antiporter AdiC is a key component of the arginine-dependent extreme acid resistance system of Escherichia coli. Phylogenetic analysis indicated that AdiC belongs to the amino acid/polyamine/ organocation (APC) transporter superfamily having sequence identities of 15-17% to eukaryotic and human APC transporters. For functional and structural characterization, we cloned, overexpressed and purified wild-type AdiC and the point mutant AdiC-W293L. Detergent-solubilized AdiC particles were dimeric in agreement with previous observations. Reconstitution experiments yielded 2D crystals of AdiC-W293L diffracting below 7 Å resolution from which we determined the projection structure at 6.5 Å resolution. The projection map showed 10-12 density peaks per monomer and suggested mainly tilted helices with the exception of one perpendicular membrane spanning α-helix. Comparison of AdiC-W293L with the projection map of the oxalate/formate antiporter from Oxalobacter formigenes, a member from the major facilitator superfamily, indicated different structures and consequently different folds. Thus, 2D crystals of AdiC-W293L yielded the first detailed view of a transport protein from the APC superfamily at sub-nanometer resolution.

2.3 INTRODUCTION

Enteric pathogens such as *Shigella*, *Salmonella*, *Yersinia* spp. and certain *Escherichia coli* (*E. coli*) strains can survive the extremely acidic conditions of the human stomach and cause intestinal diseases (1). To overcome the protective barrier of the gastric acidity, pathogenic and nonpathogenic strains of *E. coli* have developed acid resistance systems. One of these systems requires arginine to protect *E. coli* during low pH exposure. This arginine system is composed of an arginine-agmatine exchange transporter and of an acid-activated arginine decarboxylase (2). Acidification of the cytosol is prevented by the consumption of protons through decarboxylation of arginine to agmatine and carbondioxide. Agmatine is then exported out of the cytosol and new arginine imported through the arginine-agmatine transporter in a one-to-one exchange stoichiometry (2). This recently identified transport protein is the product of the *adiC* gene (3, 4). *In vitro*, AdiC-mediated exchange transport of arginine and agmatine is tightly coupled, electrogenic and acid-activated (5). AdiC forms stable homodimers in detergent and phospholipid membranes as determined by gel filtration and glutaraldehyde cross-linking experiments (5).

The origin of AdiC is somehow controversial since it was assigned to two families of transport proteins, i.e. the amino acid/polyamine/ organocation (APC) transporter superfamily (6) and the major facilitator superfamily (MFS) (5, 7). The APC superfamily of transporters consists of nearly 250 members that function as solute-cation symporters and solute-solute antiporters (6). According to hydropathy profile analysis and biochemically established topological features of most prokaryotic and eukaryotic APC superfamily members, both the N- and C- termini of the proteins are located in the cytoplasm with a twelve transmembrane segment topology (6, 8-11). Two subfamilies of the APC transporter superfamily, i.e., the cationic amino acid transporters (CATs) and the L-type amino acid transporters (LATs), have human members involved in relevant physiological functions and diseases. CAT2 provides arginine for NO synthesis (12) and for arginase (13) in classical and alternative activation of macrophages respectively, and CAT1 is required for macrophage proliferation (14). LAT1 is overexpressed in a wide spectrum of primary human cancers and plays important roles in the growth and survival of cancer cell lines (15). Mutations in the human LAT members $b^{0,+}$AT (and in its associated heavy subunit rBAT) and y^{+}LAT1 cause the primary inherited aminoacidurias cystinuria and lysinuric protein

intolerance, respectively (16-19). Another LAT transporter, xCT, is a Kaposi's sarcoma-associated herpesvirus fusion-entry receptor (20).

Structural information on amino acid transporters is sparse. Only two high-resolution structures have been reported so far, that of the bacterial glutamate transporter Glt_{ph} (21) and that of the bacterial leucine transporter $LeuT_{Aa}$ (22). Thus, besides low-resolution transmission electron microscopy (TEM) data on single detergent-solubilized Ser/Thr exchanger transport proteins (SteT) (23), no structural information is available for other members of the APC transporter superfamily. The low number of amino acid transporter structures and of membrane proteins in general is related to the difficulty in growing highly-ordered three-dimensional (3D) membrane protein crystals for structure determination by X-ray crystallography. Cryo-TEM combined with electron crystallography of 2D crystals to establish the 3D protein structure is an attractive alternative to X-ray crystallography and 3D crystals (24). An important advantage of 2D membrane protein crystals is that the protein is embedded in its native environment, the lipid bilayer. Recently, the first 2D crystals of an amino acid transporter, i.e. Glt_{ph}, have been reported (25).

Here, we present for the first time structural data at sub-nanometer resolution of the L-arginine/agmatine antiporter AdiC. Phylogenetic analysis and comparison of the projection structure of AdiC, as obtained by cryo-TEM of 2D crystals, with the projection and 3D structures of members from the MFS indicated that AdiC belongs to the APC superfamily. To grow 2D crystals, we overexpressed AdiC and the mutant AdiC-W293L in *E. coli* and purified to homogeneity. Negative-stain TEM and blue native-polyacrylamide gel electrophoresis (BN-PAGE) of both detergent-solubilized AdiC proteins revealed a homodimeric state. Reconstitution experiments yielded first well-ordered 2D crystals of AdiC-W293L suitable for structure analysis by cryo-TEM. The projection map calculated from unstained 2D crystals revealed the structure of AdiC-W293L at 6.5 Å resolution. The present work sets the basis for the structural analysis of AdiC at high-resolution using 2D crystals and cryo-TEM/electron crystallography. Importantly, AdiC represents an excellent model for understanding the molecular architecture of transporters from the APC superfamily.

2.4 MATERIALS AND METHODS

Chemicals – *E. coli* polar lipids were purchased from Avanti Polar Lipids (Alabaster, AL, USA) and n-dodecyl-β-D-maltopyranoside from Anatrace (Maumee, OH, USA). Radiolabeled L-[^{3}H] arginine was from American Radiolabeled Chemicals (St. Louis, MO, USA), and amino acids and amines from Sigma-Aldrich (St. Louis, MO, USA).

Phylogenetic analysis of AdiC – Comparative searches of the AdiC gene (NCBI's GenInfo (gi): 16131941) were performed using NCBI's BLASTp algorithm (www.ncbi.nlm.nih.gov/blast/) against the default nr protein database (GenBank CDS translations + RefSeq Proteins + PDB + SwissProt + PIR + PRF) and using default parameters. To build the phylogenetic tree, we first performed a multiple alignment using PROBCONS (26) with the AdiC protein sequence and other protein sequences collected from NCBI and http://www.tcdb.org/tcdb/superfamily.php (6) that are representatives of the following APC subfamilies: APA (PotE/gi: 26246666 from *E. coli*, CadB/gi: 536977 from *E. coli*, ArcD/gi: 254827 from *Pseudomonas aeruginosa*, LysI/gi: 38233510 from *Corynebacterium glutamicum* and YvsH/gi: 16080387 from *Bacillus subtilis*), ACT (GabA/gi: 4972245 from *Emericella nidulans* and Uga4/gi: 6319991 from *Saccharomyces cerevisiae*), CAT (CAT1/gi: 161016790 from *Mus musculus*, CAT2/gi: 113680130 from *Mus musculus*, AAT1/gi: 30685317 from *Arabidopsis thaliana* and YfnA/gi: 16077801 from *Bacillus subtilis*), GGA (XasA/gi: 33241625 from *Chlamydophila pneumoniae*),

AGT (yveA/gi: 16080500 from *Bacillus subtilis*, ybeC/gi: 16077281 from *Bacillus subtilis*) and LAT (SteT/gi: 16078351 from *Bacillus subtilis*, and eukaryotic LATs: LAT2/gi: 33286428 from *Homo sapiens*, LAT1/gi: 71979932 from *Homo sapiens*, y^+LAT-1/gi: 21361563 from *Homo sapiens*, and $b^{0,+}$AT from *Monodelphis domestica* (ENMSEBL:ENSMODP00000015349), from *Canis familiaris* (ENMSEBL: ENSCAFP0000001-1140) and from *Rattus norvegicus* (ENMSEBL: ENSRNOP00000016-919). As outgroup for the tree, we selected Mmup/gi: 172051482 of *E .coli* that belongs to the AAT family. We next evaluated this alignment and obtained the tree using the Bayesian phylogenetic method (GTR model, 5 million generations and excluding 40% of the initial trees; (27)). The tree was elaborated and prepared for proper display with the iTOL online tool (http://itol.embl.de/) (28).

Cloning, mutagenesis and overexpression of AdiC and AdiC-W293L – AdiC was cloned from genomic *E. coli* strain DH5α DNA as previously described for other transporters (23). Briefly, the following primers (5'-3') were used to amplify the AdiC open reading frame by PCR: ATGAGAATTCACTATCGAATGCAGG-CGTATG and TCATCTCGAGGTTGGCTTTT-ATGTTTGCTGGA. PCR products were digested with EcoRI and XhoI and ligated into a pBlueScript vector (Stratagene, La Jolla, CA, USA). Another PCR with the following primers (5'-3') was performed to subclone the AdiC open reading frame into the EcoRI and PstI sites of a modified version of the vector pTTQ18 (29), thereby placing its expression under the control of the *tac* promoter: TGATGAATTCGATG-TCTTCGGATGCTGATGC and ACCGCCT-GCAGAATCTTTGCTTATTGGTGCA. Mutation W293L in AdiC was introduced in AdiC cloned into pTTQ18 using QuickChange site-directed mutagenesis kit (Stratagene) with primer 5'GCTTAGGTTCACTGGGCGGC-T[T]GACGTTGCTGGCGGGTC3' and its complementary reverse primer. For ITC studies AdiC was amplified from the pBlueScript construct using as primers (5'-3') CAAT-GGATCCATGTCTTCGGATGCTGATGCTC and TCATCTCGAGGTTGGCT-TTTATGTTT-GCTGGA and subcloned into BamHI and XhoI sites of pTrcHisA vector. All DNA constructs were verified by sequencing. Expression experiments were carried out with freshly transformed *E. coli* BL21(DE3) harbouring pTTQ18-His_6-AdiC, pTTQ18-His_6-AdiC-W293L or pTrcHisA-AdiC. In cultures grown in LB medium, optimal expression of the AdiC proteins occurred after 3 h at 37 ºC after induction at A_{600} of 0.5 with 0.5 mM isopropyl-β-D-thiogalactoside. Routinely, four to six liters of medium in flasks were inoculated with the corresponding AdiC construct and after 3 h of growth at 37 ºC, cells were harvested. Cell pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 0.5 mM EDTA).

Preparation of E. coli membranes – Cell pellets were disrupted by passage through a French pressure cell (20,000 psi). Unbroken cells were removed by centrifugation (10 min at 10,000 *g*, 4 ºC). The supernatant was ultracentrifuged (1 h at 100,000 *g*, 4 ºC), and the pellet was resuspended and homogenized (30 ml glass homogenizer) in lysis buffer and ultracentrifuged again. Peripheral membrane proteins were removed by homogenization in 20 mM Tris-HCl pH 8.0, 300 mM NaCl and ultracentrifugation. Finally, the membrane pellet was resuspended in 20 mM Tris-HCl pH 8.0, 150 mM NaCl at a protein concentration between 13 and 25 mg/ml. Aliquots were frozen in liquid nitrogen and stored at -80 ºC until use.

Purification of AdiC and AdiC-W293L, and determination of protein concentration – Frozen *E .coli* total membranes containing overexpressed AdiC or AdiC-W293L were thawed and solubilized for 2 hours at 4 °C under gentle agitation in 1% DDM, 20 mM Tris-HCl pH 8, 300 mM NaCl, 10% glycerol, 0.01% NaN_3. The protein concentration during solubilization was between 2-3 mg/ml. After ultracentrifugation (100,000 *g*, 50 min at 4°C) the supernatant was diluted two-fold with 20 mM Tris-HCl pH 8,

300 mM NaCl, 0.04% DDM, 5 mM histidine, 10% glycerol, 0.01% NaN_3 (washing buffer) and bound for 2 h at 4 °C to Ni-NTA Superflow beads (Qiagen, Hilden, Germany). The beads were then loaded onto a spin column (Promega, Madison, WN, USA), washed with washing buffer and eluted with the same buffer containing 200 mM histidine. For ITC studies the beads were washed twice with 20 mM imidazole in 20 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.01% DDM. AdiC-W293L was eluted with 300 mM imidazole in the same buffer. Untagged AdiC was eluted by overnight digestion at room temperature with enterokinase (Invitrogen, Carlsbad, CA, USA) (5 U/ mg protein) in 20 mM Tris-HCl pH 7.6, 10 mM $CaCl_2$ and 0.01% DDM. Then, protein was further purified by size-exclusion chromatography with Superdex 200 (10/300) (Pharmacia) and concentrated with Vivaspin MWCO100 (Sigma-Aldrich, St. Louis, MO, USA). Protein concentration was determined spectrophotometrically measuring the absorbance at 280 nm and using a molar extinction coefficient of 85,830 M^{-1} cm^{-1} (AdiC), 91,830 M^{-1} cm^{-1} (untagged AdiC) and of 80,330 M^{-1} cm^{-1} (AdiC-W293L). These values were calculated from the amino acid sequence of the different AdiC versions using the ProtParam tool from the ExPASy proteomics server (http://www.expasy.org/tools/protparam.html).

Transport measurements – For functional studies, purified AdiC proteins were reconstituted into proteoliposomes as previously described (23). L-[^{3}H] arginine influx measurements into AdiC and AdiC-W293L proteoliposomes were performed according to Reig *et al*, 2007. All experimental values were corrected by subtracting zero time values and are reported as means ± S.E.M. (standard error of the mean).

Isothermal titration calorimetry (ITC) – All solutions were thoroughly degassed before titration in an ITC instrument (MicroCal). A solution of AdiC or AdiC-W293L (0.11 mM in 1.4 ml of buffer 150 mM NaCl, 20mM Tris-HCl, pH 7.5, 0.01% DDM) was loaded into the sample cell, 2.5 mM titrant (L-arginine in the same buffer) was loaded into the injection syringe, and the system was equilibrated at 25 °C. Titration curves were generated by 30 successive 10 µl injections at 240 sec intervals. Control injections were done in the absence of protein to determine background corrections. The data were fit to single-site binding isotherms with the Origin software.

2D crystallization of AdiC-W293L – Purified AdiC-W293L protein solubilized in DDM was mixed with *E. coli* polar lipids solubilized in the same detergent (stock solution: 5 mg/ml *E. coli* polar lipids , 1% DDM, 20 mM Tris-HCl pH 8, 150 mM NaCl, 10% glycerol, 0.01% NaN_3) to achieve a lipid-to-protein ratio (LPR) of 0.3 (w/w). The final AdiC/AdiC-W293L protein concentration was adjusted to about 1 mg/ml. The protein/lipid/detergent mix was dialyzed against detergent-free buffer, i.e. 20 mM citric acid NaOH pH 5, 250 mM NaCl, 10% glycerol, 0.01% NaN_3, 5 mM arginine for about two weeks at room temperature.

Blue native gel electrophoresis – Linear 5–12% gradient gels for BN-PAGE were prepared and run as previously described (30). Thyroglobulin (669 kDa), ferritin (440 kDa), lactate dehydrogenase (140 kDa), and bovine serum albumin (66 kDa) were used as standard proteins.

Negative stain TEM – DDM-solubilized AdiC or AdiC-W293L protein at ~3 µg/ml was adsorbed for ~10 sec to parlodion carbon-coated copper grids rendered hydrophilic by glow discharge at low pressure in air. Grids were washed with three drops of double-distilled water and stained with 2 drops of 0.75% uranyl formate. Grids of negatively stained 2D AdiC-W293L crystals (Figure 5) were prepared similarly, but included a longer adsorption time of ~60 sec. Electron micrographs were recorded at a magnification of 50,000x on Eastman Kodak Co. SO-163 sheet films with a Hitachi H-7000 electron microscope operated at 100 kV.

Cryo-TEM – 2D crystals of AdiC-W293L were

adsorbed to a thin carbon film on a molybdenum grid and frozen in liquid nitrogen in the presence of 7% (w/v) trehalose (31). Frozen grids were transferred onto a Philips CM200-FEG electron microscope using a Gatan 626 cryo-holder. Electron micrographs were recorded on Eastman Kodak Co. SO-163 films at an accelerating voltage of 200 kV, a magnification of 50,000x using a low-dose system (~10 electrons per $Å^2$) and a measured specimen temperature of about -175 °C.

Image processing – Electron micrographs of selected AdiC-W293L crystals were scanned using a Heidelberg Primescan D7100 scanner with a sampling of 1 Å/pixel at the sample level. The images were processed using the IPLT image processing toolkit (32, 33) to correct for the contrast transfer function (CTF) and crystal distortions. The parameters of the CTF and the crystal lattice were fitted manually using the graphical user interface provided by IPLT. The five best images were scaled using one image as reference and merged to produce the final data set from which the projection represented as a contour map was generated within IPLT. To generate the improved projection map of AdiC-W293L (Figure 6B), the projection of one downside-up oriented dimer was flipped, translationally and rotationally aligned with respect to the projection of one upside-down oriented dimer, averaged and two-fold symmetrized exploiting the internal symmetry of the AdiC-W293L dimer (2D crystals consist of alternating rows of downside-up and upside-down oriented dimers in the lipid bilayer).

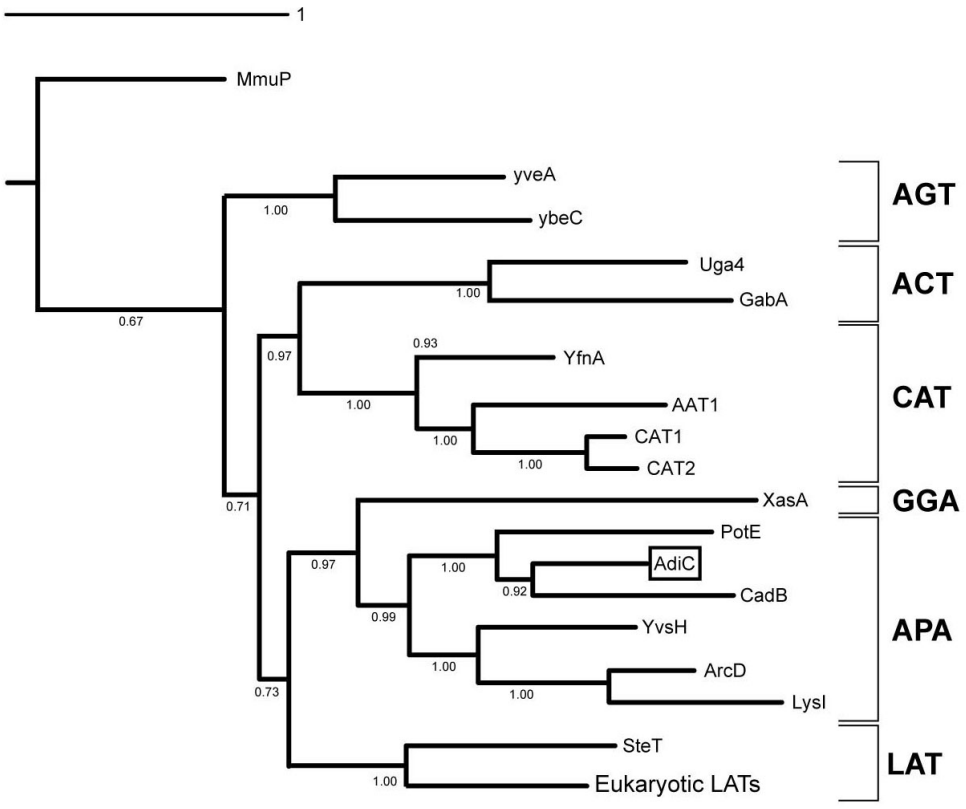

Figure 1. Origin and evolutionary relationship of AdiC with other APC transport proteins. The Bayesian tree shows the position of the AdiC protein sequence (within a box) in the context of other members of the APC transporters superfamily. The abbreviation of each of the subfamilies is on the right side of the tree: see Jack *et al*, 2000 and http://www.tcdb.org/tcdb/superfamily.php for a description of each of the subfamilies displayed. The tree was rooted using the MmuP protein of *E. coli*. Probabilities given by Bayesian analysis are displayed at each of the tree nodes. The bar shows the evolutionary distance in number of expected substitutions per site.

2.5 RESULTS

Phylogenetic analysis of AdiC – The origin and evolution of the AdiC protein appears to be controversial, as it has been reported to belong to two distant groups of membrane transporters: MFS (5, 7) and the APC superfamily (6). To clarify this issue, we have analyzed *de novo* the phylogenetic relationship of AdiC with known transport proteins. We have combined two rounds of computational searches and comparisons with different levels of sensitivities and accuracies (see materials and methods). As an initial and rough approximation, we first used BLAST to compare the AdiC protein sequence against a non-redundant protein sequence database and observed that closest matches (other than predicted, potential or hypothetical) were prokaryotic amino acid/polyamine antiporters, e.g. the proteins with the gi's: 152969070, 44886079, 10801580, 123443711 (PotE) at the National Center for Biotechnology Information (NCBI) data base. These transporters belong to the basic amino acid/polyamine antiporter (APA) family, a subgroup of the APC superfamily (6, 23). Next, and similarly to our previous work with SteT (23), we selected a number of protein sequences representing APC subfamilies that are close to APA, and calculated a Bayesian tree. As seen in Figure 1, the tree indicated that AdiC clearly clusters with members of the APA family. In addition, the phylogenetic analysis revealed that eukaryotic and human APC transporters closest related to AdiC are members from the L-amino acid transporter (LAT) family, e.g. xCT, y^+LAT1 and $b^{0,+}$AT, with sequence identities of 15-17%.

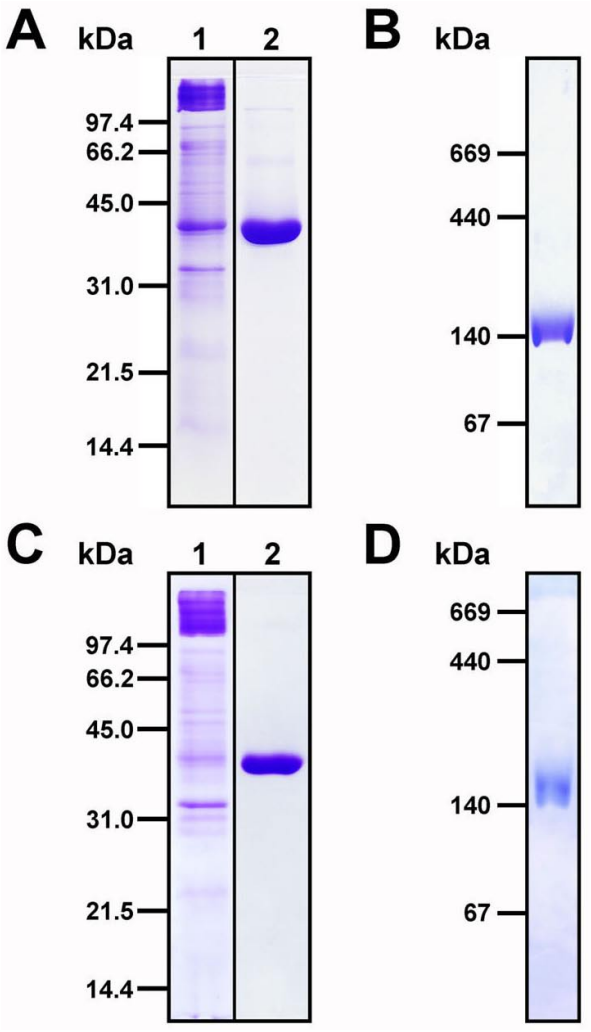

Figure 2. SDS- and BN-PAGE of AdiC and AdiC-W293L. (A) 13.5% SDS-PAGE of DDM-solubilized E. coli total membranes containing overexpressed AdiC protein (lane 1) and of purified protein after nickel affinity chromatography (lane 2). AdiC runs as a prominent band at ~39 kDa in SDS/polyacrylamide gels. (B) BN-PAGE of purified AdiC protein in a linear 5–12% gradient gel. (C) Same as in (A) but for AdiC-W293L. (D) Same as in (B) but for AdiC-W293L. All gels were stained with Coomassie Brilliant Blue R-250. Applied amount of protein per lane: ~9 µg (lane 1, panel (A)), ~6 µg (lane 2, panel (A)), ~12 µg (B), ~8 µg (lane 1, panel (C)), ~5 µg (lane 2, panel (C)) and ~5 µg (D).

Cloning, expression and purification of AdiC and AdiC-W293L – The AdiC gene from *E. coli* was amplified from genomic DNA by PCR using gene-specific primers. The W293L point mutation in transmembrane 8 was introduced in order to diminish the number of possible conformations of AdiC to foster crystallization. The corresponding substitution (W292L) in the related APA transporter PotE (putrescine/ornithine exchanger from *E. coli*) results in an inactive transporter (34). For homologous overexpression the AdiC and AdiC-W293L genes were cloned into the pTTQ18 vector fused with a C-terminal His_6-tag. Isolated *E. coli* total membranes containing overexpressed AdiC or AdiC-W293L were solubilized with n-dodecyl-β-D-maltoside (DDM), and the corresponding protein purified in a single step by nickel affinity chromatography. The high purity of the isolated

proteins is reflected in the SDS/polyacrylamide gels: AdiC (Figure 2A, lane 2) and AdiC-W293L (Figure 2C, lane 2). Both proteins run at ~39 kDa by SDS-PAGE. This was further confirmed by Western blot analysis using an anti-His antibody (data not shown). Because of the high expression levels, both AdiC forms were also visible in the DDM-solubilized *E. coli* total membranes: AdiC (Figure 2A, lane 1) and AdiC-W293L (Figure 2C, lane 1), see band at ~39 kDa. Typically, about 3 mg and 1 mg of pure AdiC and AdiC-W293L protein, respectively, were purified from 1 liter of bacterial cell culture.

Characterization of purified AdiC and AdiC-W293L by BN-PAGE and negative-stain TEM – To determine whether AdiC exists in a monomeric or oligomeric state, detergent-solubilized protein was subjected to BN-PAGE and negative stain TEM. On BN-gels, purified AdiC (Figure 2B) and AdiC-W293L (Figure 2D) migrated as single bands between 140 and 160 kDa depending on the experimental conditions. Denaturation of AdiC with SDS prior BN-PAGE resulted in an additional band at ~70 kDa (data not shown). To assess the shape, dimensions and low-resolution structure of purified AdiC and AdiC-W293L, proteins were adsorbed on a parlodion carbon-coated grid, washed, negatively stained and examined by TEM. Figure 3A shows the high homogeneity of AdiC after purification by nickel affinity chromatography. Single AdiC particles were discerned and displayed an elliptical shape with a major axis of 121±7 Å and a minor axis of 76±6 Å (n=100). The elongated particles had a dimeric appearance comparable to two connected ring-like structures. The central, stain-filled indentation in each ring was clearly visible on AdiC top views (Figure 3B, gallery). As documented in Figure 3C, AdiC-W293L had similar dimensions and appearance as AdiC.

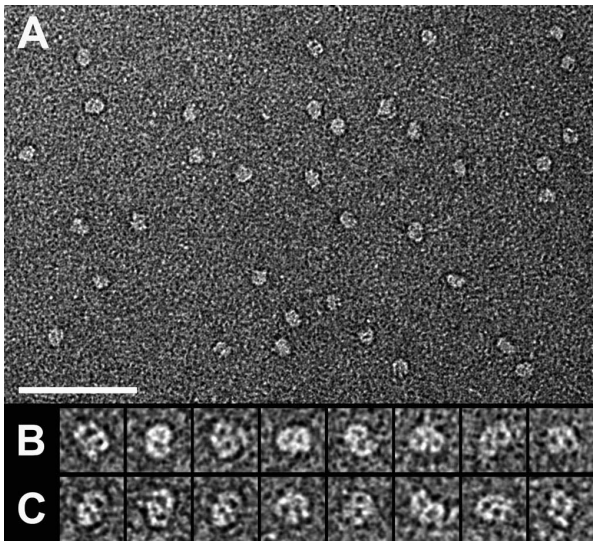

Figure 3. TEM of negatively stained AdiC and AdiC-W293L particles. (A) The homogeneity of the purified AdiC protein is reflected in the electron micrograph. Selected top views of AdiC (B) and AdiC-W293L particles (C) are displayed in the corresponding gallery. The scale bar represents 750 Å, and the frame size of the magnified particles in the galleries is 190 Å.

Transport, inhibition pattern and binding in AdiC and AdiC-W293L – Transport of L-arginine via AdiC reconstituted in proteoliposomes is characteristic of an obligatory exchanger. Thus, transport is dependent on the presence of substrate (L-arginine) inside AdiC proteoliposomes (Figure 4A). In the absence of substrate inside AdiC proteoliposomes influx of L-[^{3}H] arginine most probably occurs by simple diffusion (Figure 4A). Thus, this transport is very low and similar to that of L-serine with SteT (the LAT exchanger of Ser/Thr from *B. subtilis*) proteoliposomes under similar conditions (23). The exchange L-[^{3}H] arginine/L-arginine has a larger flux that the exchange L-[^{3}H] arginine/agmatine (Figure 4B) in agreement with previous results (5). The inhibition profile of L-[^{3}H] arginine/arginine exchange by 5 mM substrate analogs suggest the following affinity pattern: L-arginine, agmatine > cadaverine > putrescine > L-lysine, L-ornithine, L-glutamate, 1,3-diaminopropane > L-glutamine, aminoguanidine (Figure 4C). This is in agreement with the previously reported transport competence profile for reconstituted AdiC: L-arginine > cadaverine > L-lysine, L-ornithine (5). Moreover, our results indicate that substitution of the agmatine guanidine group by amine is recognized by AdiC depending on the size of the aliphatic chain (Figure 4C; cadaverine > putrescine > 1,3-diaminopropane).

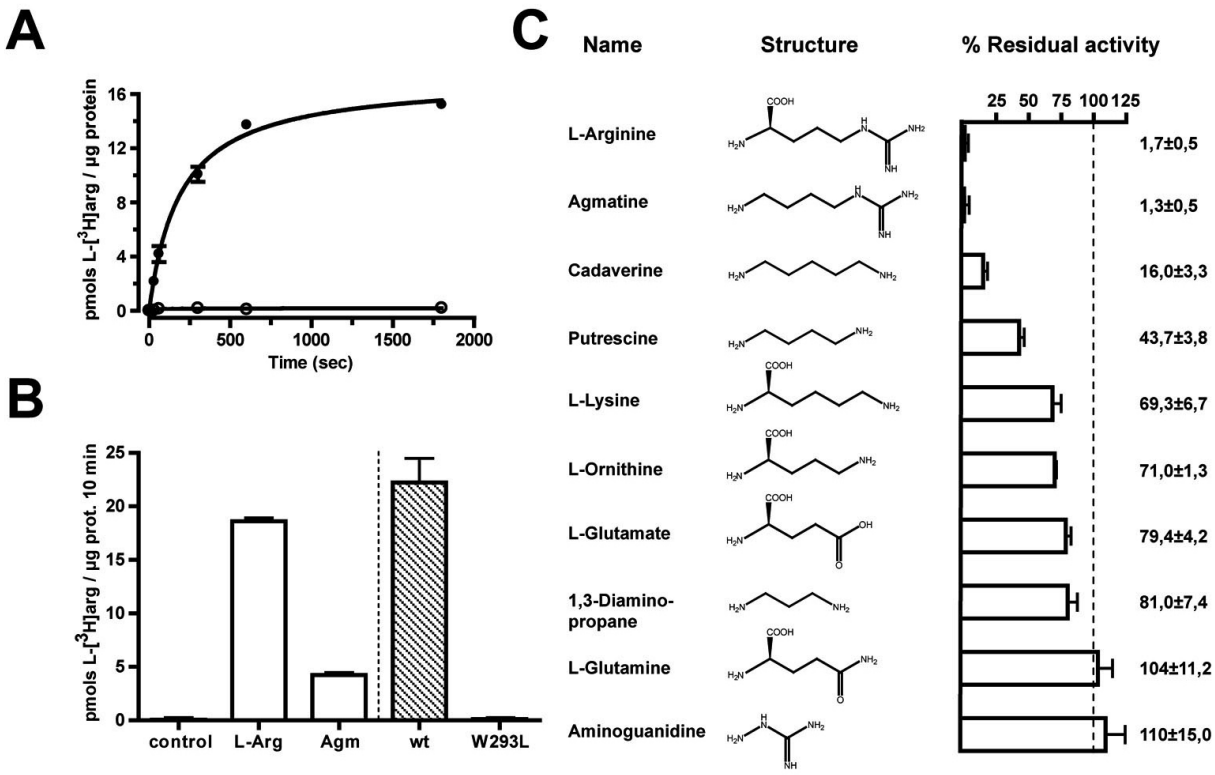

Figure 4. AdiC transport activity. (A) Time course of L-arginine transport in AdiC proteoliposomes. Influx of 10 µM L-[^{3}H] arginine into AdiC proteoliposomes lacking (open circles) or containing 2 mM L-arginine (closed circles). Data (mean ± S.E.M) correspond to a representative experiment with three replicas. Error bars when not visible are smaller than symbols. (B) Influx of 10 µM L-[^{3}H] arginine into AdiC and AdiC-W293L proteoliposomes. Transport was measured in AdiC proteoliposomes (left) containing no substrate (control), 2 mM L-arginine (L-Arg) or 2 mM agmatine (Agm). Exchange activity in AdiC (wt) and AdiC-W293L (W293L) proteoliposomes (right) was calculated by subtracting transport in the corresponding proteoliposomes with no substrate inside to that in proteoliposomes containing 2 mM L-arginine. Data (mean ± S.E.M.) correspond to a representative experiment with three replicas. A second experiment gave similar results. (C) Inhibition pattern of AdiC transport in proteoliposomes. The residual exchange activity of 10 µM L-[^{3}H] arginine (outside) and 2 mM L-arginine (inside) in the presence of the indicated substrate analogs (5 mM) in the external medium is shown. Exchange activity was calculated as in (B) and is expressed as the percentage of transport in AdiC proteoliposomes in the absence of inhibitors. Data are from four to six experiments with three replicates per condition.

AdiC-W293L proteoliposomes showed no L-[^{3}H] arginine/L-arginine exchange (Figure 4B, right). Similarly mutation to cysteine in the same residue (W293C) render a transporter that is expressed but inactive (data not shown). In order to dissect whether W293L mutation disrupts substrate recognition or translocation, substrate binding to AdiC-W293L in detergent solution was analyzed by ITC. AdiC without His-tag, with full exchange activity (data not shown), was used as a control in these studies. In agreement with a previous report using untagged AdiC (5) for both Arg and Agm, substrate titrations produce signals of heat absorption, showing that binding is enthalpically unfavorable ($\Delta H^o > 0$), and thus entropy-driven ($\Delta S^o = 24$ and 22 cal/mol-K for Arg and Agm respectively), with Arg showing substantially larger binding enthalpy (1.5 kcal/mol) than Agm (0.4 kcal/mol). The titration data are fit well by isotherms saturating at one binding site (0.96) per AdiC monomer, with equilibrium dissociation constants of 100 µM for Arg and 30 µM for Agm. AdiC-W293L showed head adsorption upon Arg and Agm titration with a equilibrium dissociation constant similar to AdiC (data not shown).

2D crystallization and projection structure of AdiC – Purified AdiC and AdiC-W293L protein was reconstituted into lipid bilayers as described in materials and methods. Both AdiC proteins yielded tubular 2D crystals diffracting better than 20 Å by negative stain. However, AdiC-W293L crystals, in contrast to AdiC, diffracted

significantly better by cryo-TEM and were, importantly, much more reproducible. Therefore, we focused on the determination of the projection structure of AdiC-W293L. Figure 5A shows typical 2D crystals of AdiC-W293L, measuring 0.4-0.6 µm in width and up to 2 µm in length. The lattice lines of such 2D crystals are barely, but still visible on noisy electron micrographs (Figure 5B). Upon adsorption on carbon film, tubular crystals became flattened into two layers, one usually diffracting better than the other in cryo-TEM. The best micrographs from cryo-TEM displayed reflections below 7 Å resolution (Table 1).

A 6.5 Å projection density map of AdiC-W293L was calculated by merging the data from five images (Figure 6A; see also Table 1 for statistics). The unit cell, which harbored four AdiC dimers, had dimensions of a=184 Å, b=119 Å, γ= 90° and a crystallographic *p*2 symmetry. As visible in Figure 6A, 2D crystals consisted of horizontal, alternating rows of upside-down and downside-up oriented AdiC-W293L dimers (upside-down is the mirror image of the downside-up dimer and vice versa). Therefore, 2D AdiC-W293L crystals expose both protein surfaces, i.e. the cytoplasmic and periplasmic side. Although not evident at first glance, small packing differences between these in the unit cell differently oriented dimers excluded a screw axis of symmetry. To further improve the projection structure of AdiC-W293L, one upside-down oriented dimer was flipped, translationally and rotationally aligned with respect to a downside-up oriented dimer, averaged and two-fold symmetrized exploiting the intrinsic symmetry of the AdiC-W293L dimer. The latter, improved the resolution of the projection map as documented in Figure 6B.

The overall shape of the AdiC dimer was elliptical with dimensions of ~95 × ~55 Å. From the 10-12 density peaks in the projection structure of the AdiC-W293L monomer, most of the densities were elongated with the exception of one single strong density peak (Figure 6B, arrowheads). The dimer interface appeared to consist of two main interacting sites between the monomers (Figure 6B, stars).

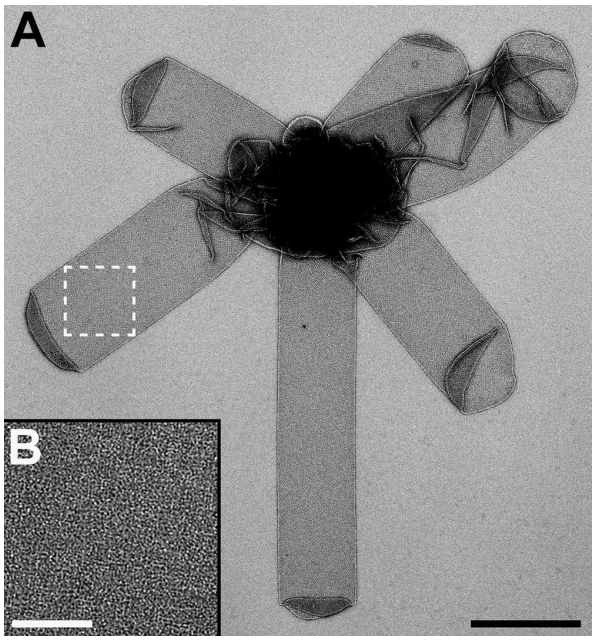

Figure 5. TEM of negatively stained 2D crystals of AdiC-W293L. (A) Overview electron micrograph of tubular AdiC crystals. The area marked by the white broken box was magnified and is displayed in panel (B). The scale bars represent 0.6 µm (A) and 0.15 µm (B).

Table I. Electron crystallographic data and statistics

Plane group symmetry	*p*2
Unit cell dimensions	a=184 Å, b=119 Å, γ= 90°
Number of processed electron micrographs	5
Resolution limit for merging	6.5 Å
Total number of observed reflections (IQ ≤5)	1262
Number of unique reflections	536
Weighted phase error	27.5° (overall)
	18.0° (200 – 15.8 Å)
	28.8° (15.8 – 11.2 Å)
	26.3° (11.2 – 9.1 Å)
	38.3° (9.1 Å – 7.9 Å)
	40.6° (7.9 Å – 7.1 Å)
	38.2° (7.1 Å – 6.5 Å)
	46.5° (6.5 Å – 6.0 Å)
	48.6° (6.0 Å – 5.6 Å)

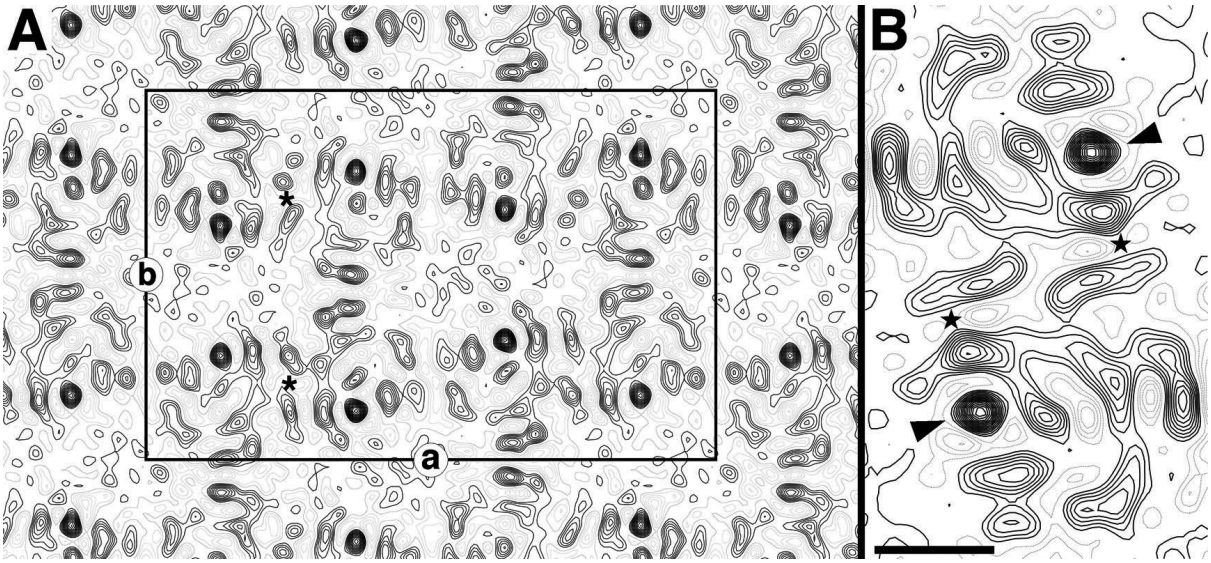

Figure 6. Projection structure of AdiC-W293L. (A) p2-symmetrized projection map of AdiC-W293L at 6.5 Å resolution calculated from five electron micrographs. The black rectangle marks the unit cell, lattice dimensions: a=184 Å, b=119 Å, γ= 90°, which contains four AdiC-W293L dimers (two upside-down and two downside-up oriented dimers). (B) Improved projection map of AdiC-W293L after averaging of the two in the lipid bilayer differently oriented AdiC-W293L dimers (e.g. the two dimers marked by asterisks) and subsequent two-fold symmetrization. The only strong density peak in the projection structure of the AdiC-W293L monomer is marked by an arrowhead. The intradimeric contact sites are indicated by stars. The scale bar represents 20 Å.

2.6 DISCUSSION

From the literature, the origin of the AdiC protein is unclear, as it has been reported to belong to two distant groups of membrane transporters: the MFS (5, 7) and the APC superfamily (6). A *de novo* analysis of the phylogenetic relationship of AdiC with known transport proteins clearly indicated that AdiC belongs to the APC superfamily (Figure 1). On our aim to elucidate the structure of membrane transporters from the APC superfamily using 2D crystals and cryo-TEM, we cloned, overexpressed, purified, and biochemically and functionally characterized wild-type AdiC and the point mutant AdiC-W293L. For 2D crystallization, overexpression of AdiC and AdiC-W293L yielded large quantities of pure protein (Figure 2A and 2C), i.e. 1-3 mg per liter of bacterial cell culture. BN-PAGE of purified AdiC and AdiC-W293L indicated an apparent molecular mass between 140 and 160 kDa, depending on the experimental conditions. Assuming an average mass of ~48.5 kDa for both AdiC forms and a DDM/coomassie brilliant blue G-250 micelle of ~45 kDa bound to the proteins during BN-PAGE (35), AdiC and AdiC-W293L exist as dimers when solubilized in DDM. The latter was supported by TEM of negatively stain AdiC and AdiC-W293L proteins that showed elongated particles resembling to two connected rings (Figure 3B and 3C). Furthermore, the elliptical shape of AdiC with dimensions of ~120 × ~75 Å is about twice as long as that of detergent-solubilized SteT monomers (SteT is also a member of the APC superfamily), which have dimensions of ~70 × ~60 Å (23). Interestingly,

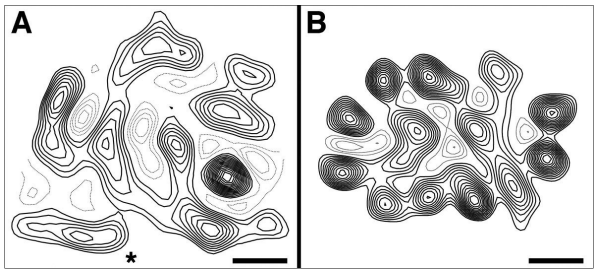

Figure 7. Comparison of the AdiC-W293L and OxlT projection structures. (A) Projection map of the AdiC monomer at a resolution of 6.5 Å, a member of the APC transporters superfamily. The asterisk indicates the two-fold symmetry axis of the AdiC dimer. (B) Projection map of the OxlT monomer at 6 Å resolution (41), a member of the MFS. The scale bar represents 10 Å.

also the overall low-resolution structures of SteT and AdiC monomers are similar being round-shaped with a central indentation. Finally, our results on the dimeric nature of AdiC are in line with a recent report by Fang *et al.*, 2007.

Reconstitution of AdiC into proteoliposomes confirmed the L-arginine/agmatine exchanger activity of the transporter (Figures 4A and 4B) and indicated relevant substrate recognition similarities to PotE. Thus, in both transporters the inhibition competence of substrate analogs depends on the size of the aliphatic chain connecting the two primary amino groups: cadaverine > putrescine > 1,3-diaminopropane in AdiC (Figure 4C) and putrescine > 1,3-diaminopropane in PotE (34). In PotE the mutation W292L abolishes the transport activity of the protein (34). The same was the case for AdiC when introducing the corresponding point mutation, i.e. W293L (Figure 4B). Based on the crystal structure and mutational analysis of PotF, a periplasmic substrate-binding protein of the putrescine-specific ABC transport system in *E. coli* (36), and on the absence of transport activity in the mutant PotE-W292L, it was proposed that W292 is one of the key residues in the recognition of the butane moiety of putrescine in PotE (34).

The high amounts of pure and homogeneous protein formed an excellent basis for the 2D crystallization of AdiC and AdiC-W293L. Both AdiC proteins had the propensity to form 2D crystals. However, AdiC-W293L 2D crystals were much more reproducible and better ordered diffracting below 7 Å resolution. Similar to the lactose permease mutant LacY-C154G (37), the introduction of a point mutation in AdiC, i.e. AdiC-W293L, improved the crystal quality and reproducibility considerably compared to wild-type AdiC crystals. Therefore, we focused on the determination of the projection structure of AdiC-W293L.

2D crystals of AdiC-W293L yielded the first detailed view in projection of a transport protein from the APC superfamily. Although the exact number of transmembrane helices and their relative orientations within the molecule cannot be determined from the projection structure of AdiC-W293L, meaningful information is gained when comparing with known membrane protein structures. For instance, the elongated density peaks suggest that the fold of AdiC consists mainly of tilted helices with the exception of one single strong density peak (Figure 6B, arrowheads) which suggests a perpendicular membrane spanning α-helix. This is based on the projection and 3D structures of aquaporins (38) and bacteriorhodopsin (39) that mainly consist of tilted or vertical transmembrane helices, respectively. As mentioned above, it is unclear from the literature, if AdiC belongs to the MFS or APC superfamily. To verify our results from phylogenetic analysis (Figure 1), indicating that AdiC is a member of the APC superfamily, we compared our AdiC projection structure with that of the oxalate/formate antiporter from *Oxalobacter formigenes* (OxlT). OxlT is a typical twelve transmembrane helix transport protein from the MFS (40) (Figure 7) and the molecular mass is comparable to that of AdiC. The difference of the two proteins is striking supporting the notion that AdiC does not belong to the MFS. Interestingly, the packing of the structural elements in OxlT is dense in contrast to AdiC. This is also reflected in the protein density area in the projection map of AdiC, which is distinctly larger than that of OxlT (Figure 7). The OxlT monomer contains a near-two-fold symmetry that is evident in the projection map (Figure 7B), relating two sets of six membrane-spanning helices, consistent with the clear homology between the first six and last six transmembrane segments in the MFS (33). This intrinsic near-two-fold symmetry is not present in the AdiC-W293L monomer (Figure 7A). In summary, we predict based on this comparison a markedly different fold for APC family members compared to transporter proteins from the MFS.

2.7 ACKNOWLEDGMENTS

The authors are thankful to Teruhisa Hirai and Sriram Subramaniam for providing the projection structure of OxlT and to Susanna Bial for technical assistance in preparing *E. coli* membranes expressing the AdiC protein versions used. This work was supported by the EC Project Grant 502802 EUGINDAT (to D.F. and M.P.), by the Berne University Research Foundation (to D.F.), by the Spanish Ministry of Science and Education Grant BFU2006-14600 (to M.P.), by the Swiss National Center of Competence in Research for Structural Biology (to A.E.) and by the Maurice E. Müller Foundation of Switzerland (to A.E.). M.R. is a recipient of a pre-doctoral fellowship from the Spanish Ministry of Science and Education.

2.8 REFERENCES

1. Donnenberg MS. (2000) *Nature* **17** 768-74
2. Foster JW. (2004) *Nat. Rev. Microbiol* **2** 898-907
3. Gong S., Richard H., Foster JW. (2003) *J. Bacteriol* **185** 4402-9
4. Iyer R., Williams C., Miller C. (2003*) J. Bacteriol* **185** 6556-61
5. Fang Y., Kolmakova-Partensky L., Miller C. (2007) *J Biol Chem* **282** 176-82
6. Jack DL., Paulsen IT., Saier MH. (2000) *Microbiology* **146** 1797-814
7. Saier MH Jr., Beatty JT., Goffeau A., Harley KT., Heijne WH., Huang SC., Jack DL., Jähn PS., Lew K., Liu J., Pao SS., Paulsen IT., Tseng TT., Virk PS. (1999) *J Mol Microbiol Biotechnol* **1** 257-79
8. Hu LA., King SC. (1998) *Biochem J* **336** 69-76
9. Ellis J., Carlin A., Steffes C., Wu J., Liu J., Rosen BP. (1995) *Microbiology* **141** 1927-35
10. Cosgriff AJ., Pittard AJ. (1997) *J Bacteriol* **179** 3317-23
11. Gasol E., Jiménez-Vidal M., Chillarón J., Zorzano A., Palacín M. (2004) *J Biol Chem* **279** 31228-36
12. Nicholson B., Manner CK., Kleeman J., MacLeod CL. (2001) *J Biol Chem* **276** 15881-5
13. Yeramian A., Martin L., Arpa L., Bertran J., Soler C., McLeod C., Modolell M., Palacín M., Lloberas J., Celada A. (2006) *Eur J Immunol.* **36** 1516-26
14. Yeramian A., Martin L., Serrat N., Arpa L., Soler C., Bertran J., McLeod C., Palacín M., Modolell M., Lloberas J., Celada A. (2006) *J Immunol* **176** 5918-24
15. Fuchs BC., Bode BP. (2005) *Semin Cancer Biol* **15** 254-66
16. Calonge MJ., Gasparini P., Chillarón J., Chillón M., Gallucci M., Rousaud F., Zelante L., Testar X., Dallapiccola B., Di Silverio F. (1994) *Nat Genet* **6** 420-5
17. Feliubadaló L., Font M., Purroy J., Rousaud F., Estivill X., Nunes V., Golomb E., Centola M., Aksentijevich I., Kreiss Y., Goldman B., Pras M., Kastner DL., Pras E., Gasparini P., Bisceglia L., Beccia E., Gallucci M., de Sanctis L., Ponzone A., Rizzoni GF., Zelante L., Bassi MT., George AL Jr., Manzoni M., De Grandi A., Riboni M., Endsley JK., Ballabio A., Borsani G., Reig N., Fernández E., Estévez R., Pineda M., Torrents D., Camps M., Lloberas J., Zorzano A., Palacín M.; International Cystinuria Consortium (1999) *Nat Genet* **23** 52-7
18. Torrents D., Mykkänen J., Pineda M., Feliubadaló L., Estévez R., de Cid R., Sanjurjo P., Zorzano A., Nunes V., Huoponen K., Reinikainen A., Simell O., Savontaus ML., Aula P., Palacín M. (1999) *Nat Genet* **21** 293-6
19. Borsani G., Bassi MT., Sperandeo MP., De Grandi A., Buoninconti A., Riboni M., Manzoni M., Incerti B., Pepe A., Andria G., Ballabio A., Sebastio G. (1999) *Nat Genet* **21** 297-301
20. Kaleeba JA., Berger EA. (2006) *Virology* **354** 7-14
21. Yernool D., Boudker O., Jin Y., Gouaux E. (2004) *Nature* **431** 811-8
22. Yamashita A., Singh SK., Kawate T., Jin Y., Gouaux E. (2005) *Nature* **437** 203-5
23. Reig N., del Rio C., Casagrande F., Ratera M., Gelpí JL., Torrents D., Henderson PJ., Xie H., Baldwin SA., Zorzano A., Fotiadis D., Palacín M. (2007) *J Biol Chem* **282** 13270-81
24. Stahlberg H., Fotiadis D., Scheuring S., Rémigy H., Braun T., Mitsuoka K., Fujiyoshi Y., Engel A. (2001) *FEBS Lett* **504** 166-72
25. Raunser S., Appel M., Ganea C., Geldmacher-Kaufer U., Fendler K., Kühlbrandt W. (2006) *Biochemistry* **45** 12796-805
26. Do CB., Mahabhashyam MS., Brudno M., Batzoglou S. (2005) *Genome Res* **15** 330-40
27. Huelsenbeck JP., Ronquist F. (2001) *Bioinformatics* **17** 754-5
28. Letunic I., Bork P. (2007) *Bioinformatics* **23** 127-8
29. Xie H., Patching SG., Gallagher MP., Litherland GJ., Brough AR., Venter H., Yao SY., Ng AM.,

Young JD., Herbert RB., Henderson PJ. and Baldwin SA. (2004) *Mol Membr Biol* **21** 323-336
30. Schägger H., von Jagow G. (1991) *Anal Biochem* **199** 223-31
31. Hirai T., Murata K., Mitsuoka K., Kimura Y., Fujiyoshi Y. (1999) *J Electron Microsc* **48** 653-8
32. Philippsen A., Schenk AD., Signorell GA., Mariani V., Berneche S., Engel A. (2007) *J Struct Biol* **157** 28-37
33. Philippsen A., Schenk AD., Stahlberg H., Engel A. (2003) *J Struct Biol* **144** 4-12
34. Kashiwagi K., Kuraishi A., Tomitori H., Igarashi A., Nishimura K., Shirahata A., Igarashi K. (2000) *J Biol Chem* **275** 36007-12
35. Suda K., Filipek S., Palczewski K., Engel A., Fotiadis D. (2004) *Mol Membr Biol* **21** 435-46
36. Vassylyev DG., Tomitori H., Kashiwagi K., Morikawa K., Igarashi K. (1998) *J Biol Chem* **273** 17604-9
37. Abramson J., Smirnova I., Kasho V., Verner G., Iwata S., Kaback HR. (2003) *FEBS Lett* **555** 96-101
38. Engel A. (2000) *Nephrol Dial Transplant* **6** 23-5
39. Subramaniam S., Hirai T., Henderson R. (2002) *Philos Transact A Math Phys Eng Sci* **360** 859-74
40. Heymann JA., Sarker R., Hirai T., Shi D., Milne JL., Maloney PC., Subramaniam S. (2001) *EMBO J* **20** 4408-13

3. High -Throughput Single Molecule Force Spectroscopy for Membrane Proteins

Patrick D Bosshart[1], Fabio Casagrande[1], Patrick LTM Frederix[1], Merce Ratera[2], Christian A Bippes[3], Daniel J Müller[3], Manuel Palacin[2], Andreas Engel[1] and Dimitrios Fotiadis[1,4]

[1] M. E. Müller Institute for Structural Biology, Biozentrum of the University of Basel, CH-4056 Basel, Switzerland

[2] Institute for Research in Biomedicine, Barcelona Science Park, Department of Biochemistry and Molecular Biology, Faculty of Biology, University of Barcelona and Centro de Investigacion Biomedica en Red de Enfermedades Raras, E-08028 Barcelona, Spain

[3] BioTechnology Center, Technical University, Tatzberg 47, D-01307 Dresden, Germany

[4] Institute for Biochemistry and Molecular Medicine of the University of Berne, CH-3012 Berne, Switzerland

3.1 ABSTRACT

Atomic force microscopy-based single molecule force spectroscopy (SMFS) is a powerful tool to study the mechanical properties, inter- and intra-molecular interactions, unfolding pathways and energy landscapes of membrane proteins. One limiting factor for the large-scale applicability of SMFS on membrane proteins is its low efficiency in data acquisition. We have developed a semi-automated high-throughput (HT)-SMFS procedure for efficient data acquisition. In addition, we present a coarse filter to efficiently extract protein unfolding events from large data sets. The HT-SMFS procedure and the coarse filter were validated using the proton pump bacteriorhodopsin (BR) from *Halobacterium salinarum* and the L-arginine/agmatine antiporter AdiC from the bacterium *Escherichia coli*. To screen for molecular interactions between AdiC and its substrates, we recorded data sets in the absence and in the presence of L-arginine, D-arginine and agmatine. Altogether ~400,000 force-distance curves were recorded. Application of coarse filtering to this wealth of data yielded six data sets with ~200 (AdiC) and ~400 (BR) force-distance spectra in each. Importantly, the raw data for most of these data sets were acquired in one to two days opening new perspectives for HT-SMFS applications.

3.2 INTRODUCTION

The atomic force microscope (AFM) has become a versatile tool to image the surface topography of biological membranes at a lateral resolution of ~5 Å and a vertical resolution of ~1 Å under near native conditions (1). In addition, the AFM is successfully used to probe mechanical properties of proteins down to the single molecule level. Such single molecule force spectroscopy (SMFS) has developed into a powerful technique to study the unfolding of individual membrane proteins incorporated into lipid bilayers (2). SMFS unfolding curves have a saw-tooth pattern appearance that is characteristic of the investigated protein. The unique 'fingerprint' of a specific protein allows localizing and quantifying molecular interactions that stabilize the assembly of secondary structure elements. Furthermore, SMFS has been used to measure interactions of membrane proteins with ligands and small molecules as well as their activation (2). Additionally, kinetical parameters of membrane protein unfolding, e.g. unfolding rates, were extracted from such experiments (3). Sapra and co-workers (4) have demonstrated by SMFS that the oligomeric state of bacteriorhodopsin influences the intrinsic stability of the structural elements within the monomer but not its unfolding pathway. However, the unfolding pathway of an entire protein complex cannot be studied by SMFS, because of the lack of covalent catenation between the proteins forming the complex.

The procedure of SMFS on membrane proteins can be divided into several steps: membranes containing the integral protein of interest have first to be localized by AFM imaging. Then, force-distance (F-D) curves are recorded by pushing the AFM tip onto membrane proteins and retraction after a short delay. If a protein adheres strongly to the cantilever with one of its termini, it can be mechanically unfolded upon cantilever retraction. Subsequent data processing involves (i) data coarse filtering to extract F-D curves that show unfolding events, (ii) classification of the latter, (iii) alignment of F-D curves and (iv) their analysis based on polymer chain models. Large-scale applicability of SMFS on membrane proteins is, however, restricted by its low efficiency: not every F-D curve leads to an unfolding event. For the bacterial sodium/proton antiporter it was estimated that complete protein unfolding occurred in only ~3% of the approaches (5), and this when using two-dimensional (2D) crystals where the membrane proteins are packed at the highest possible density. When applying SMFS to a more realistic system, i.e. non-crystalline proteoliposomes

or even native membranes, the probability for protein adhesion to the tip is considerably lower than with 2D crystals. In addition, membrane proteins may follow different unfolding pathways with occurrence probabilities below 10% (2). This requires performing a considerable number of experiments and makes the analysis of such large data sets challenging. Furthermore, ions, small molecules or the environment, e.g. temperature, all influence the unfolding behavior of membrane proteins, giving rise to subtle changes in the distribution of unfolding pathways (6, 7). Consequently, not only one but several large data sets have to be recorded, i.e. one for each experimental condition. For these reasons, high-throughput (HT) protocols for each of the SMFS procedure steps are desirable. To date, fully automated SMFS procedures on membrane proteins are not feasible. The main reason for this is that lipid membranes containing the protein of interest have first to be located by AFM imaging and eventually manipulated before SMFS, e.g. to remove aggregates or the upper layer of a collapsed vesicle to expose areas where force spectra can be recorded. Therefore, imaging and sample manipulation by AFM prior to force spectroscopy are currently the limiting steps towards a fully automated HT-SMFS procedure for membrane proteins.

Automated alignment and pattern-recognition algorithms have been described in the past (8-10). However, the preceding data recording and coarse filtering has not been optimized yet. Data acquisition is a bottleneck towards HT-SMFS: recording of a single F-D curve takes about one second, but the manual recording of a complete data set, consisting of ~200 spectra, with an unfolding efficiency of less than 1% is laborious work. Furthermore, manual recording comprises online filtering by the operator leading to the loss of relevant force spectra and consequently to a decrease in the data acquisition efficiency. In the worst case, one or several classes of force spectra might even be ignored if the operator is biased towards a specific 'fingerprint'. Such bias is eliminated by the application of an automated acquisition and saving procedure. The acquired data is subsequently filtered to extract the F-D curves of completely unfolded proteins, i.e. pulled from one of their termini. This filtered data is then used for further processing and classification.

In this work, we present a HT-SMFS procedure, which addresses the semi-automated recording of F-D curves and the off-line coarse filtering of the acquired data. The method was validated by its application to 2D crystals of the proton pump bacteriorhodopsin (BR) from *Halobacterium salinarum* and to proteoliposomes of the L-arginine/agmatine antiporter AdiC from *Escherichia coli* (*E. coli*). While the atomic structure of BR has been elucidated (11, 12), the structure of AdiC is unknown. Furthermore, BR has been extensively characterized by SMFS (6, 13-15) and other methods, whereas for AdiC only very few functional and structural data are available. AdiC is a component of the bacterial acid resistance mechanism (16) and a member of the amino acid/polyamine/organocation (APC) transporter superfamily. It is a 445 amino acid (aa) protein with a molecular mass of ~47 kDa and expected to fold into 12 transmembrane α-helices with cytoplasmic N- and C-termini (17, 18). Recently, Fang *et al.* (19) have over-expressed, and functionally and biochemically characterized the AdiC exchange transporter showing among other things that it forms homodimers.

We have taken BR as well-characterized control and the AdiC protein as test case for our HT-SMFS procedure. Two main classes of force spectra were obtained for AdiC and attributed to proteins unfolded by pulling from the N- and from the C-terminus, respectively. Importantly, the assignment of a main class to a terminus was made possible by the use of two different recombinant forms of the AdiC protein differing in the length of their termini. To screen for molecular interactions between AdiC and its substrates, we recorded data sets in the absence and in the presence of L-arginine, D-arginine and agmatine. This means that ~400,000 F-D curves had to be recorded to obtain six data sets (including the BR control),

with ~200 force spectra in each (except for BR with ~400 curves), demonstrating the usefulness of the HT-SMFS protocol presented here.

3.3 MATERIALS AND METHODS

Cloning, expression, purification, reconstitution and transport measurement of AdiC – AdiC was cloned from genomic *E. coli* strain DH5α DNA as previously described for other transporters (20). Shortly, the following primers (5'-3') were used to amplify the AdiC open reading frame by PCR: ATGAGAATTCACTATCGA-ATGCAGGCGTATG and TCATCTCGAGGTT-GGCTTTTATGTTTGCTGGA. PCR products were digested with EcoRI and XhoI and ligated into a pBlueScript vector (Stratagene, La Jolla, CA, USA). Another PCR with the following primers (5'-3') was performed to subclone the AdiC open reading frame into the EcoRI and PstI sites of a modified version of the vector pTTQ18 (21), thereby placing its expression under the control of the *tac* promoter: TGATGAATTCGATGTCTTCGGATGCT-GATGC and ACCGCCTGCAGAATCTTTGC-TTATTGGTGCA. To subclone AdiC into the pTrcHisA vector another PCR was performed with the pBlueScript construct using as primers (5'-3'): CAATGGATCCATGTCTTCGGATGCTG-ATGCTC and TCATCTCGAGGTTGGCTTTT-AT-GTTTGCTGGA. Digestion of the PCR product with BamHI and XhoI and ligation allowed to subclone AdiC into the new vector. All DNA constructs were verified by sequencing. The two AdiC constructs with either an N-terminal (pTrcHis-construct: N-His_6-AdiC; 481 aa long) or C-terminal (pTTQ18-construct: AdiC-His_6-C; 461 aa long) His_6-tag were expressed in *E. coli*, purified by nickel affinity chromatography, reconstituted into *E. coli* lipid bilayers and assayed for transport function as previously described (20). Proteoliposomes for AFM and SMFS were reconstituted by dialysis (dialysis buffer: 150 mM NaCl, 10% glycerol, 0.01% NaN_3, 20 mM Tris-HCl pH 8) at lipid-to-protein ratios between 0.5 and 1.0 using dodecylmaltoside-solubilized AdiC and *E. coli* polar lipids (Avanti Polar Lipids Inc., Alabaster, AL, USA).

High-throughput single molecule force spectroscopy (HT-SMFS) – HT-SMFS experiments were performed on a commercial AFM (NanoWizard® II Ultra, JPK Instruments, Berlin, Germany) equipped with a 30 µm *x* and *y* closed-loop controlled piezo-scanner. To reduce noise from the surroundings, the AFM was placed on an active damping table (MOD-1L plus, Halcyonics GmbH, Göttingen, Germany) and acoustically shielded. Rectangular silicon nitride cantilevers (ORC8-PS, $k_{nominal}$ = 0.05 $N \cdot m^{-1}$, Olympus Ltd., Tokyo, Japan) were used. The deflection sensitivity (*DS*) was determined from the slope of the linear regime of F-D curves recorded on bare mica in buffer solution (150 mM NaCl, 20 mM citric acid pH 5.0). The spring constant, *k*, of the cantilevers was determined using the thermal noise method (22). BR crystals were adsorbed onto mica as previously described (23). Lipid vesicles containing densely packed AdiC were adsorbed onto freshly cleaved mica in buffer solution (150 mM NaCl, 20 mM citric acid pH 5.0) at room temperature. After adsorption, the sample was rinsed several times with the same buffer to remove non-adsorbed membranes. HT-SMFS experiments were performed in the following buffer solutions: 150 mM KCl, 20 mM Tris-HCl pH 7.8 (BR) and 150 mM NaCl, 20 mM citric acid pH 5.0 (AdiC). To investigate AdiC-substrate interaction, vesicles containing densely packed AdiC were incubated in substrate containing buffer solution (150 mM NaCl, 20 mM citric acid pH 5.0, 10 mM substrate; substrate: L-arginine, agmatine or D-arginine) for 30 min at 4 °C and then adsorbed on mica in the same substrate containing buffer at room temperature. All buffers were prepared from ultrapure water from a NANOpure Diamond™ system (Barnstead, Dubuque, IA, USA, resistivity ≥ 18 MΩ·cm). Proteoliposomes densely packed with AdiC were located by contact mode AFM prior to force spectroscopy. Where necessary, the AFM tip was used as a nanoscalpel to remove aggregates

or the upper layer of a collapsed vesicle (24) thus exposing suitable membrane areas for force spectroscopy. For SMFS experiments the force spectroscopy tool of the NanoWizard® II Ultra AFM was used. A point grid with an edge length of 150 - 300 nm and a linear point density of 0.125 nm^{-1} was placed on 2D BR crystals (a kind gift of Dr. Georg Büldt, Jülich, Germany) or single-layered densely packed AdiC membranes (Figure 1). At each point of the grid ten successive measurements were performed. Membrane proteins were attached to the AFM tip by pushing the cantilever ten times onto the membrane at each grid point with a force of 1 nN for 0.1 - 0.6 s. The cantilever was then retracted for 0.25 s with a velocity of 0.53 $\mu m \cdot s^{-1}$ (BR) and 0.78 $\mu m \cdot s^{-1}$ (AdiC), respectively. All F-D curves were saved and consisted of 4096 data points corresponding to a sampling frequency of 16.4 kHz. The cantilever deflection was detected with a position-sensitive photodiode (PSD). Its output voltage, V_{PSD}, was converted into a metric deflection by multiplication with the cantilever *DS*, and into a force by subsequent multiplication with the spring constant.

Data filtering and data analysis – For F-D curve analysis only retraction data were used. F-D curves were shifted vertically so that the average force in the baseline, defined as the final 5% of the data points in the tail part of the spectrum, was zero. The actual cantilever *DS* of each F-D curve was calculated from the slope of the deflection regime between 300 pN and 1 nN where the force increased linearly with piezo movement. Tip-sample separation values (*tss*) were computed by subtracting the corrected metric cantilever deflection from the position of the *z*-piezo. The average *tss* value of the points used for the actual *DS* determination was defined as zero distance between tip and surface, i.e. *tss* = 0. For representation purpose all force spectra were flipped vertically so that pulling force peaks faced upwards. Therefore, negative and positive forces in force spectra represent pushing and pulling forces, respectively.

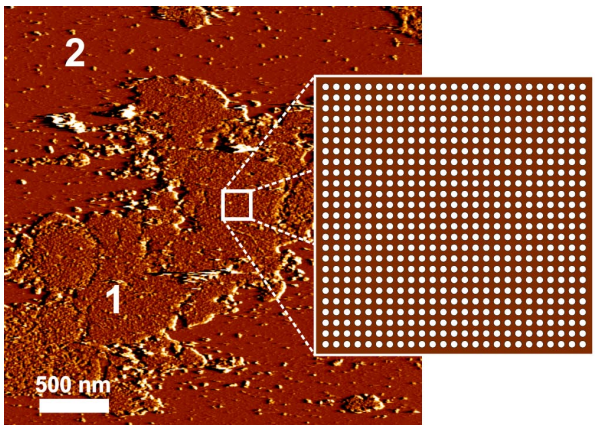

Figure 1. The cantilever deflection image shows a single-layered membrane (marked 1), densely packed with AdiC proteins (dots on this surface type) and the mica surface (marked 2). The white box on the AdiC proteoliposome shows the area (frame size: 200 nm), where F-D curves were recorded. The zoomed area contains 625 positions at which ten measurements per position were performed yielding 6250 F-D curves per area.

A coarse data filtering process (Figure 2) was applied to the acquired data sets to extract unfolding events. Filtering was performed using a home-written macro for IGOR Pro (IGOR Pro 6.0.1.0, Wavemetrics Inc., Portland, OR, USA) run on a PC with an Intel Core™2 CPU T7400 2.16 GHz processor equipped with 2 GB RAM. The following filtering criteria were used:

1) The standard deviation of the force, σ_F, of the final 5% of the data points in F-D curves was taken as a measure of the flatness of the non-contact region. To ensure that only F-D curves with flat non-contact region were imported, the maximum allowed σ_F was set to F_{thr1} = 20 pN.

2) The maximal *tss* beyond which no significant pulling forces should be found was set to tss_{max}. This value is based on the contour length, $L_{C,max}$, of the corresponding protein amino acid sequence. Because the final unfolding event occurs before the protein is fully extended, tss_{max} was chosen slightly below $L_{C,max}$. The introduction of tss_{max} allowed omitting F-D curves having force peaks at positions beyond the full-length of the protein.

3) Because the AFM tip can adsorb at different positions along the protein termini, e.g. first or last amino acid of a terminus, individual force

spectra are shifted horizontally with respect to each other in a given tss range. This range is defined by the tss_{min} and tss_{max} parameters in the filtering algorithm. Naturally, the last force peak indicating complete unfolding of the protein is located between tss_{min} and tss_{max}.

4) F-D curves revealing pushing (negative) force peaks should not be considered in the analysis. Therefore, the *tss* interval between tss_{low} = 5 nm and tss_{max} was searched for pushing (negative) forces below a critical threshold. tss_{low} was set to 5 nm because the front part of a F-D curve is often dominated by nonspecific tip-sample interactions that impede correct threshold determination.

With the exception of the determination of the force noise σ_F, deflection data were flattened by a 5 points binomial smoothing before threshold testing. In summary, curves were identified as possible unfolding events when they fulfilled the following selection criteria:

$$\sigma_F < F_{thr1} \quad \text{for } (1\text{-}frac)\cdot tss_{end} < tss < tss_{end}$$

$$F_{SMFS,max} < F_{thr2}\cdot\sigma_F \quad \text{for } tss_{max} < tss < tss_{end}$$

$$F_{SMFS,max} > F_{thr3}\cdot\sigma_F \quad \text{for } tss_{min} < tss < tss_{max}$$

$$F_{SMFS,min} > -F_{thr4}\cdot\sigma_F \quad \text{for } tss_{low} < tss < tss_{max}$$

where *frac* is the fraction of points, which is used to determine the baseline of the F-D curve. $F_{thr2\text{-}4}$ are dimensionless numbers, which yield filtering force thresholds in picoNewtons when multiplied with σ_F. Here, tss_{end} is the maximal *tss* value of

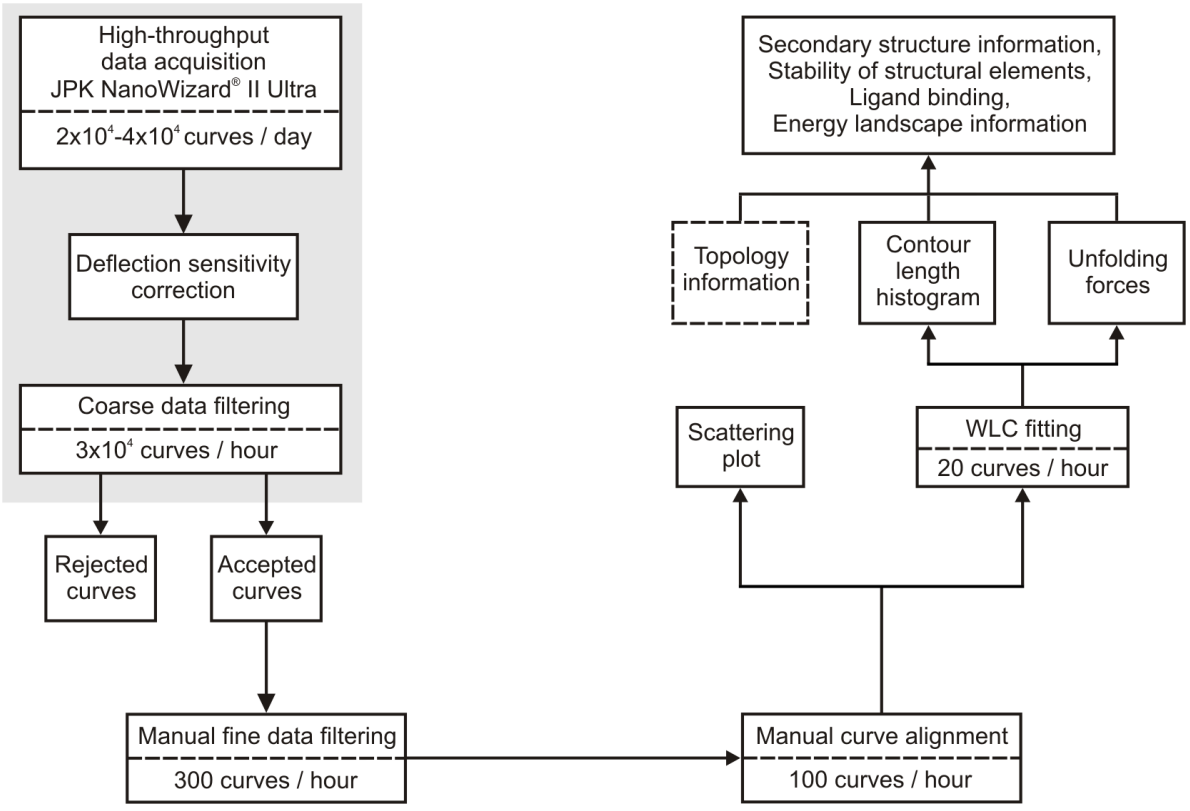

Figure 2. Flow chart describing the HT-SMFS procedure. It starts with data acquisition and an automated/manual data filtering part. Force spectra of the protein of interest are aligned with respect to an exemplary curve and then the WLC model (25, 26) is fitted to each curve to determine the contour length, L_C, of each stretched polypeptide segment as well as the required unfolding force. For visualization, a scattering plot of the aligned curves is computed. The contour length in number of amino acids is compared with a sequence-based topology model of the investigated membrane protein (e.g. Figure 10 for AdiC) to map the locations of the unfolding barriers. The automation of the shaded part is addressed in this work.

the force spectrum and $F_{SMFS,max}$ and $F_{SMFS,min}$ are the maximal and minimal forces of a spectrum after binomial smoothing on the selected interval. F-D curves, which fulfilled the above selection criteria, were automatically exported as force versus tip-sample separation data for further analysis.

After coarse filtering, a manual fine filtering and classification (Figure 2) based on the repetitively found 'fingerprint' pattern of AdiC was performed. This step removed curves with not interpretable unfolding patterns that had passed the coarse filtering. Data acquisition was continued until ~200 classifiable AdiC unfolding traces were recorded. All traces were aligned with respect to the most prominent unfolding peaks at *tss* ≈ 25 nm for BR, *tss* ≈ 40 nm for N-His_6-AdiC and *tss* ≈ 90 nm for AdiC-His_6-C (Figure 6). Deflection data were binomially smoothed and subsequently all force peaks were fitted using the WLC model choosing a persistence length of 0.4 nm. The contour length, L_C, was the only free fitting parameter in the worm-like chain (WLC) model (25, 26). The contour length was divided by the backbone length per aa, 0.36 nm·aa^{-1}, to obtain a contour length in aa. Histograms of the obtained L_C values and unfolding forces were computed, and Gaussian curves were fitted to the peaks in the histograms. Contour lengths were correlated with the sequence-based topology model of AdiC (Figure 10) to map the unfolding barriers within the protein. Certain unfolding barriers were found on the opposite side of the pulling tip (periplasmic side). Since the polypeptide chain goes through the membrane that has a thickness of ~4.8 nm, unfolding barriers were shifted by 13 aa (4.8 nm / 0.36 nm·aa^{-1} ≈ 13 aa) to compensate for the thickness of the membrane (14, 27). This procedure allowed the entire length of the stretched polypeptide to be calculated. Scattering plots of the aligned BR and AdiC force spectra were used to highlight common features (Figure 6). They were computed using a home-written macro for IGOR Pro with a box size of 1 pN·0.25 nm for BR and 1 pN·0.5 nm for AdiC. The number of pixels in each box was counted and converted into a gray scale value. The gray scale range of all scattering plots of AdiC was normalized with respect to the number of force spectra in the data set.

3.4 RESULTS

Cloning, expression, purification, reconstitution and functional characterization of AdiC – Both recombinant forms of AdiC, i.e. N-His_6-AdiC and AdiC-His_6-C, were successfully cloned, over-expressed, purified, reconstituted and assayed for transport function according to Reig *et al.* (20). Transport activity for L-arginine and agmatine was for both AdiC constructs similar to the values obtained by Fang *et al.* (19). Thus, transport of 10 µM L-[^{3}H] arginine into proteoliposomes containing either tagged version of AdiC was completely dependent on the presence of L-arginine (~1-2 pmols / µg protein and min) or agmatine (~0.4 pmols / µg protein and min) in the inside (data not shown).

Data acquisition and validation of the filtering process – Before HT-SMFS data acquisition, reconstituted lipid membranes containing densely packed AdiC were adsorbed onto mica and located by contact mode AFM (Figure 1). If necessary, collapsed vesicles or aggregated membranes were dissected using the AFM tip as a nanoscalpel to expose flat membrane areas suitable for SMFS. This localization and manipulation step of target proteoliposomes limited the full automation of our HT-SMFS data acquisition.

Figure 1 describes a typical HT-SMFS experiment with AdiC: an area densely packed with protein was first selected (white square). The selected area was then subdivided into 625 positions (highlighted by the grid in Figure 1), at which ten F-D curves per position were recorded yielding 6250 F-D per area. The *DS* (in nm/volt) was determined for each recorded F-D curve by linear regression to the first 5-10% of the deflection curve during retraction of the cantilever. Figure

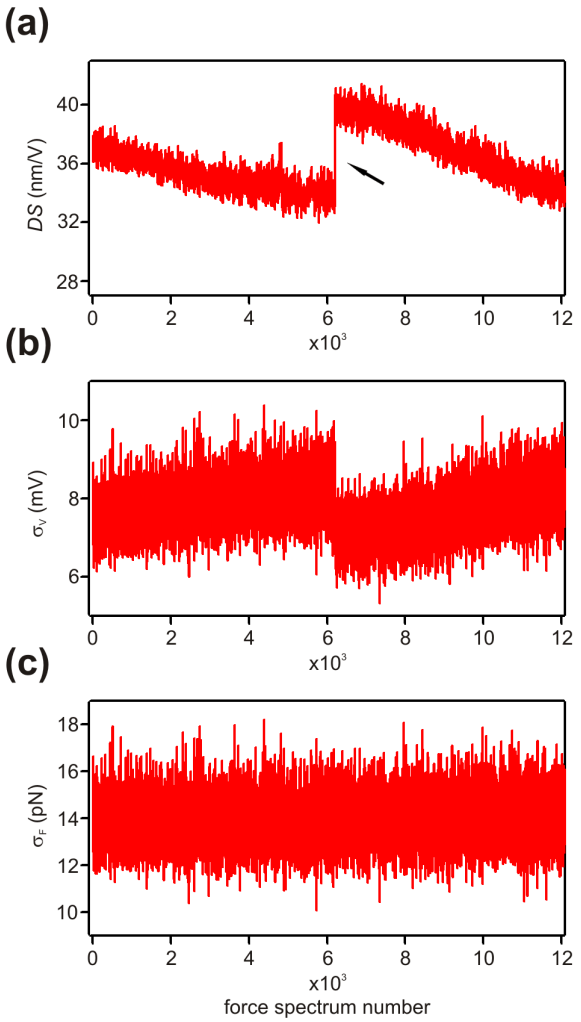

Figure 3. Changes in cantilever DS (a), cantilever deflection noise (b) and force noise (c) extracted from F-D curves taken on densely packed AdiC membranes. The step in (a) at force spectrum no. 6250 (arrow) originates from manually adjusting the laser beam to the original position on the photodiode. The force noise (σ_F), representing the real deflection noise of the cantilever, was between 10 and 18 pN.

3(a) illustrates the long-term drift and the short-term fluctuations of *DS,* both likely to be related primarily to short- and long-term variations of the laser beam position on the photodiode. For the latter, ~12'000 F-D curves were recorded in batches of 625 points on a densely packed AdiC membrane and the *DS* values determined (see Figure 3(a)). The step in the *DS* around 6250 (arrow) originated from the manual adjustment of the laser beam to its original position, which was done after F-D curves had been recorded on an area comprising 625 points. The comparable drift after repositioning of the laser beam highlights the reproducible character of the *DS* change. Monitoring the cantilever deflection noise σ_V (in volts) by analysing the last 5% of the F-D curve (~200 values) exhibits a long-term drift as well, including the step after readjustment of the laser beam (Figure 3(b)). However, the average noise level increases for decreasing *DS*. The force noise (σ_F) of the cantilever (Figure 3(c)), which is calculated from the cantilever deflection noise σ_V and the *DS* of the same F-D curve, does not exhibit a drift over time. This demonstrates that forces calculated by our approach are not impaired by long-term drifts.

The coarse filter (see materials and methods) was validated by comparing two sets of filtering parameters. The first filter was only sensitive for the presence of a pulling force over a certain threshold ($F_{SMFS,thr}$) at a minimal tip-sample separation (tss_{min}), as indication for a completely unfolded protein. Figure 4(a) shows the filtering result with tss_{min} = 95 nm and $F_{SMFS,thr} = 3.5 \cdot \sigma_F$. Such loosely defined selection criteria imported all useful unfolding traces from the data set, but also useless curves that arose from contaminants attached to the AFM stylus, e.g. unfolding curves longer than the fully stretched AdiC protein. Curves exhibiting tip contaminations were partly filtered out by excluding F-D curves with negative (pushing) forces between 5 – 150 nm (compare Figure 4(a), arrow with Figure 4(b) where negative (pushing) forces have been filtered out). The first part of the F-D curve (tss < 5 nm) was not considered because it is often dominated by negative forces arising from non-specific tip-sample interactions. The introduction of an additional parameter tss_{max} beyond which no peak should be present led to further improvement by excluding more non-interpretable F-D curves (Figure 4(b)). The value for tss_{max} was defined based on the working topology model of AdiC (see Figure 10). Filtering the AdiC data sets using the parameters: tss_{low} = 5 nm, tss_{min} = 90 nm, tss_{max} = 150 nm, F_{thr1} = 20 pN, F_{thr2} = 3, F_{thr3} = 3.5, F_{thr4} = -3 yielded a considerably smaller set of force spectra without losing good F-D curves (Figure 4(b)). Figure 4(c) shows the useful force spectra

contained in the two preceding data sets (Figures 4(a) and (b)) after manual filtering. BR data sets were filtered with the parameters: tss_{low} = 5 nm, tss_{min} = 60 nm, tss_{max} = 80 nm, F_{thr1} = 20 pN, F_{thr2} = 3, F_{thr3} = 3.5, F_{thr4} = -3. To collect enough data for unfolding of AdiC (and the control, BR) under all selected conditions (see above) a total of ~400,000 F-D curves were recorded. Our pre-processing algorithm allowed ~99% of all recorded F-D curve to be eliminated without loss of useful data. About 70% of the remaining F-D curves were then rejected by manual fine filtering yielding five data sets for AdiC consisting of ~200 spectra each and one data set for BR with 398 spectra. Since all data are stored, they are amenable to data refined pre-processing if one of the filter parameters was suspected to be too tight. Thus, our method combines the advantage of recording all F-D curves and the unbiased reduction of this overwhelming data set to a number of force spectra that can be analyzed manually.

HT-SMFS results – Up to now, the transport protein AdiC has not been studied by SMFS. Therefore, we also tested our HT-SMFS procedure on BR, a membrane protein that has been extensively characterized by force spectroscopy (6, 13-15). As shown in Figure 6(a), our procedure yielded the same results as those published previously (13, 14), validating our HT-approach. In contrast to the latter papers, our data set (Figure 6(a)) consisting of 398 complete unfolding events of BR was recorded in only one day. With this amount of curves even unfolding pathways with a probability as low as 2.5% should be detected at least 4 times (binomial distribution with $P > 0.99$).

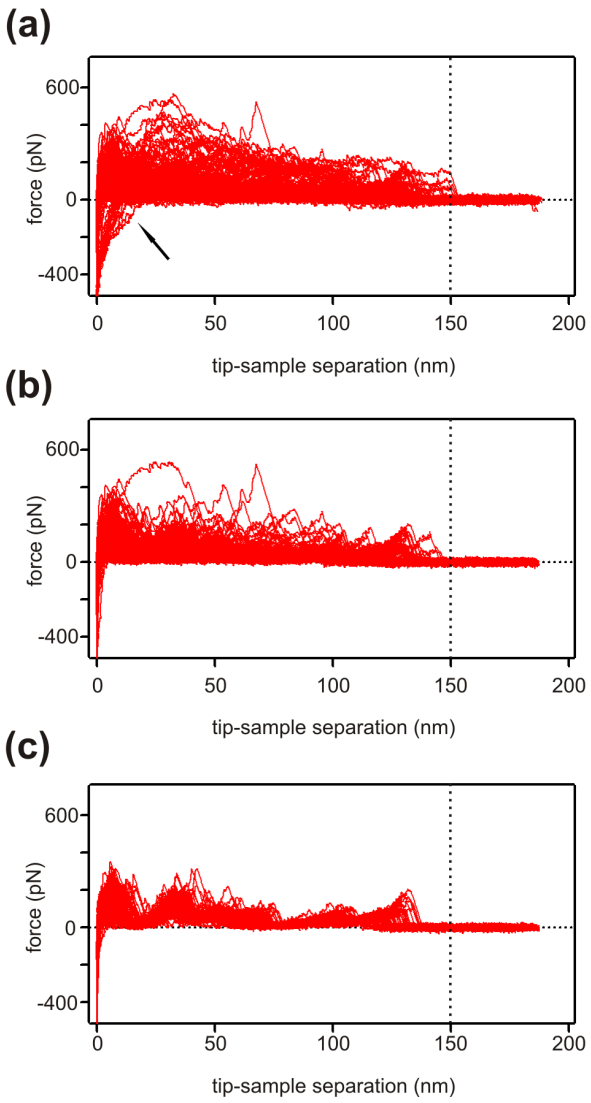

Figure 4. Exemplary filtering of AdiC F-D curves using different filtering parameters. (a) Filtering with tss_{min} = 95 nm and a minimal force of $F_{SMFS,thr}$ = 3.5·σ_F resulted in the import of 190 out of 22,000 F-D curves. Curves where the tip was contaminated are highlighted by the arrow. (b) Filtering of the same data set as in (a) using tss_{low} = 5 nm, tss_{min} = 90 nm, tss_{max} = 150 nm, F_{thr1} = 20 pN, F_{thr2} = 3, F_{thr3} = 3.5, F_{thr4} = -3. These filtering parameters led to the import of 135 curves. (c) 75 Force spectra of N-terminally unfolded AdiC extracted from (a) and (b) by manual fine filtering. The dashed line indicates tss_{max} = 150 nm.

In a second step, our HT-SMFS procedure consisting of semi-automated data acquisition and efficient data filtering was used to study the mechanical unfolding of the bacterial L-arginine/agmatine antiporter AdiC. Although 2D membrane protein crystals were used in most SMFS studies (5, 13, 28, 29), crystals are not a prerequisite for SMFS. Since the probability for an unfolding event decreases with decreasing membrane protein density in lipid bilayers, semi-automated procedures for fast data acquisition are particularly important for SMFS on non-crystalline proteoliposomes. With densely packed vesicles of AdiC, the HT-SMFS procedure described here allowed collecting data sets comprised of ~200 F-D spectra (after filtering) in only one to two days.

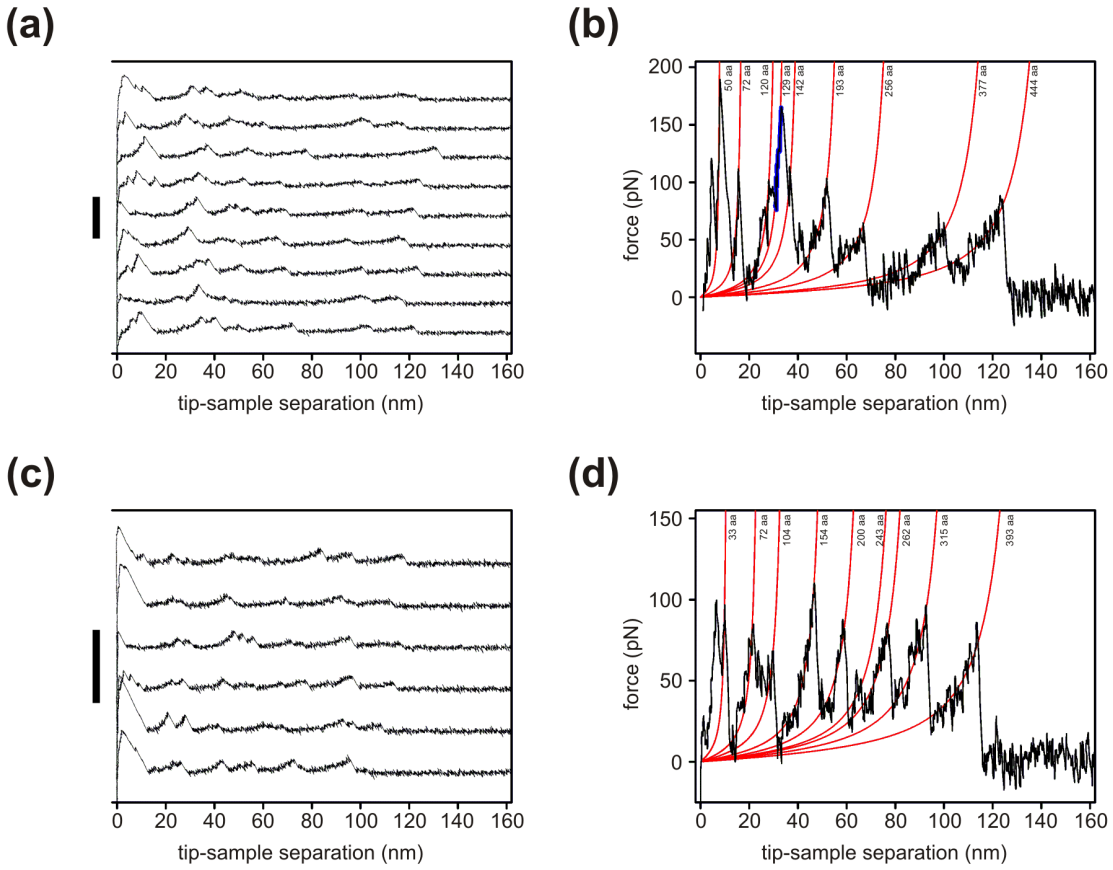

Figure 5. Representative force spectra obtained with the N-His_6-AdiC construct. (a) Force spectra of the major unfolding class revealing different unfolding pathways. (b) Exemplary force spectrum of the major unfolding class. Each peak in the spectrum corresponds to the unfolding of one or more secondary structure elements. Worm-like-chain (WLC) curves (25, 26) were fitted to all peaks to determine the contour length, L_C, in number of amino acids of the unfolded structural element(s). The given L_C values next to the WLC curve were so that the blue peak had an L_C of 129 aa. (c) Unfolding traces of the minor unfolding class of the N-His_6-AdiC construct with a representative force spectrum (d). The scale bars in (a) and (c) represent 500 pN.

Unfolding of N-His_6-AdiC resulted in two classes of F-D spectra, a major (Figure 5(a) and (b)) and a minor class (Figure 5(c) and (d)). Both classes of F-D spectra had a similar length suggesting complete protein unfolding from both termini. From a probabilistic point of view the attachment of the N-terminus to the cantilever is more likely than that of the C-terminus, because the former is predicted to be ~2.6 times longer than the latter. Another indication favouring pulling from the N-terminus was the pH-dependent interaction between AFM tip and the terminus: at pH 8 N-terminal unfolding events were observed in ~0.06% of the cases in contrast to pH 5, where the efficiency was ~0.32%. Protonation of the His_6-tag ($pK_{a,His}$ ~6) at the N-terminal end introduces a positive charge. This might promote adhesion of the protein to the silicon nitride tip, which is negatively charged even at low pH (30). However, to prove that the curves shown in Figure 5(c) and (d) corresponded indeed to unfolding events pulled from the C-terminus of N-His_6-AdiC, a second construct, AdiC-His_6-C, with a long C-terminal (containing a His_6-tag) and short N-terminal end (no His_6-tag) was subjected to HT-SMFS. As for N-His_6-AdiC the attachment of the protein to the AFM tip was pH-dependent: at pH 8 C-terminal force spectra were obtained in ~0.05% of the cases in contrast to at pH 5, where the efficiency was ~0.29%. This resulted in a single class of F-D spectra (Figure 6(e)) with a length of ~110 nm and the same 'fingerprint' as previously observed for the putative C-terminal unfolding of the N-His_6-AdiC construct (Figure 5(c) and (d)).

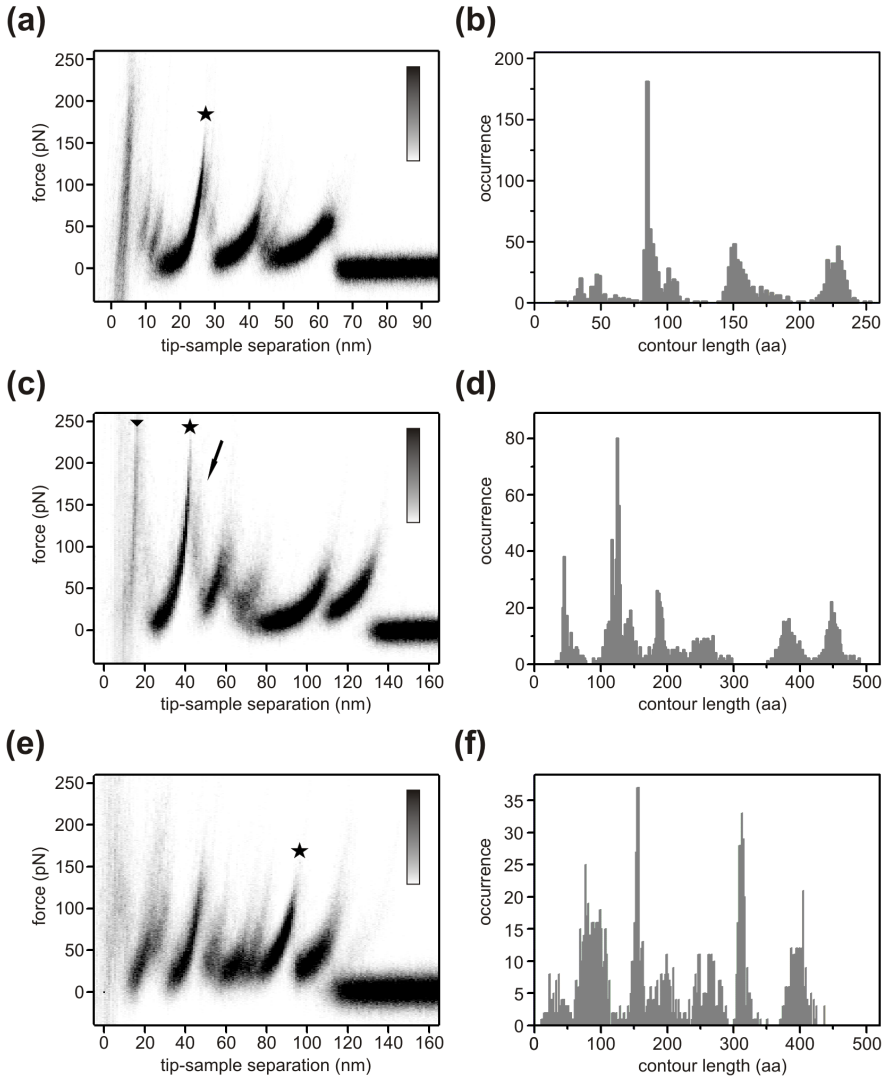

Figure 6. Scattering plots and contour length histograms for BR, and the N-His$_6$-AdiC and AdiC-His$_6$-C constructs. Scattering plot (a) and contour length histogram (b) of 398 force spectra of BR unfolded from its C-terminal end. Used filtering parameters: tss_{low} = 5 nm, tss_{min} = 60 nm, tss_{max} = 80 nm, F_{thr1} = 20 pN, F_{thr2} = 3, F_{thr3} = 3.5, F_{thr4} = -3. Force spectra were aligned with respect to the unfolding peak at $tss \approx 25$ nm (denoted by star). (c) Scattering plot of 205 force spectra of the N-His$_6$-AdiC construct unfolded from its N-terminal end. Force spectra were aligned with respect to the unfolding peak at $tss \approx 40$ nm (denoted by star). The arrowhead highlights the presence of an unfolding peak with L_C = 50 aa and the arrow shows the presence of a side peak with L_C = 147 aa. (d) Contour length histogram obtained from fitting WLC curves to each force peak. Average L_C values and peak heights are given in table 1. (e) Scattering plot of 190 force spectra of the AdiC-His$_6$-C construct unfolded from its C-terminal end. All force spectra were aligned with respect to the unfolding peak at $tss \approx 90$ nm (denoted by star). (f) Contour length histogram, values are given in table 2. The gray scale of the bars correspond to 70 (a), 90 (c) and 83 (e) counts.

AdiC unfolding experiments were performed until ~200 classifiable F-D spectra were obtained ensuring a statistically reliable data set. Figure 6(c) shows a scattering plot of ~200 force spectra for the N-terminal unfolding of N-His$_6$-AdiC, recorded during a single day using three different cantilevers. The histogram of the corresponding contour length values exhibited well-resolved peaks (Figure 6(d)). The mean contour length values of the main force peaks together with the average height are displayed in table 1. Three force peaks with L_C = 129 aa ($tss \approx 40$ nm), 387 aa ($tss \approx 110$ nm) and 455 aa ($tss \approx 130$ nm) were highly conserved, while the other events had lower probabilities (table 1). In the diffuse front part ($0 < tss < 20$ nm) of the scattering plot in Figure 6(c) a peak with L_C = 50 aa was found. Statistical analysis of the occurrence probability of this peak yielded a probability of 52%. However, this probability represents a lower

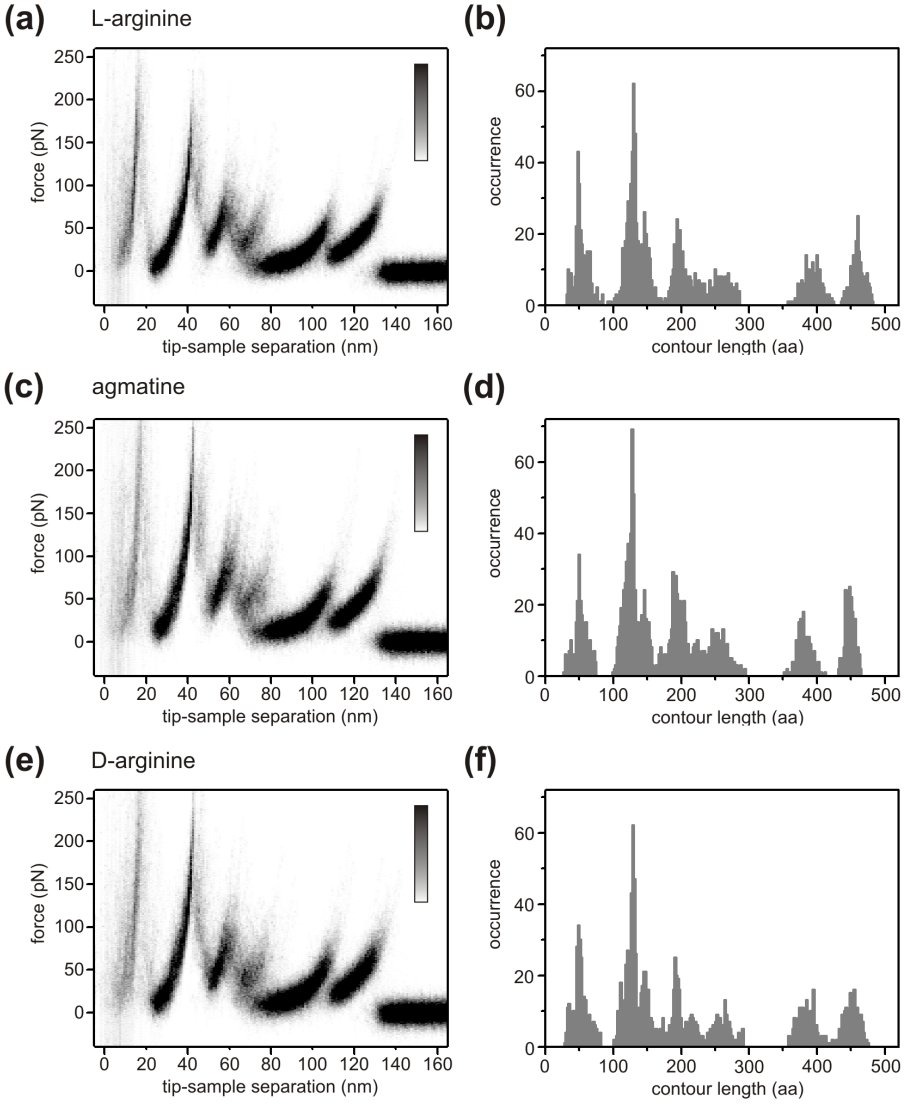

Figure 7. Scattering plot and distribution of contour length values for N-terminal unfolding of N-His_6-AdiC in the presence of 10 mM L-arginine, *n* = 193, panels (a)-(b); 10 mM agmatine, *n* = 197, panels (c)-(d); 10 mM D-arginine, *n* = 196, panels (e)-(f). The contour length values and the corresponding forces obtained from the WLC fit are given in table 3. The gray scale of the scattering plots corresponds to 90 counts.

limit because this *tss*-regime is often obscured by non-specific adhesion between sample and AFM tip. Careful analysis of individual force spectra revealed that the unfolding peak at L_C = 387 aa (*tss* ≈ 110 nm) had a side peak at 404±5 aa with low occurrence. Due to its low probability (~10%) and close proximity to the main peak at 387 aa it can neither be observed in the scattering plot nor in the contour length histogram.

Peak heights being between 210 and 70 pN in average were interpreted as a direct measure of the forces that hold the structural elements together (table 1). The height of the force peaks decreased with increasing number of unfolded elements. Comparison of the scattering plot (Figure 6(c)) with the contour length histogram (Figure 6(d)) revealed that the peak at L_C = 147 aa (*tss* ≈ 45 nm, black arrow) was not clearly visible in the scattering plot. It generally occurred as a side event following the main unfolding peak at L_C = 129 aa. Figure 6(e) shows a scattering plot of 190 aligned force spectra obtained from C-terminal unfolding of AdiC-His_6-C. The peaks in the corresponding contour length histogram (Figure 6(f)) highlight the presence of unfolding

barriers with L_C and force values given in (table 2). The separation between the peaks at 158 aa ($tss \approx 45$ nm) and 315 aa ($tss \approx 90$ nm) was highly conserved, which is reflected by the low standard deviation of only ±8 aa. Unfolding forces of all peaks were in the range of values previously published for other membrane proteins (5, 13, 28, 29, 31). However, unfolding barriers in the front part of F-D spectra were significantly weaker than in the case of N-terminal unfolding of the N-His_6-AdiC construct. The highest force, 135 pN, was measured for the peak with $L_C = 100$ aa ($tss \approx 30$ nm), while the lowest barrier height was 57 pN for $L_C = 424$ aa ($tss \approx 130$ nm). Furthermore, HT-SMFS was used to detect whether these forces in N-His_6-AdiC may change in the presence of L-arginine, D-arginine and agmatine. No major changes were observed in the corresponding scattering plots (Figure 7(a)-(f)). The analysis of occurrence and average forces stabilizing the structural regions unfolded did not reveal major changes either (table 3).

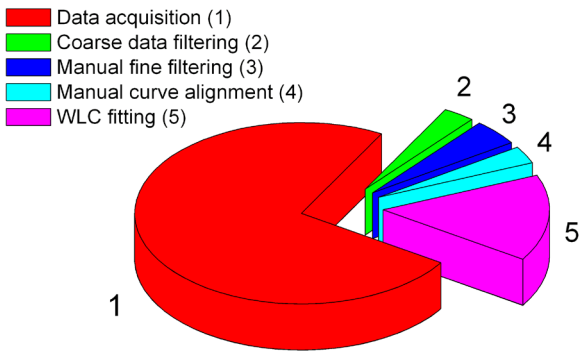

Figure 8. Representation of the relative time required for each step of the here presented HT-SMFS procedure (see Figure 2) when applied to AdiC. As indicated in the chart, most of the time was required for data acquisition. However, it should be considered that during ~90% of the data acquisition time the presence of the operator was not required. Coarse data filtering (2) is completely automated in contrast to steps (3)-(5).

3.5 DISCUSSION

We have introduced and applied a new HT-SMFS procedure to the bacterial proton pump BR and the L-arginine/agmatine antiporter AdiC from *E. coli*. As illustrated in Figure 8, data acquisition required most of the time in our HT-SMFS procedure. However, because of semi-automation, the presence of the operator was required at most for ~10% of data acquisition time. In contrast, pure manual data acquisition would have been an order of magnitude more time consuming and drudgery considering that AdiC attachment to the tip occurred in <1% of the cases under the given experimental conditions. The semi-automated method presented here allowed collecting up to ~40,000 curves per day. The immense amount of data, thus generated, makes an efficient filter for extracting useful force spectra an absolute prerequisite.

Statistical estimates are helpful to plan HT-SMFS experiments so that an appropriate number of F-D curves is available for analysis. The peak probability, p, and its detection at a given certainty, P, define the minimum number of F-D curves, which have to be collected. The number of F-D spectra, n, that have to be recorded so that at least κ force peaks that occur with probability, p, can be found with a certainty $P(k \geq \kappa)$ (k: number of observed force peaks) is calculated from binomial statistics:

$$P(k \geq \kappa) = 1 - \sum_{i=0}^{\kappa-1} \binom{n}{i} p^i (1-p)^{n-i}$$

From numerous published F-D curves, it is reasonable to set up the SMFS experiment such that side peaks, which occur with $p = 10\%$ should be detectable. To find at least one occurrence of such a side peak with a certainty of $P = 95\%$, only $n > 29$ F-D spectra revealing all conserved peaks have to be acquired. To reduce the standard error of the mean for the contour length measurement by a factor of 4, 16 such side peaks should be detected, in turn requiring the acquisition of at

least 227 F-D spectra. Based on this, we included ~200 force spectra of fully unfolded proteins per data set.

In addition to unbiased data acquisition, HT-SMFS considerably reduces the time required to collect a complete data set. This is particularly important when working with non-crystalline samples, e.g. native membranes or reconstituted proteoliposomes, where the density of the protein of interest is smaller than in 2D crystals. Our HT-SMFS method has proven its ability to provide large data sets in a relatively short period of time on non-crystalline samples. Nevertheless, the measurements required continuous operation of the JPK AFM over many hours, and we had to investigate whether long-term drifts of pertinent calibration parameters would introduce uncontrolled experimental errors. As documented in Figure 3, the constant monitoring of *DS* ensured the force measurement to be correctly calibrated throughout the experiment. Interestingly, short-term fluctuations of *DS* were not related to the accuracy of the linear regression, but to fluctuations of the deflection sensor. However, this noise turned out to be negligible compared to the noise introduced by the cantilever itself. The correction of the cantilever *DS* does not yet include the elasticity of the proteoliposomes. The latter over-estimates the *DS* by around 5% compared to measurements on mica (data not shown). However, the resulting difference is constant and smaller than the error from long-term drift (Figure 3(a)). Furthermore, we have validated our approach using bacteriorhodopsin, which yielded the same results like published in earlier studies (6, 13-15). The design of the recombinant target protein is an important issue for successful data acquisition and interpretation. On the one hand, the identification of the terminus from which AdiC was unfolded was made possible by the use of two recombinant forms of AdiC, i.e. N-His_6-AdiC and AdiC-His_6-C, differing in the position of the His_6-tag and the length of the termini. On the other hand, pH-induced protonation of the N- or C-terminally located His_6-tag promoted the interaction between AFM tip and AdiC. Consequently, tags, which are often introduced into recombinant proteins for affinity purification, are also very helpful for the efficient collection of force spectra. Furthermore, we have demonstrated that the direction from which a membrane protein is unfolded can be controlled by the properties of its termini, e.g. length and His_6-tag.

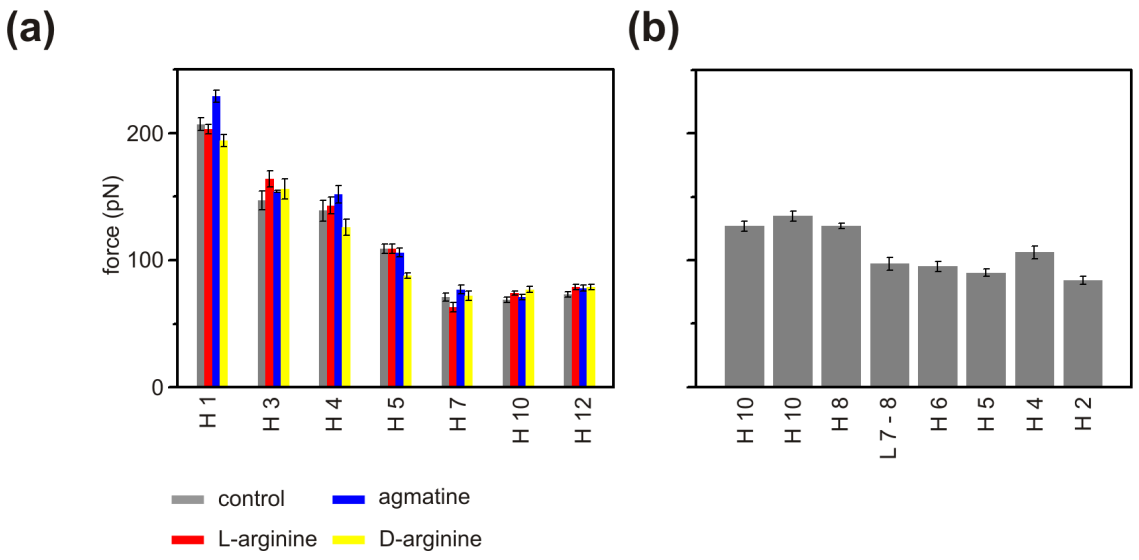

Figure 9. Unfolding forces for the N-His_6-AdiC (a) and AdiC-His_6-C (b) constructs in the absence (control) and presence of substrates (see Figure for substrates). Error bars represent the standard error of the mean. See also tables 1-3.

Mechanical unfolding of AdiC from both termini provided highly reproducible F-D spectra. Since no major differences between unfolding in the absence and presence of substrates were observed (Figures 6, 7 and 9), AdiC SMFS experiments are discussed generally and not with respect to the substrates explored. Comparison of the contour length values of the most prominent force peaks with the working topology model (Figure 10) allowed the mapping of unfolding barriers within the N-His_6-AdiC (Figure 10(a)) and AdiC-His_6-C proteins (Figure 10(b)). For N-His_6-AdiC, the main unfolding barriers were located at the cytoplasmic moieties of helices 1, 3, 5 and 7, at the periplasmic moiety of helix 10 and in the transmembrane part of the predicted helix 12 (see table 1 for details). The peak at L_C = 147 aa (*tss* ≈ 45 nm), which occurred as a side peak of the main peak at L_C = 129 aa (*tss* ≈ 40 nm), indicated an unfolding barrier at the periplasmic moiety of helix 4. This illustrates that helices 3 and 4 do not exclusively unfold pairwise in a single step, but also individually whenever the side peak at 147 aa (*tss* ≈ 45 nm) is observed. Therefore, helices 3 and 4 do not always act as one mechanical unit. According to our SMFS data, helices 7, 8 and 9 and their connecting loops were unfolded in a single step. A consensus amphipathic region (CAR) in the C-terminal half of helix 8 and the loop connecting helices 8 and 9, which are relevant for transport function, has been identified in several APC transporters (32-34). Minchin and McCoubrie proposed that PotE, a member of the same protein family (APC family) as AdiC, contains a re-entrant loop between helices 7 and 8 based on the presence of hydrophobic helix structures, which will preclude complete exposure to the aqueous periplasmic space (35). Thus, possibly, helix 8 is strongly interacting with neighbouring parts of the AdiC protein. Interestingly, joint unfolding of the "three-helix-bundle" formed by helices 7, 8 and 9 was only observed when AdiC was unfolded from its N-terminal end. However, there is an unfolding barrier located in this region, near the cytoplasmic moiety of helix 8, in force spectra from C-terminally unfolded AdiC. The interaction of helices 7 - 9 with helices 10, 11 and 12 may contribute to the stability of the "three-helix-bundle" formed by helices 7 – 9. When the C-terminal part is first removed, the bundle would thus not be stable any more. Moreover, in the C-terminal unfolding of AdiC, several helices unfold in a single step. This supports the hypothesis that the stability of the "three-helix-bundle" is related to interacting partners, i.e. helices 10 – 12, for N-terminal unfolding. This would mean that helices 7 – 9 are intrinsically not a stable unit. However, this interpretation for the joint unfolding of helices 7, 8 and 9 remains to be confirmed, e.g. by the high-resolution structure of AdiC or a related APC family member.

Comparison of N- and C-terminal unfolding provided insights into the topology and stability of AdiC. C-terminal unfolding resulted in main unfolding barriers at the cytoplasmic side of the predicted helices 2, 4, 8 and 10 (see table 2 for details). Several force spectra exhibited a plateau-like unfolding event between L_C = 80 aa (*tss* ≈ 20 nm) and 100 aa (*tss* ≈ 30 nm) (see Figure 5(d)): these L_C values correspond to the cytoplasmic and periplasmic moieties of helix 10 (Figure 10(b)). The plateau is tentatively interpreted as an unresolved force peak, because the force did not drop back as would be expected for the unfolding of a complete structural element (compare with the other force peaks). This assumption is corroborated by the existence of a low-probability side peak (~10%, L_C = 404 aa) of the main unfolding peak with L_C = 387 aa also in helix 10, when N-His_6-AdiC was unfolded from its N-terminal end. Detection and analysis of this low probability peak at L_C = 404 aa was only possible when working with large data sets. This example illustrates the need of large data sets and therefore HT-SMFS procedures to allow the detection of low probability unfolding events.

HT-SMFS experiments did not only yield topological information, but also information on the mechanical stability of structural elements within AdiC. The stability of secondary structure elements is reflected by the height of unfolding peaks. As demonstrated here for AdiC and

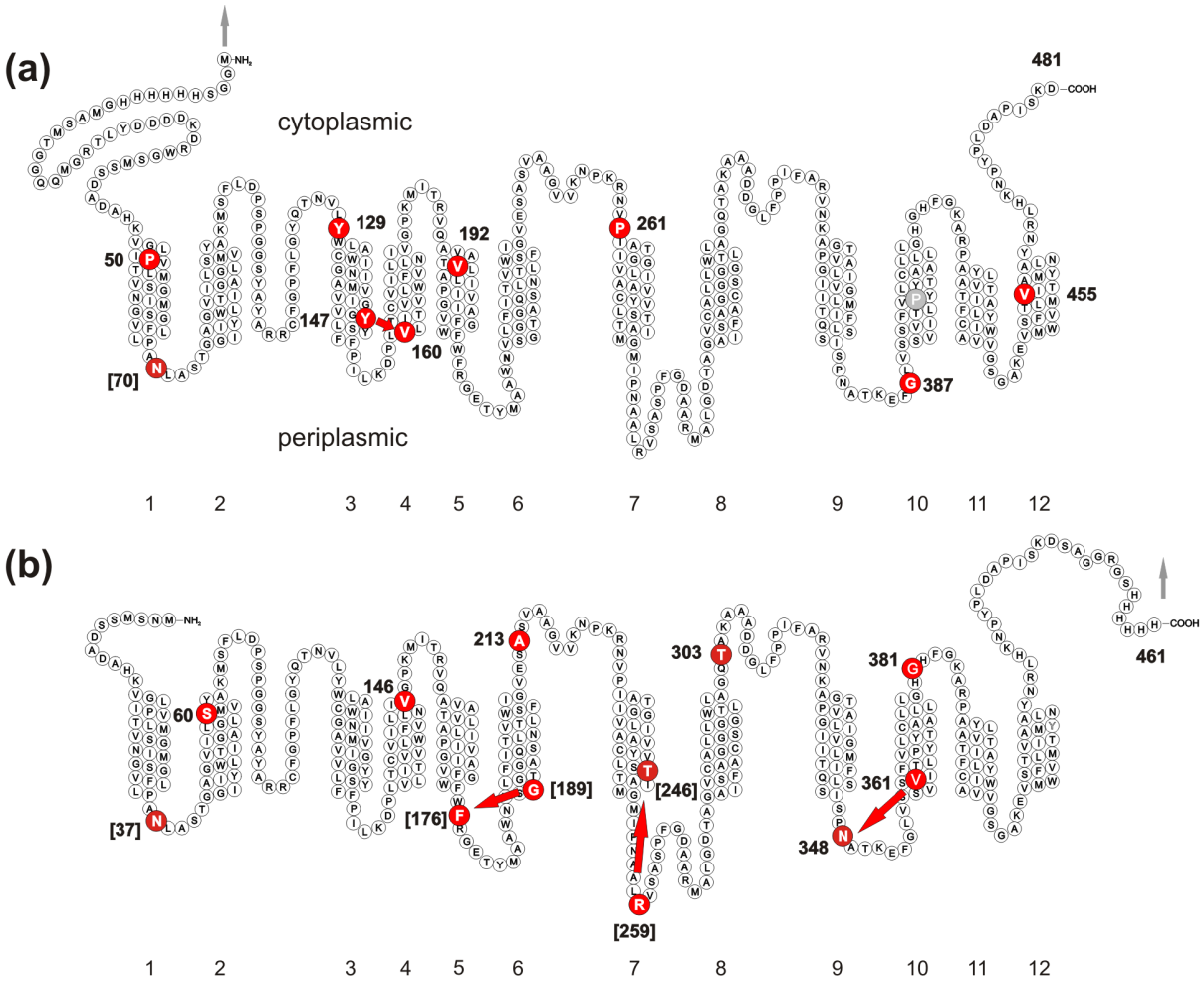

Figure 10. Topology predictions and locations of unfolding barriers for the unfolding of (a) N-His_6-AdiC and (b) AdiC-His_6-C. Residues highlighted by red circles reflect the average position of unfolding barriers (tables 1-3). Numbers in brackets show unfolding barriers that are found at low probabilities. The proline residue in helix 10 (P400 in N-His_6-AdiC) that is highlighted in gray illustrates the location of an unfolding barrier corresponding to a side peak of the main unfolding peak at 387 aa. Gray arrows show the pulling direction. Red arrows show the shift in the position of the unfolding barriers caused by compensation for the thickness of the membrane. The sequence based topology prediction was computed using a hidden Markov model (43) and the AdiC amino acid sequence with the accession number spP60061 supplemented with the residues that were introduced by protein engineering. Loop 2 - 3 is assumed to fold back into the membrane because of AdiC's sequence identity to xCT (35% amino acid identity within this loop), where this has been demonstrated (18).

elsewhere for BR (14, 36), the height of force peaks decreases with the number of unfolded structural elements (Figure 9): for example, the force needed to unfold helix 1 differs whether it is the first or the last structural element to be pulled out of the membrane. In the former case the force is ~2.5 times higher than in the latter case, where helix 1 is the last unfolded structure (see tables 1 and 2 for details). This reflects the high stability of a membrane protein with (nearly) completely established intramolecular interactions. The more interactions are ruptured during unfolding, the lower the stability of the remaining protein. The steepest decrease was observed for N-terminal unfolding of N-His_6-AdiC. Therefore, the force difference between N- and C-terminal unfolding gives a qualitative impression of how strong helix 1 is stabilized by the other elements (5, 36). This observation is corroborated by the three-stage model of membrane protein folding describing the membrane insertion of transmembrane helices as independently stable units which are further condensed and stabilized in their functional 3D structure (37).

Major differences between AdiC F-D spectra recorded in the absence and presence of substrates were not detected (Figures 6 and 7), although substrate concentrations were much higher than the K_d values for L-arginine and agmatine, which are ~100 µM and ~30 µM, respectively (19). Generally, molecules do not bind very strongly to a transporter since they must be released from the protein after transport. Often, substrates are bound to the protein by hydrogen bonds (e.g. in the lactose permease LacY (38)) whose rupture force are weak. If significant interactions between substrates and AdiC are located somewhere in helices 7 – 9 and their connecting loops, they would not be detected since these structural elements are pulled out of the membrane in a single step. It is likely that at least one of these structural elements interacts with the substrates, because several residues in this region have been shown to be essential for transport activity or relevant for substrate interaction in certain members of the APC family (e.g. CadB, PotE and xCT) (35, 39-41). It is possible that dynamic HT-SMFS, which probes the stability of molecular interactions at different time scales, may give insights into how the energy barriers, stabilizing structural regions in AdiC, change their position, height or width (42). Such insights into the energy landscape of AdiC may reveal how substrate binding favours or disfavours functionally related conformations of the transporter.

3.6 SUMMARY AND PERSPECTIVES

With BR and AdiC we demonstrate that SMFS data acquisition and pre-processing can be automated to a large extent. Our procedure was used to acquire a data set comprising ~400,000 F-D curves, which were corrected for variations in cantilever *DS* and filtered by a straightforward algorithm to select ~4,000 F-D curves. Subsequent manual selection yielded ~1,400 interpretable spectra with high efficiency. This semi-automated data acquisition and processing protocol allowed a relevant data set to be collected ideally within a single day using a single sample of the reconstituted target protein. Therefore, HT-SMFS can be performed speedily and with small amounts of sample. Although we applied coarse filtering to data acquired with an *x* and *y* closed-loop controlled piezo-scanner, this procedure can also be applied to F-D curves recorded with an open-loop piezo-scanner.

As demonstrated with AdiC, HT-SMFS made the acquisition of AdiC unfolding spectra in the absence and presence of different compounds possible. Considering the wealth of data collected and the lack of any significant difference in the force spectra between different conditions suggest that AdiC and the substrates tested interact only weakly. Our HT-SMFS approach can be generally applied to a wide range of reconstituted or even native membrane proteins and specific mutants to study their fold as well as ligand binding, and

Table 1. Analysis of main conserved unfolding curves of the N-His_6-AdiC construct without substrate.

L_C (aa, ave±SD)	Probability (%)	Unfolding force (pN, ave±SE)	Topological location	*tss* (nm)
50±5	52	208±10	Helix 1 (c)	15
129±5	100	147±7	Helix 3	40
147±12	55	139±8	Helix 4 (p)*	45
193±10	77	87±3	Helix 5 (c)	60
261±30	70	71±3	Helix 7	75
387±24	100	69±2	Loop 9 – 10	110
455±17	100	73±2	Helix 12 (m)	130

Abbr.: cytoplasmic part (c), intramembranous part (m), periplasmic part (p), average (ave), tip-sample separation (*tss*). SE is the standard error of the mean, SD the standard deviation (n = 205). * Topological position was corrected for the thickness of the membrane.

Table 2. Analysis of main conserved unfolding curves of the AdiC-His$_6$-C construct without substrate.

L_C (aa, ave±SD)	Probability (%)	Unfolding force (pN, ave±SE)	aa from N-terminus (ave±SD)	Topological location	*tss* (nm)
80±20	100	127±4	381±20	Loop 10 – 12	20
100±17	100	135±4	361±17	Loop 9 – 10*	30
158±8	100	127±2	303±8	Loop 8 – 9	45
202±6	25	97±5	259±6	Helix 7 (p)*	55
248±8	26	95±4	213±8	Loop 6 – 7	N.V.
272±22	65	90±3	189±22	Loop 5 - 6 (p)*	70
315±8	100	106±2	146±8	Helix 4	80
401±24	100	84±3	60±24	Helix 2 (c)	110
424±2	7	57±4	37±2	Loop 1 – 2	130

Abbr.: cytoplasmic part (c), intramembranous part (m), periplasmic part (p), average (ave), tip-sample separation (*tss*). SE is the standard error of the mean, SD the standard deviation (n = 190). N.V.: not visible in the scattering plot. * Topological position was corrected for the thickness of the membrane.

Table 3. Analysis of main conserved unfolding events of N-His$_6$-AdiC in the presence of 10mM substrate.

L_C (aa, ave±SD)	Probability (%)	Unfolding force (pN, ave±SE)	Topological location	*tss* (nm)
L-arginine (n = 193)				
49±7	88	191±7	Helix 1 (c)	15
129±6	100	140±3	Helix 3	40
148±12	70	140±3	Helix 4 (p)*	45
196±14	77	88±4	Helix 5 (c)	60
262±30	70	63±4	Helix 7	75
394±26	90	65±2	Helix 10 (p)	110
459±17	100	73±2	Helix 12 (c)	130
agmatine, (n = 197)				
50±8	67	243±10	Helix 1 (c)	15
129±7	100	154±4	Helix 3	40
145±14	69	152±7	Helix 4 (p)*	45
191±14	82	87±3	Helix 5 (c)	60
257±26	81	77±3	Helix 7	75
381±20	90	71±2	Loop 10 – 11	110
448±14	100	78±2	Helix 12 (p)	130
D-arginine (n = 196)				
50±7	70	258±6	Helix 1 (c)	15
129±6	100	150±7	Helix 3	40
145±12	65	141±7	Helix 4 (p)*	45
193±9	69	102±5	Helix 5 (c)	60
259±27	65	71±4	Helix 7	75
384±28	90	73±3	Loop 10 – 11	110
451±23	100	79±2	Helix 12 (m)	130

Abbr.: cytoplasmic part (c), intramembranous part (m), periplasmic part (p), average (ave), tip-sample separation (tss). SE is the standard error of the mean (n = see corresponding value in table).
* Topological position was corrected for the thickness of the membrane.

specifically to transport proteins for assessing their interaction with substrates or inhibitors, and their behaviour under different environmental conditions such as pH, ionic strength and temperature. The presented semi-automated HT-SMFS procedure can also be adapted for dynamic SMFS experiments in order to explore the energy landscape of membrane proteins.

3.7 ACKNOWLEDGMENTS

The authors are thankful to Dr. Georg Büldt, Jülich, Germany, for the kind gift of purple membranes, and to Susanna Bial for technical assistance in preparing *E. coli* membranes expressing tagged versions of AdiC. This work was supported by the Maurice E. Müller Foundation of Switzerland, by the Berne University Research Foundation (to D.F.), by the Swiss National Center of Competence in Research on Nanoscale Science (to A.E.), by the EC Project 29084 NANOMOT (to A.E. and D.J.M.), by the EC Project Grant 502802 EUGINDAT (to D.F. and M.P.) and by the Spanish Ministry of Science and Education Grant BFU2006-14600 (to M.P.). M.R. is a recipient of a pre-doctoral fellowship from the Spanish Ministry of Science and Education. D.J.M. acknowledges the financial support of the BMBF and the DFG. The used AFM-facility was built up with contributions of the Swiss University Conference and JPK-Instruments AG, Berlin, Germany.

3.8 REFERENCES

1. Engel A. and Gaub HE. (2008) *Ann Rev Biochem* **77** 1-22
2. Kedrov A., Janovjak H., Sapra KT. and Mueller DJ. (2007) *Ann Rev Biophys Biomol Struct* **36** 233-60
3. Janovjak H., Sapra KT., Kedrov A. and Mueller DJ. (2008) *Chem Phys Chem* (in press)
4. Sapra KT., Besir H., Oesterhelt D. and Mueller DJ. (2006) *J Mol Biol* **355** 640-50
5. Kedrov A., Ziegler C., Janovjak H., Kuehlbrandt W. and Mueller DJ. (2004) *J Mol Biol* **340** 1143-52
6. Janovjak H., Kessler M., Gaub H., Oesterhelt D. and Mueller DJ. (2003) *EMBO J* **22** 5220-29
7. Mueller DJ., Sapra KT., Scheuring S., Kedrov A., Frederix PLT. M., Fotiadis D. and Engel A. (2006) *Curr Opin Struct Biol* **16** 489-95
8. Kuhn M., Janovjak H., Hubain M. and Mueller DJ. (2005) *J Microsc* **218** 125-32
9. Marsico A., Labudde D., Sapra T., Mueller DJ. and Schroeder M. (2007) *Bioinformatics* **23** 231-36
10. Dietz H. and Rief M. (2007) *Jap J Appl Phys* **46** 5540-42
11. Kimura Y., Vassylyev DG., Miyazawa A., Kidera A., Matsushima M., Mitsuoka K., Murata K., Hirai T. and Fujiyoshi Y. (1997) *Nature* **389** 206-11
12. Essen L-O., Siegert R., Lehmann WD. and Oesterhelt D. (1998) *Proc Natl Acad Sci USA* **95** 11673-78
13. Oesterhelt F., Oesterhelt D., Pfeiffer M., Engel A., Gaub HE. and Mueller DJ. (2000) *Science* **288** 143-6
14. Mueller DJ., Kessler M., Oesterhelt F., Moeller C., Oesterhelt D. and Gaub HE. (2002) *Biophys J* **88** 3578-88
15. Janovjak H., Struckmeier J., Hubain M., Kedrov A., Kessler M. and Mueller DJ. (2004) *Strucure* **12** 871-9
16. Foster JW. (2004) *Nat Rev Micro* **2** 898-907
17. Jack DL., Paulsen IT. and Saier MH. (2000) *Microbiology* **146** 1797-814
18. Gasol E., Jimenez-Vidal M., Chillaron J., Zorzano A. and Palacin M. (2004) *J Biol Chem* **279** 31228-36
19. Fang Y., Kolmakova-Partensky L. and Miller C. (2007) *J Biol Chem* **282** 176-82
20. Reig N., del Rio C., Casagrande F., Ratera M., Gelpi JL., Torrents D., Henderson PJF., Xie H., Baldwin SA., Zorzano A., Fotiadis D. and Palacin M. (2007) *J Biol Chem* **282** 13270-81
21. Xie H., Patching SG., Gallagher MP., Litherland GJ., Brough AR., Venter H., Yao SY., Ng AM., Young JD., Herbert RB., Henderson PJ. and Baldwin SA. (2004) *Mol Membr Biol* **21** 323-336
22. Butt H-J. and Jaschke M. (1995) *Nanotechnology* **6** 1-7
23. Engel A. and Mueller DJ. (2007) *Nat Prot* **2** 2191-7
24. Fotiadis D., Scheuring S., Mueller SA., Engel A. and Mueller DJ. (2002) *Micron* **33** 385-97
25. Bustamante C., Marko JF. and Smith ED. (1994) *Science* **265** 1599-600
26. Rief M., Gautel M., Oesterhelt F., Fernandez JM. and Gaub HE. (1997) *Science* **276** 1109-12
27. Moeller C., Fotiadis D., Suda K., Engel A., Kessler M. and Mueller DJ. (2003) *J Struct Biol* **142** 369-78
28. Cisneros DA., Oesterhelt D. and Mueller DJ. (2005) *Structure* **13** 235-42
29. Klyszejko AL., Shastri S., Mari SA., Grubmueller H., Mueller DJ. and Glaubitz CL. (2008) *J Mol Biol* **376** 35-41
30. Butt H-J. (1992) *Biophys J* **63** 578-82
31. Kedrov A., Wegmann S., Smits SHJ., Goswami P., Baumann H. and Mueller DJ. (2007) *J Struct Biol* **159** 290-301
32. Closs EI., Lyons CR., Kelly C. and Cunningham JM. (1993) *J Biol Chem* **268** 20796-20800
33. Hu LA. and King SC. (1998) *Biochem J* **336** 69-76
34. King SC., Hu LA. and Pugh A. (2003)

Biochem J **376** 645-653
35. Minchin RF. and McCoubrie JE. (2004) *Int J Biochem Cell Biol* **36** 271-80
36. Kessler M. and Gaub HE. (2006) *Structure* **14** 521–7
37. Engelman DM., Chen Y., Chin C-N., Curran AR., Dixon AM., Dupuy AD., Lee AS., Lehnert U., Matthews EE., Reshetnyak YK., Senes A. and Popot J-L. (2003) *FEBS Letters* **555** 122-5
38. Abramson J., Smirnova I., Kasho V., Verner G., Kaback HR. and Iwata S. (2003) *Science* **301** 610-5
39. Kashiwagi K., Kuraishi A., Tomitori H., Igarashi A., Nishimura K., Shirahata A. and Igarashi K. (2000) *J Biol Chem* **275** 36007-12
40. Soksawatmaekhin W., Uemura T., Fukiwake N., Kashiwagi K. and Igarashi K. (2006) *J Biol Chem* **281** 29213-20
41. Jimenez-Vidal M., Gasol E., Zorzano A., Nunes V., Palacin M. and Chillaron J. (2004) *J Biol Chem* **279** 11214-11221
42. Kedrov A., Appel M., Baumann H., Ziegler C. and Mueller DJ. (2008) *J Mol Biol* **375** 1258-1266
43. Krogh A., Larsson B., von Heijne G. and Sonnhammer ELL. (2001) *J Mol Biol* **305** 567-80

4. Functional and Structural Characterization of the first Prokaryotic Member of the LAT Family: a Model for APC Transporters

Núria Reig[1*], César del Rio[1*], Fabio Casagrande[2*], Mercè Ratera[1], Josep Lluís Gelpí[1], David Torrents[4], Peter JF. Henderson[5], Hao Xie[5], Stephen A. Baldwin[5], Antonio Zorzano[1], Dimitrios Fotiadis[2] and Manuel Palacín[1,6]

[1] Institute for Research in Biomedicine, Barcelona Science Park, Department of Biochemistry and Molecular Biology, Faculty of Biology, University of Barcelona, E-08028 Barcelona, Spain

[2] M. E. Müller Institute for Structural Biology, Biozentrum of the University of Basel, CH-4056 Basel, Switzerland

[3] Molecular Modeling and Bioinformatics, Institute for Research in Biomedicine, E-08028 Barcelona, Spain

[4] Barcelona Supercomputing Center, Life Science Program-ICREA, E-08028 Barcelona, Spain

[5] Institute of Membrane and Systems Biology, University of Leeds, LS29JT Leeds, UK

[6] Centre for Biomedical Research on Rare Diseases (CIBERER), E-08028 Barcelona, Spain

*) These three authors contributed equally to this work

4.1 ABBREVIATIONS

BN-PAGE, blue native-polyacrylamide gel electrophoresis; **DDM**, n-dodecyl-ß-D-maltopyranoside; **DM**, n-decyl-ß-D-maltopyranoside; **DMS**, dimethyl suberimidate; **BM(PEO)4**, bis-maleimidotetraethylene glycol; **DSP**, dithiobis(succinimidyl)propionate; **DST**, disuccinimidyl tartarate; **MTSET**, 2-(trimethylammonium) ethylmethanethiosulfonate bromide; **OG**, ß-D-octylglucoside; **PL**, proteoliposome; **TEM**, transmission electron microscopy.

4.2 ABSTRACT

We have identified YkbA from *Bacillus subtilis* as a novel member of the LAT family of amino acid transporters. The protein is ~30% identical in amino acid sequence to the light subunits of human Heteromeric Amino acid Transporters (HATs). Purified His-tagged YkbA from *Escherichia coli* membranes reconstituted in proteoliposomes exhibited sodium-independent, obligatory exchange activity for L-serine and L-threonine, and also for aromatic amino acids, albeit with less activity. Thus, we propose that YkbA be renamed SteT (Ser/Thr Exchanger Transporter). Kinetic analysis supports a sequential mechanism of exchange for SteT. Freeze-fracture analysis of purified, functionally active SteT in proteoliposomes, together with blue native-polyacrylamide gel electrophoresis and transmission electron microscopy (TEM) of detergent-solubilized purified SteT, suggest that the transporter exists in a monomeric form. Freeze-fracture analysis showed spherical particles of diameter 7.4 nm. TEM revealed elliptical particles (diameters: 6 x 7 nm) with a distinct central depression. To our knowledge this is the first functional characterization of a prokaryotic member of the LAT family and the first structural data on an APC transporter. SteT represents an excellent model to study the molecular architecture of the light subunits of HATs and other APC transporters.

4.3 INTRODUCTION

The APC (amino acids, polyamines and choline for organocations) superfamily of transport proteins includes nearly 250 members that function as solute:cation symporters and solute:solute antiporters ((1) ; see also http://www.tcdb.org). They occur in all phyla from prokaryotes to higher eukaryotes, and vary in length between 350-850 amino acid residues. The smaller proteins are generally of prokaryotic origin while the larger ones are of eukaryotic origin and have N- and C-terminal hydrophilic extensions. Most of APC members are predicted to possess 12 transmembrane (TM) α-helical domains.

The LAT (L-amino acid transporter) family belongs to the APC superfamily. LAT family members correspond to the light subunits of the Heteromeric Amino acid Transporters (HATs), also called glycoprotein-associated Amino acid Transporters (gpaAT) (2-4). HATs are composed of two subunits, a polytopic membrane protein (the light subunit) and a disulfide-linked N-glycosylated type II membrane glycoprotein (the heavy subunit). The light subunit is the catalytic component of the transporter, whereas the heavy subunit appears to be essential only for trafficking to the plasma membrane. Two types of heavy subunit (4F2hc and rBAT) and 10 types of light subunit have so far been identified. A number of human pathologies have highlighted the physiological roles of HATs. For example, two transporters of this family are responsible for inherited aminoacidurias: mutations in any of the two genes coding for the subunits of system $b^{0,+}$ ($b^{0,+}$AT and rBAT) lead to cystinuria (MIM 220100) (5,6), whereas mutations in y^+LAT1 (a 4F2hc-associated system y^+L) result in lysinuric protein intolerance (LPI) (MIM222700) (7,8). In addition, xCT, a LAT transporter that in association with 4F2hc mediates cystine uptake and glutamate efflux (9,10), has been recently identified as the receptor of Kaposi's sarcoma-associated Herpesvirus (KSHV, human herpesvirus 8) (11). *In vivo*, this transport system is involved in cocaine relapse through the control of the basal levels of extrasynaptic glutamate (12), and it contributes to the maintenance of the plasma redox balance (13).

Despite the important roles attributed to HATs, only a few studies have addressed the structure-function relationships of these transporters: i) most HATs are obligate antiporters with a 1:1 stoichiometry (14), and a sequential mode of exchange has been proposed for system $b^{0,+}$ (15); ii) light subunits appear to be sufficient for

transport activity, as demonstrated for $b^{0,+}$AT (16); iii) using xCT as a model for the light subunits, a membrane topology with 12 transmembrane segments and with a re-entrant loop between transmembrane segments 2 and 3 has been reported (17); and iv) the xCT residues His110 and Cys327 have been shown to be crucial for function (17,18) while the cystinuria-specific mutation Ala354Thr inactivates $b^{0,+}$AT (16). Similarly, structure-function studies on the APC superfamily as a whole have been very limited, and primarily related to membrane topology studies and the identification of relevant residues for substrate interaction (19-21).

The recent elucidation of atomic structures for several prokaryotic transporters has given key insights into the molecular dynamics of fundamental transport processes. Towards a similar increase in our understanding of amino acid transport, we describe here the identification and characterization, at the functional and structural levels, of the first example of a prokaryotic member of the LAT family: the orphan transporter YkbA (from here on referred to as SteT, see below) from *B. subtilis*. SteT shows significant similarity to the light subunits of eukaryotic HATs not only in amino acid sequence and putative membrane topology, but also in the characteristic sequential mode of obligatory exchange. Structural analyses of SteT revealed a monomeric structure with a "donut-like shape" of 7 x 6 nm in diameters and a distinct central depression. SteT represents an excellent model for structural studies of the LAT family, and of APC transporters in general.

4.4 MATERIALS AND METHODS

Identification and sequence analysis of SteT – Initial sequence analysis that led to the identification of prokaryotic homologs of LAT family transporters, was performed by using the BLAST algorithm (default settings; (22)) to compare all known members of the LAT family against a non-redundant protein sequence database (non-redundant GenBank CDS translations + RefSeq Proteins + PDB + SwissProt + PIR + PRF). In order to study the phylogenetic relationship of SteT and PotE relative to the APC superfamily of transporters (23), we first constructed a multiple protein sequence alignment of all known APC members (listed at http://www.tcdb.org/tcdb/superfamily.php) with test sequences using ProbCons (24) in combination with manual refinements. All APC members that appeared to be too distant to provide information (i.e. that could not be properly aligned) were not further considered. The resulting alignment was further evaluated with G-BLOCKS (25) in order to select the most conserved, and hence, most informative regions of the multiple alignment. The latter were subsequently processed with CLUSTAL (26) to obtain a neighbor-joining phylogenetic tree (1000 bootstraps with 111 random seeds).

Cloning of prokaryotic APC transporters – Genomic DNA from *E. coli* strain DH5α and from *B. subtilis* strain 168t+ was prepared from cells collected after an overnight liquid culture. Cell pellets were incubated for 1 h at 50ºC in 10 mM Tris-HCl pH 8, 0.1 M EDTA, 0.5% SDS with 0.1 mg/ml of Proteinase K. Three extractions with the same volume of phenol were performed with centrifugations at 5,000 *g*. Next, a chloroform/ethanol precipitation gave final pellets which were resuspended in 10 mM Tris-HCl pH 8.0 and 1 mM EDTA. The following primers (5′-3′) were used to amplify open reading frames encoding the indicated proteins from genomic DNA by PCR: ATCTGAATTCTGATGCGCAAAGCACCCT-GTT and CATTCTCGAGAGAAAGGGCGA-TCATTCAATC for PotE, and ATCTGAA-TTC-TCCTCCACATTACATAACATCA and CATTCTCGAGAGTCCACGGTGCTTTTA-TCAAT for YhfM from *E. coli*; and ATCAGAATTCGAAGCTTCAACATCA-TAGGAG and CTATCTCGAGTCCTGTCAA-CTTT-TATCTTCTG for SteT and ATGTGA-ATTCCAATATAATACA-ACAAGAACTGC and TCTACTCGAGTATCGCTTCATCTGTG-TGTC for YfnA from *B. subtilis*. PCR products were digested with *Eco*RI and *Xho*I and ligated

into a pBlueScript vector (Stratagene). Another PCR with the following primers (5′-3′) was performed to subclone each transporter open reading frame into the *Eco*RI and *Pst*I sites of a modified version of the vector pTTQ18 (27), thereby placing its expression under the control of the *tac* promoter: TGATGAATTCGATGAGTCAGGCTAAATC and ACCGCCTGCAGAACCGTGTTTATTTTTCAGT for PotE, TGATGAATTCGATGACTGCAAACTCTCCCCTA and ACCGCCTGCAGACGACAAAGCGTTGAGCTGGC for YhfM, TGATGAATTCGATGCATACTGAAGACAACGG and ACCGCCTGCAGAGCTTGCTTTT-CGTTTTTTCA for SteT and CGATGAATTCGATGAGTTCATTATTTAGA and ACCGCCTGCAGATTTGTTTAATAAGCTGTGTT for YfnA. All DNA constructs were verified by sequencing.

Expression of prokaryotic APC transporters in E. coli – Expression experiments were carried out with freshly transformed cultures of *E. coli* strain BL21(DE3). Initially, the expression of the four APC transporters (PotE, YhfM, SteT and YfnA) was investigated in 50 ml samples cultured in LB medium containing 50 µg/ml ampicillin. When the OD_{600} had reached 0.5, protein expression was induced by addition of 0.5 mM isopropyl-ß-D-thiogalactoside (IPTG; Roche). To estimate protein expression cells were harvested after incubation for different time periods (1, 2, 3, 4 h and overnight) at 30ºC or 37ºC. For biochemical or functional analyses of SteT and PotE, either 6 x 0.8 L of medium in flasks or 10 L of medium in a fermentor were inoculated with *E. coli* cultures harbouring pTTQ18-His_6-YkbA or pTTQ18-His_6-PotE. Cells were induced with 0.5 mM IPTG at an OD_{600} of 0.5 and harvested after 3 h growth at 30ºC and 37ºC, respectively. In all cases, the cell pellet was resuspended in Lysis buffer (20 mM Tris-HCl pH 8, 0.5 mM EDTA) and stored frozen at -20ºC.

Preparation of E. coli membranes - Cell pellets were thawed and disrupted by passage through a French pressure cell (20,000 psi, three times). Unbroken cells were removed by centrifugation (10 min at 10,000 *g* and 4ºC). The supernatant was ultracentrifuged (1 h at 100,000 *g* and 4ºC), the pellet resuspended and homogenised (30 ml glass homogenizer for 2 min) in Lysis buffer and ultracentrifuged again. Peripheral membrane proteins were removed by homogenization in 20 mM Tris-HCl pH 8, 300 mM NaCl and ultracentrifugation. Finally, the membrane pellet was resuspended in 20 mM Tris-HCl pH 8, 150 mM NaCl at a protein concentration between 13-25 mg/ml. Aliquots were frozen in liquid nitrogen and stored at -80ºC until use.

Purification of APC transporters – To estimate protein expression and functional reconstitution of SteT and PotE, frozen membranes were thawed and solubilized in 0.5% DDM (n-dodecyl-ß-D-maltopyranoside; Anatrace, USA), 20 mM Tris-HCl pH 8, 20% glycerol, 50 mM NaCl on a roller shaker (1 h, 4ºC) at a protein concentration of ~2 mg/ml. The supernatant after ultracentrifugation (1 h at 100,000 *g*, 4ºC) was incubated for 2 h at 4ºC with equilibrated Ni-NTA Superflow beads (Qiagen) with washing buffer (20 mM Tris-HCl pH 8, 20% glycerol, 200 mM NaCl, 0.05% DDM, 10 mM imidazole). The supernatant fraction (slurry) was removed by centrifugation (1 min at ~160 *g*). Protein-bound beads were washed three times with 10 ml of washing buffer and centrifuged as before. Then, columns were packed with 5 ml protein-bound beads each. Two more washes were performed with 15 ml washing buffer before elution with 10 ml of elution buffer (washing buffer supplemented with 500 mM imidazole). The purified protein was concentrated by centrifugation in an Amicon Ultra (10,000 MWCO Millipore) at 3,220 *g* down to a volume of 1 ml. Imidazole was removed by addition of 10 ml of 20 mM Tris-HCl pH 8, 20% glycerol, 200 mM NaCl, 0.05% DDM and re-concentration to the desired final volume.

For negative stain transmission electron microscopy studies, frozen SteT-containing membranes were thawed and solubilized for 1 h

at 4 °C in 1% DM (n-decyl-ß-D-maltopyranoside; Anatrace, USA), 20 mM Tris-HCl pH 8, 300 mM NaCl, 10% glycerol, 0.01% NaN_3. The protein concentration during solubilization was between 2-3 mg/ml. After ultracentrifugation (100,000 *g*, 45 min at 4°C) the supernatant was diluted two-fold with 20 mM Tris-HCl pH 8, 300 mM NaCl, 0.3% DM, 3 mM histidine, 10% glycerol, 0.01% NaN_3 (washing buffer) and bound for 2 h at 4 °C to Ni-NTA Superflow beads (Qiagen). The beads were then loaded onto a spin column (Promega), washed with washing buffer and eluted with the same buffer containing 200 mM histidine.

Reconstitution into proteoliposomes – *E. coli* polar lipid extract (Avanti Polar Lipids) solubilized in chloroform (50 mg/ml) was dried under a stream of nitrogen to remove the solvent and to obtain a thin layer of dry lipids in a glass tube. The dried lipids were resuspended in dialysis buffer (120 mM KPi pH 7.4, 0.5 mM EDTA, 1 mM $MgSO_4$, 5 mM $TrisSO_4$, 1% glycerol and 4 mM of the desired amino acid, unless otherwise indicated) to yield a final lipid concentration of 40 mg/ml. After four 30 s sonication and vortexing cycles, the liposomes were extruded in a LiposoFast-Pneumatic Actuator (Avestin) through a 400 nm polycarbonate filter (Avestin) to obtain unilamellar vesicles of homogeneous size. Liposomes were mixed with purified protein at a 1:100 (occasionally 1:40) protein-to-lipid ratio (w/w). To destabilize the liposomes, 1.25% β-D-octylglucoside (OG; Roche) was added and incubated in ice with occasional agitation for 5 min. DDM and OG were removed by dialysis for 40 h at 4ºC against 100 volumes of dialysis buffer. Finally, proteoliposomes were ultracentrifuged (100,000 *g*, 1 h at 4ºC) and the pellet resuspended in a third of the initial volume of dialysis buffer without amino acids.

Transport measurements – Influx measurements in proteoliposomes were made as described (16) with minor changes. Cold proteoliposomes (10 µl) were mixed with 180 µl of transport buffer (150 mM choline chloride, 10 mM Tris-HEPES pH 7.4, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 0.5 µCi radiolabeled L-amino acid, and unlabeled amino acid to the desired final concentration) and incubated at room temperature for different periods of time. To test the effect of an imposed membrane potential, 2.8 µM valinomycin (Sigma) was added to the transport buffer. Reactions were stopped by the addition of 850 µl of ice-cold stop buffer (150 mM choline chloride, 10 mM Tris-HEPES and 5 mM L-serine or putrescine for SteT- and PotE-containing proteoliposomes, respectively) and filtration through membrane filters (Sartorius, 0.45 µm pore size). Filters were then washed three times with 2 ml of stop buffer and dried, and the trapped radioactivity was counted. All experimental values were corrected by subtracting zero-time values obtained by adding the stop solution before the proteoliposomes into the transport buffer. The following radiolabeled compounds (American Radiolabelled Chemicals) were used in this study: L-[^{3}H] serine, [^{3}H] putrescine, L-[^{3}H] ornithine, L-[^{3}H] arginine, L-[^{3}H] lysine, [^{3}H] glycine, L-[^{3}H] proline, L-[^{3}H] leucine, L-[^{3}H] isoleucine, L-[^{3}H] methionine, L-[^{3}H] glutamate and L-[^{3}H] alanine.

For efflux measurements, proteoliposomes (400 µl) were mixed with 200 µl of three-fold concentrated transport buffer (450 mM choline chloride, 30 mM Tris-HEPES pH 7.4, 3 mM $MgCl_2$, 3 mM $CaCl_2$, 35 µCi L-[^{3}H] serine at a final concentration of 10 µM) and incubated at room temperature for 2 h. Then, the proteoliposome suspensions were divided into two halves (295 µl each) and these were diluted 13-fold with transport buffer (without L-[^{3}H] serine) with or without 7 mM L-serine. At the indicated times, aliquots (195 µl) were mixed with ice-cold stop buffer (850 µl), vortexed and filtered (Sartorius, 0.45 µm pore size). Filters were then washed three times with 2 ml of stop buffer, dried and the trapped radioactivity counted.

Simulation of SteT exchange activity – Simulation of the time-course of transport of radiolabeled L-serine into proteoliposomes containing purified SteT (SteT-PLs) was performed using a model based on the following premises: 1) The induced

amino acid transport activity is an obligatory exchange process with a 1:1 stoichiometry. 2) An additional diffusive, protein-independent flux is necessary to explain transport observed in the absence of transporter (i.e., in proteoliposomes containing no SteT), or in the absence of exchangeable amino acids (SteT-PLs containing L-arginine instead of L-serine). The computer program, previously designed to simulate transport in oocytes (28) and membrane vesicles (15), was adapted to model amino acid exchange and simulated the experimental influx and efflux rates governed by a concerted (i.e., sequential) mechanism of exchange, as described in a previous study for system $b^{0,+}$ (15). Different types of mechanisms (e.g., ping-pong), however, would produce similar results (data not shown). The computer program is available upon request. The simulations were set up using the following rules: i) Characteristics of the simulated system. The experimental setup was reproduced as two separate compartments. The outside compartment volume was set to 86 µl (taking 1 µg protein as reference). The inside volume, 62 nl, together with the Fick parameter, was deduced from extrapolation to infinity of non-specific influx measurements (SteT-PLs containing L-arginine), fitted to the Fick law by standard non-linear regression. ii) Kinetic parameters: Experimental parameters were used when available. Best values for the unknown parameters and transporter concentration were determined after a systematic search of the appropriate range of values. iii) Simulation procedure: Transport rates were evaluated from the relative concentrations of transporter complexes. Transport rates were estimated on the assumption that the limiting step is the translocation of the transporter bound to the substrate, as described for system y^+L (29), and therefore all binding steps were considered as being at equilibrium. Possible inactivation of the transporter was included as an exponential first order law with the appropriate parameters. During simulation amino acid concentrations were calculated by numerical integration of the transport rates. Integration was performed with a constant time step of 0.02 min during the desired period.

Freeze fracture and electron microscopy of proteoliposomes – Freeze-fracture electron microscopy study was performed as described (30). The suspension was sandwiched between two copper platelets using a 400-mesh gold grid as spacer. The samples were frozen by liquid propane immersion, at -189ºC, and fractured at -150ºC and 10^{-8} mbar in a BAL-TEC BAF 060 freeze-etching system. The replicas were obtained by unidirectional shadowing at 45º with 2 nm of Pt/C and at 90º with 20 nm of C, and subsequently floated on distilled water for 5 min. Electron micrographs (at 50,000x) were recorded in a Jeol 1010 electron microscope operated at 80 kV.

Measurement of freeze-fracture particles – Particle diameters were measured from scanned electron micrographs using AnalySIS software. The diameter was obtained by measuring the width of the particle edge-to-edge in a direction perpendicular to the direction of the shadow. Duplicate measurements of individual particles of SteT and PotE revealed a ± 0.2 nm error of measurement. The accuracy of the diameter measurements was tested by measuring the diameter of 10-nm gold particles (Chemicon) placed directly on Formvar-coated copper grids (9.94 ± 0.04 nm; n = 125). All values are reported as mean ± standard error of the mean (SEM).

Negative stain TEM – DM-solubilized SteT protein as eluted from the Ni-NTA column was adsorbed for 10 s to parlodion carbon-coated copper grids rendered hydrophilic by glow discharge at low pressure in air. Grids were washed with four drops of double-distilled water and stained with 2 drops of 0.75% uranyl formate. This washing step is to effectively remove the buffer solution and not adsorbed protein. The former, if not removed, can lead to precipitation of the uranyl salts with buffer components. Electron micrographs were recorded at a magnification of 50,000x and an underfocus of ~400 nm on Eastman Kodak Co. S0-163 sheet films with a Hitachi H-7000 electron microscope operated at 100 kV.

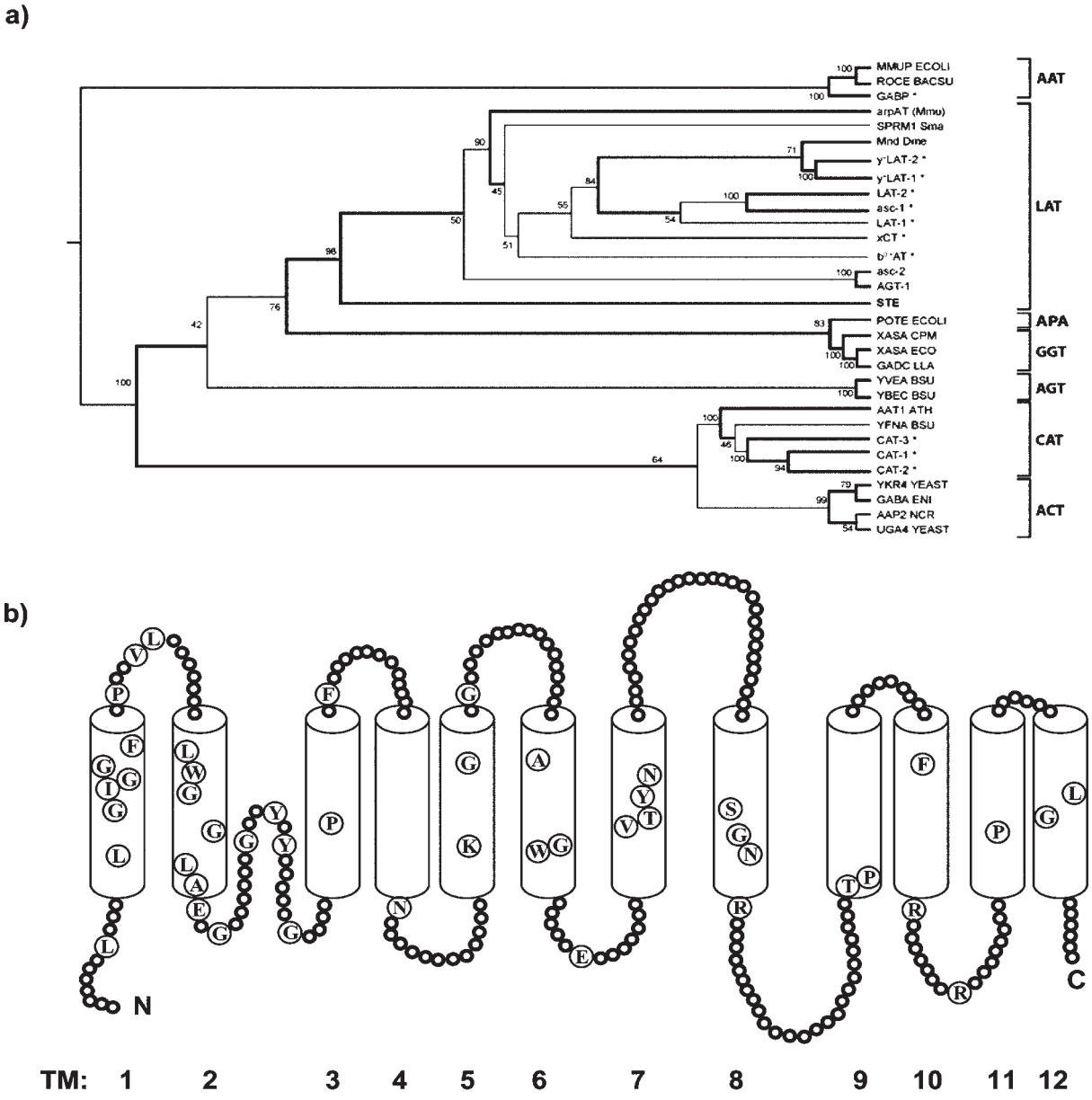

Figure 1. a) Phylogenetic relationship of SteT with other prokaryotic and eukaryotic members of the APC transporter superfamily. The neighbor-joining tree illustrates the phylogenetic relationships of SteT (YkbA from Bacillus subtilis) with all alignable members of the APC superfamily that are either functionally characterized or putative amino acid transporters. The abbreviation of each of the subfamilies is in bold; see Ref. (23) and http://www.tcdb.org/tcdb/superfamily.php for a description of the subfamilies. The tree was rooted using yeast VAL1 permease (YAT family). Bootstrap values (as a percentage of 1000 replicates) are indicated for each of the nodes. Thicker strokes denote branches supported with bootstrap values larger than 70%. In order to simplify the tree, clades with terminal branches that include clear orthologs were collapsed and labeled with the general name of the transporter followed by *. We have identified other bacterial protein sequences that were annotated during the course of our studies: the sequences were closely related to SteT (probably orthologs) and clustered with LAT members as well (data not shown; GenInfo Identifiers, gi: 37521919, 77543889, 67931606, 29898613 and 41582385). b) Membrane topology model of SteT. This model is based on the alignment of the amino acid sequence of SteT with the light subunits of mammalian HATs that interact with the heavy subunits rBAT or 4F2hc (see Suppl. Figure 1), the experimentally probed membrane topology of human xCT (17) and the predicted membrane topology of SteT (Suppl. Figure 1). The membrane topology model of SteT contains 12 transmembrane domains (TM 1 to 12) with intracellular N- and C-termini. Due to the high amino acid sequence similarity to human xCT (see Suppl. Figure 1) the intracellular loop connecting TM2 and TM3 is depicted as a re-entrant loop as demonstrated for human xCT (17). Conserved residues (circled) in SteT and the mammalian light subunits that interact functionally with rBAT or 4F2hc are distributed along the protein but are more frequent in the first part of the protein (TM1, TM2 and the re-entrant loop) (see also Suppl. Figure 1). The conserved Cys residue in the external loop between TM3 and TM4 involved in the disulfide bridge connecting the light and the heavy subunits of eukaryotic HATs is not present in SteT (see also Suppl. Figure 1).

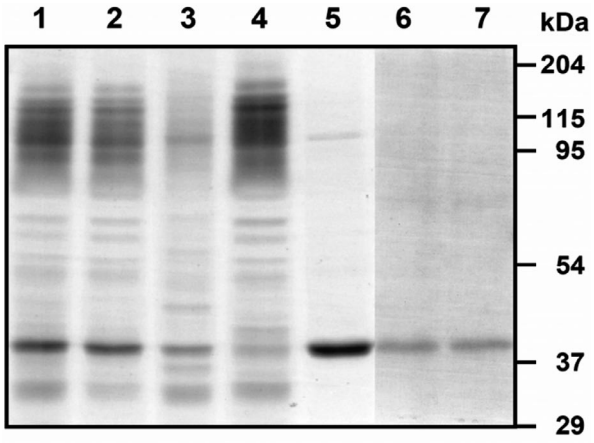

Figure 2. SteT purification and reconstitution into proteoliposomes. Purification of His-tagged SteT by nickel affinity chromatography. SDS-PAGE of the different purification steps: lane 1, *E. coli* total membranes containing overexpressed SteT; lane 2, supernatant after solubilization with 0.5% DDM and ultracentrifugation; lane 3, pellet after ultracentrifugation; lane 4, Ni^{2+}-NTA flowthrough; lane 5, Ni^{2+}-NTA elution. SDS-PAGE of SteT reconstituted into proteoliposomes (1:100 protein-to-lipid ratio) with (lane 6) or without 4 mM L-serine (lane 7). The displayed gels were stained with Coomassie Brilliant Blue. Applied protein per lane: 8 µg (lanes 1-4), 4 µg (lane 5), 2 µg (lanes 6-7). SteT migrates as a prominent band at ~40 kDa (see lanes 5-7).

Blue native gel electrophoresis – Linear 5-12% gradient gels for blue native-polyacrylamide gel electrophoresis (BN-PAGE) were prepared and run as previously described (31). Thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), lactate dehydrogenase (140 kDa), and bovine serum albumin (66 kDa) were used as standard proteins.

4.5 RESULTS

Screening for an appropriate prokaryotic homolog candidate of LAT transporters – In order to identify appropriate prokaryotic homolog candidates for studying the structure-function relationships of the light subunits of HATs, an exhaustive search was made on several available protein sequence databases. Among the most significant matches, four candidates were selected for further investigation: PotE and YhfM from *E. coli*, and YkbA (SteT) and YfnA from *B. subtilis*. These proteins had amino acid sequence identities to the light subunits of HATs ranging from 17% to 29%. All these candidates were hypothetical proteins (32,33) with the exception of PotE, the putrescine/ornithine exchanger (34), from *E. coli*. In preliminary trials, each of these APC members was C-terminally tagged with His_6 and successfully expressed in *E. coli* as described in Materials and Methods. However, detailed investigation was subsequently confined to PotE and SteT because of the high yield of purified protein: up to 3 and 1 mg/L of bacterial culture, respectively. In contrast, purified protein yield was lower than 0.5 mg/L for the other APC members (data not shown).

In a wide phylogenetic study of the APC superfamily of transporters, the PotE and SteT proteins were initially classified as basic amino acid/polyamine antiporters (APA) within one of the eleven independently defined APC subfamilies (1). To our surprise, and in disagreement with this existing classification, our BLAST comparisons of the SteT protein sequence with those of all other APC members clearly showed a preferred match with members of the LAT subfamily (with identities ranging between 31 and 29%) rather than with any other APC transporters, including PotE (with 22% identity). To solve this discrepancy we performed a new phylogenetic analysis of SteT and PotE using a restricted collection of closely related APC subfamilies that, in contrast to wide APC studies, allowed, not only the usage of a more accurate protein alignment algorithms (ProbCons), but also a manual examination and verification of the alignment results, which is therefore expected to provide a more consistent phylogenetic classification. The sequence comparisons and the constructed phylogenetic tree indicated that SteT clearly clusters with all the members of the LAT family (Figure 1a). Furthermore, these results are also in agreement with a revised analysis recently performed by Saier and co-workers (Milton H. Saier Jr., personal communication). Thus, SteT and the other related bacterial sequences described in the legend to Figure 1a can be considered as the first identified prokaryotic members of the LAT family. SteT is predicted to contain 438 amino acid residues (estimated molecular mass 48,879 Da)

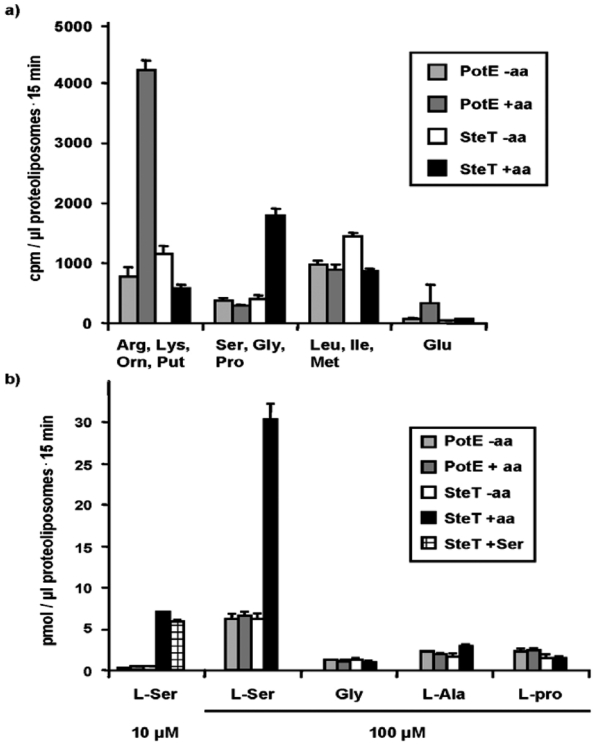

Figure 3. Screening of amino acid substrates for SteT transport activity. a) Transport of the indicated mixture of 10 µM radioactive labeled amino acid (pooled by structural similarity) was tested in proteoliposomes containing SteT or PotE. Proteoliposomes were loaded (PotE +aa and SteT +aa) or not (PotE –aa and SteT –aa) with a mixture of ten representative L-amino acids, i.e., Arg, Orn, Gly, Pro, Ala, Leu, Met, Phe, Tyr, Glu; each at 2 mM. b) Transport of the indicated radiolabeled amino acid (at 10 or 100 µM) into proteoliposomes containing no amino acids (PotE –aa and SteT –aa), a mixture of ten representative amino acids as in a) (PotE +aa and SteT +aa) or 2 mM Ser (SteT +Ser). Transport (a, b) was measured for 15 min and is expressed as cpm / µl of proteoliposomes per 15 min in a) and as pmol / µl of proteoliposomes per 15 min in b). A representative experiment (a, b) with 3 replicates per condition is shown.

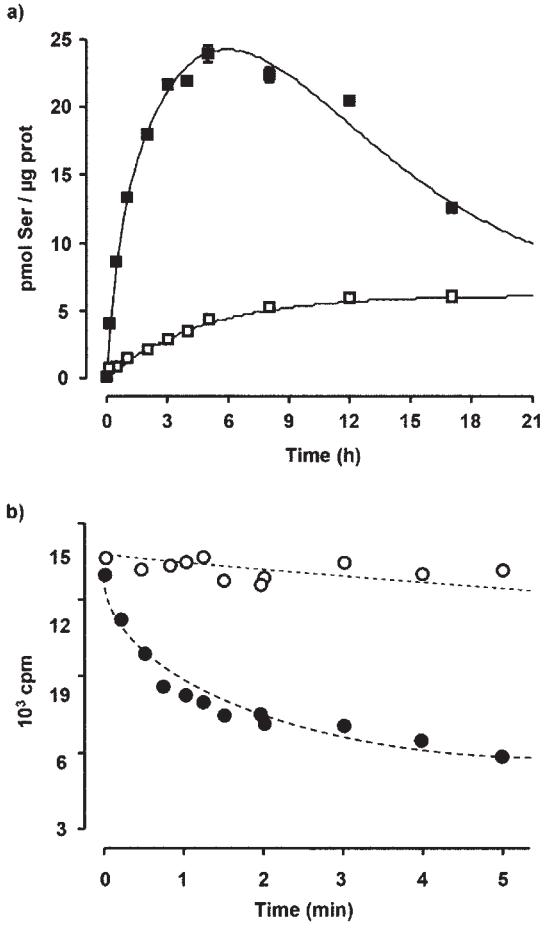

Figure 4. Time-course of L-serine transport in SteT proteoliposomes. a) Influx (pmol / µg protein) of 10 µM radiolabeled L-serine into SteT proteoliposomes (SteT-PLs) containing 4 mM L-serine (closed squares) or L-arginine (open squares). L-arginine is not a substrate for SteT (see Figure 6a, b) and was used as a negative control. Data correspond to a representative experiment, performed using 3 replicates. Errors bars (SEM) when not visible are smaller than symbols. Three additional experiments gave similar results. The continuous lines correspond to simulated progress curves assuming a 1:1 substrate exchange stoichiometry plus simple diffusion for transport in SteT-PLs containing 4 mM L-serine and simple diffusion for transport in SteT-PLs containing 4 mM L-arginine (see experimental procedures for details). b) Efflux of L-serine through SteT transporter: SteT-PLs containing 4 mM L-serine were loaded with 10 µM radiolabeled L-serine for 2 h. Then, SteT-PLs were diluted 13-fold with medium containing (closed circles) or lacking (open circles) unlabelled 7 mM L-serine. Efflux of L-serine was dramatically stimulated by L-serine in the external medium. A representative experiment is shown. A second experiment gave similar results.

and its sequence can be aligned in its entirety with almost the full lengths of the sequences of all the eukaryotic LAT family members (Suppl. Figure 1). The estimated membrane topology of SteT (Suppl. Figure 1) is similar to that of these transporters and fits the experimental topology model of xCT, which contains 12 transmembrane (TM) domains (17). Totally conserved residues in SteT and the mammalian rBAT- or 4F2hc-associated LAT family members (see selected transporters in Suppl. Figure 1) are indicated in Figure 1b. The regions most highly conserved between SteT and the eukaryotic LAT family members correspond to TM1, TM2 and the re-entrant loop between TM2 and TM3. The main differences are shorter N- and C- termini in SteT and the absence of the cysteine residue, in the loop between TM3 and TM4, that is involved in the disulfide bridge between the light and the heavy subunits in the eukaryotic LAT family members.

Functional characterization of SteT – In order to characterize the function of SteT, C-terminally His_6-tagged SteT was overexpressed in *E. coli* and purified by nickel affinity chromatography (Figure 2, lanes 1 to 5). Purified SteT migrates as a band of ~40 kDa, similar in its mobility to monomers of the mammalian LAT family members, which have a molecular mass of ~50 kDa (17,35,36). The faster migration of these proteins in SDS-PAGE than would be expected from their predicted molecular masses probably reflects their hydrophobic nature and is commonly observed for membrane transport proteins. A higher molecular weight band (~115 kDa; Figure 2, lane 5), visible in some but not all protein preparations, corresponded to a low-abundance contaminant since it was not detected with HisProbe-HRP (Qiagen) (i.e., His tag detection) (data not shown). From scanning densitometry of multiple Coomassie blue-stained gels, the purity of a typical SteT preparation was estimated to be >86%. Purified SteT was reconstituted in the presence of *E. coli* lipids to form proteoliposomes (SteT-PLs) at a protein-to-lipid ratio of 1:100. Again the ~40 kDa band is visible in SteT-PLs (Figure 2, lanes 6 and 7).

SteT was a hypothetical protein identified in the *B. subtilis* genome sequence (32). Based on its phylogenetic relationship with eukaryotic LAT family members we reasoned that SteT could be an amino acid exchanger. To identify the SteT transport activity a fast screening approach for amino acid exchanger activity was set up. The first step consisted in preparing SteT-PLs with or without a mixture of ten representative amino acids (L-Arg, L-Orn, Gly, L-Pro, L-Ala, L-Leu, L-Met, L-Phe, L-Tyr, L-Glu, at 1 mM each) contained inside the proteoliposomes. Then, influx of a group of radiolabeled, structurally-related amino acids was measured. In agreement with the proposed transport function determined in whole bacteria expressing PotE and in membrane vesicles derived from these cells (19,34), PotE reconstituted in liposomes (PotE-PLs) showed putrescine/ornithine and ornithine/ornithine exchange (Suppl. Figure 2). As expected, PotE-PLs showed higher transport in the amino acid-filled than in the empty proteoliposomes (i.e., trans-stimulation) with the group of radiolabeled L-arginine, L-lysine, L-ornithine and putrescine (Figure 3a). In contrast, SteT-PLs showed trans-stimulated uptake only with the group of

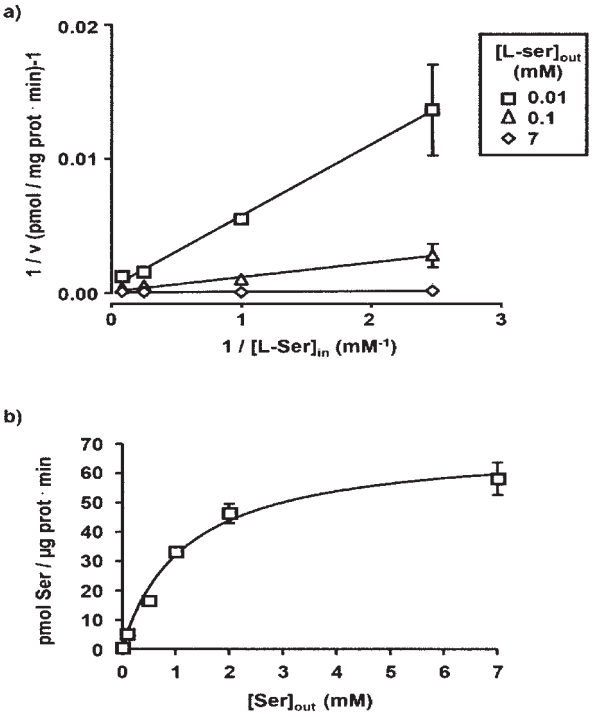

Figure 5. Kinetic analysis of SteT transport. a) Transport of radiolabeled L-serine at different concentrations in the external (as indicated) and the internal (0.4, 1, 4 and 12 mM) medium was measured in SteT proteoliposomes (SteT-PLs) for 1 min under linear conditions. Trans-stimulated transport was calculated by subtracting transport in SteT-PLs containing no L-serine in the internal medium assayed at the three indicated external concentrations of L-serine. Lineweaver-Burk linearization of the trans-stimulated transport with the different external and internal L-serine concentrations is shown. The lines converge to negative X-axis values supporting a sequential model of exchange. The reverse plot (external L-serine concentrations in the X-axis for each internal concentration of L-serine) gave a similar plot. The estimated apparent K_m values (calculated from the intersection in the X-axis for each line) ranged between 1.2 to 5.3 and 1.5 to 3.1 mM for internal and external L-serine, respectively. The estimated V_{max} with 12 mM internal L-serine was 67 ± 3 pmol / µg protein · min (mean ± SEM). A representative experiment is shown with 3 replicates per condition. Errors bars (SEM) when not visible are smaller than symbols. b) Transport of radiolabeled L-serine at different external concentrations (0.01, 0.1, 0.5, 1, 2, and 7 mM) was measured in SteT-PLs containing 15 mM L-serine or L-arginine in linear conditions. Trans-stimulated transport was calculated by subtracting transport in SteT-PLs containing L-arginine from that in those containing L-serine. Non-parametric analysis (GraphPad; Sigma) estimated an external apparent K_m of 1.2 ± 0.2 mM. The estimated V_{max} was 67 ± 4 pmol / µg protein · min (mean ± SEM). A representative experiment is shown with 3 replicates per condition. Errors bars (SEM) when not visible are smaller than symbols.

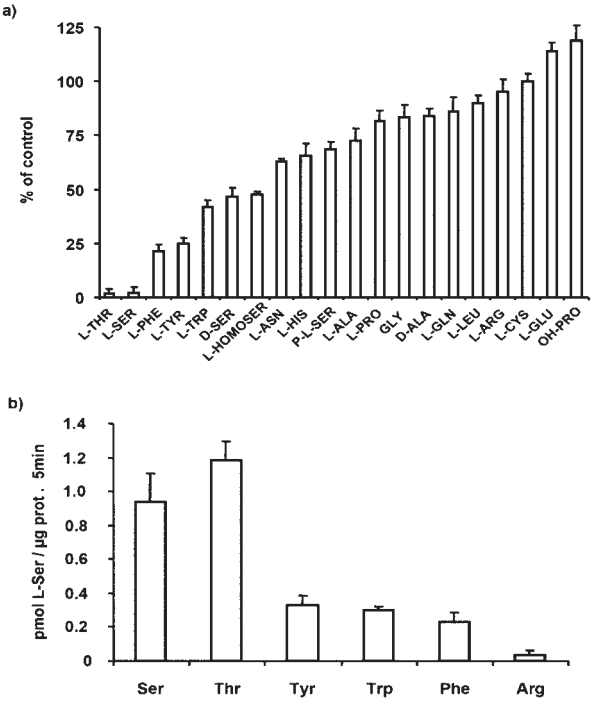

Figure 6. Substrate specificity of SteT transport activity. a) Amino acid cis-inhibition pattern of SteT transport activity. Transport of 10 µM radiolabeled L-serine was measured in SteT proteoliposomes (SteT-PLs) containing no amino acids or 4 mM L-serine for 5 min (linear conditions). Transport was measured in the absence (no amino acids) or presence of the indicated amino acids at 5 mM in the external medium (2 mM in the case of L-tyrosine). Trans-stimulated transport was calculated by subtracting transport in SteT-PLs containing no L-serine from that in SteT-PLs containing L-serine. Transport is expressed as the percentage of the transport in SteT proteoliposomes containing 4 mM L-serine and with no cis-inhibitors (no amino acids) (5.2 ± 1.2 pmol / µg protein · 15 min). Data are from 2-3 experiments with 3 replicates per condition. b) Trans-stimulation of SteT transport activity. Transport of 10 µM radiolabeled L-serine was measured in SteT proteoliposomes containing 4 mM of the indicated amino acids (2 mM in the case of L-tyrosine) for 5 min. L-arginine was used as a negative control because this amino acid did not inhibit L-serine transport via SteT (see Figure 6a). Transport is expressed in pmol / µg protein in 5 min. A representative experiment with 3 replicates per condition is shown. Two additional experiments gave similar results.

L-serine, glycine and L-proline (Figure 3a). Next, trans-stimulation of transport of these amino acids, including L-alanine, was measured individually with SteT-PLs (Figure 3b). Only L-serine transport was trans-stimulated in amino acid-filled SteT-PLs. Similarly, substitution of the intra-proteoliposome mixture of amino acids by L-serine as the unique internal substrate also showed trans-stimulation of L-serine transport in SteT-PLs (Figure 3b). Indeed, the initial velocity of transport of 10 µM L-serine was 11-fold higher in L-serine-filled than in L-arginine-filled SteT-PLs (2.12 ± 0.15 and 0.19 ± 0.05 pmol / µg protein per 10 min respectively; n=5 independent experiments). As expected, transport of the amino acids from this group was not trans-stimulated in PotE-PLs (Figure 3b).

Trans-stimulation can result either from allosteric trans-activation or exchange activity. In order to distinguish between these two possible mechanisms, L-serine transport was examined over time in SteT-PLs containing L-serine or L-arginine (used as a negative control) (Figure 4a). SteT showed accumulation of L-serine over the equilibrium levels, as indicated by the overshoot observed in L-serine-containing SteT-PLs, which was not observed in L-arginine-containing SteT-PLs. SteT capacity for active accumulation is characteristic of an exchanger and is not compatible with an allosteric trans-activation mechanism. Moreover, the transport activity of SteT was bidirectional, as shown by L-serine efflux measurements from SteT-PLs (Figure 4b). Efflux of radiolabeled L-serine from preloaded SteT-PLs was dramatically trans-stimulated by L-serine in the external medium. In contrast, L-serine was apparently transported by simple diffusion in SteT-PLs containing L-arginine: i) L-serine transport in PotE-PLs was indistinguishable from that in SteT-PLs when only L-arginine was present inside the proteoliposomes (data not shown). ii) Moreover, pretreatment with the cysteine-reactive reagent MTSET (1 mM for 5 min) inactivated L-serine transport in L-serine-containing (65% inhibition) but not in L-arginine-containing SteT-PLs (data not shown). SteT contains five cysteine residues, which have been mutated to serine to construct a cysteine-less version of the transporter (SteT-Cysless). This mutated protein retains ~25% of the transport activity of the wild-type SteT (data not shown). Pretreatment with MTSET (1 mM for 5 min) did not affect L-serine transport in SteT-Cysless-PLs containing L-serine (data not shown). These findings indicate that L-serine transport in SteT-PLs containing L-serine is SteT-mediated, whereas that in SteT-PLs containing

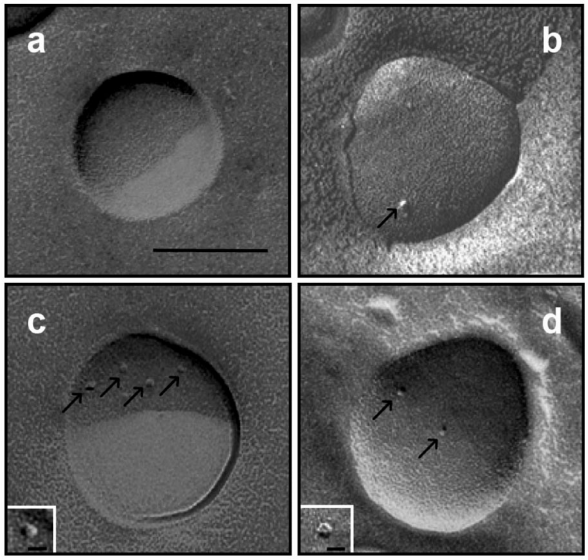

Figure 7. Freeze-fracture TEM of purified SteT and PotE reconstituted in liposomes. a) control liposomes are devoid of intramembrane particles. b–c) SteT proteoliposomes prepared for transport assays at a protein-to-lipid ratio of 1:100 (b) and 1:40 (c). Intramembrane particles (arrows) appear in both concave (b) and convex faces (c) of unilamellar vesicles, indicating a random insertion of the protein in the bilayer. d) PotE proteoliposomes prepared for transport assays at a protein-to-lipid ratio of 1:100. Intramembrane particles (arrows) are visible. Scale bar: 100 nm. Insets: high magnification view of one SteT (c) and one PotE (d) particle. Scale bar in insets: 10 nm.

L-arginine is not SteT-mediated. Taken together, all these results demonstrate that SteT is an obligate exchanger.

The kinetics of L-serine/L-serine exchange in SteT-PLs were examined by varying L-serine concentrations inside and outside proteoliposomes (Figure 5a). Lineweaver-Burk plots showed straight lines with different slopes where internal and external apparent K_m values were independent of the substrate concentration in trans. This kinetic behavior supports a sequential mechanism and rules out a ping-pong mechanism, which is characterized by a non-intersecting pattern (37). This initial kinetic analysis estimated the apparent K_m values from inside and outside the proteoliposomes to lie in the range of 1-5 and 1-3 mM, respectively. Subsequently, a second kinetic analysis was performed with 15 mM L-serine inside SteT-PLs and varying the external L-serine concentration (Figure 5b). In agreement with these previous results, the external apparent K_m was 1.2 ± 0.2 mM. In these two experiments (with up to 15 mM L-serine inside SteT-PLs) the V_{max} of L-serine/L-serine exchange was 67 pmol/ µg protein · min. Thus, the turnover rate of SteT L-serine/L-serine exchange was 0.06 s^{-1}.

In order to test whether ions were cotransported via SteT, we studied the effect of membrane potential on L-serine/L-serine exchange (Suppl. Figure 3). A membrane potential (negative inside) was generated by inclusion of 120 mM KPi in the internal medium of the proteoliposomes followed by the addition of valinomycin to the external medium. PotE-PLs were used as a positive control. PotE exchanges putrescine and L-ornithine with 1:1 substrate stoichiometry (34). At neutral pH putrescine/L-ornithine exchange would be electrogenic (two charges (putrescine) will be exchanged for one charge (L-ornithine)). Consistent with this stoichiometry, imposing a membrane potential increased the putrescine (influx)/L-ornithine (efflux) exchange in PotE-PLs by ~80%. In contrast, L-serine/L-serine exchange in SteT-PLs was independent of the membrane potential. Moreover, SteT exchange activity was independent of Na^+ and Cl^- in the external medium (i.e,. substitution by choline and acetate respectively had no effect), and was not affected by variation of the external medium pH (from 5.5 to 7.5; internal proteoliposome pH = 7.4) (data not shown). Similarly, SteT exchange activity was independent of K^+ gradient (i.e., equilibration of the K^+ gradient by the addition of 120 mM K^+ to the external medium; data not shown).

To further confirm that SteT mediates obligate exchange, the time-course of L-serine/L-serine exchange via SteT was simulated (see Experimental Procedures). Transport of 10 µM radiolabeled L-serine into SteT-PLs containing 4 mM L-arginine fitted well to a model of simple diffusion with a diffusion coefficient of 2.5 x 10^{-6} min^{-1} and an internal SteT-PLs volume estimated at equilibrium of 62 nl / µg of purified protein (Figure 4a). Transport of 10 µM radiolabeled L-serine into SteT-PLs containing 4 mM L-serine

was simulated using two transport components: a) electroneutral Ser/Ser exchange with 1:1 exchange stoichiometry and apparent internal and external substrate affinities of 2.0 and 1.2 mM respectively plus b) the simple diffusion component (Figure 4a). The internal SteT-PLs volume estimated at equilibrium was 62 nl / µg of protein. Interestingly, the best agreement with the experimental data was obtained when a decrease in transport activity over time was included in the simulation (V_{max} from 70 to 22 pmols /µg protein · min within the first 2 h with a $t_{1/2}$ of 21 min). The nature of this inactivation during time-course transport measurements is unknown. However, these simulations support an electroneutral and obligate exchange of L-serine by SteT, and indicate that this exchange is balanced (probably with 1:1 stoichiometry).

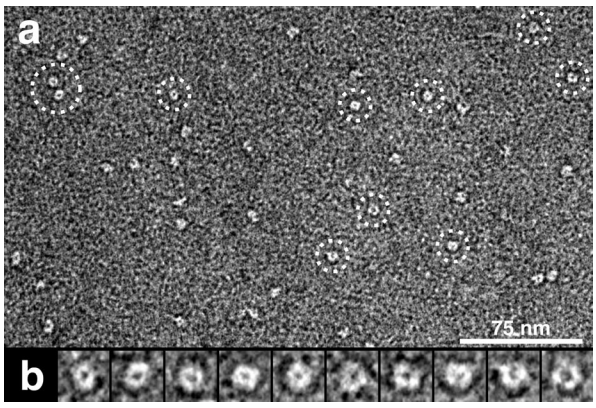

Figure 8. TEM of negatively stained SteT. The homogeneity of the DM-solubilized SteT proteins is reflected in the electron micrograph (a). The selected top view particles marked by broken circles were magnified and are displayed in the gallery (b). SteT proteins are elliptical (diameters: ~6 x ~7 nm) and donut-shaped with a central depression. The frame size of the magnified particles in the gallery is 12 nm.

Finally, the amino acid specificity of SteT transport activity was characterized by cis-inhibition experiments (Figure 6a). A 500-fold excess of the L-stereoisomers of serine and threonine almost completely abolished L-serine/L-serine exchange. Amino acids structurally related to L-serine (D-serine and L-homoserine) also inhibited transport, but to a much lower extent. Similarly, the phosphorylated form of L-serine (L-phosphoserine) was a much poorer inhibitor. Many other amino acids, including glycine, proline, arginine, cysteine, lysine and glutamate, inhibited transport very weakly or not at all. Surprisingly, L-aromatic amino acids (phenylalanine, tyrosine and tryptophan) significantly inhibited SteT transport. In order to confirm the role of threonine and the aromatic amino acids as substrates, SteT-PLs were filled with these amino acids and trans-stimulation of L-serine transport was measured (Figure 6b). Transport was clearly trans-stimulated by L-serine and L-threonine, and to a lesser extent by the L-aromatic amino acids. In contrast, L-arginine, used as a negative control, did not trans-stimulate L-serine transport.

The results presented demonstrate that SteT exhibits an obligate exchange activity for serine, threonine and aromatic amino acids. Because serine and threonine were the more efficiently transported, we propose that YkbA be re-designated as a Serine/Threonine Exchanger Transporter (SteT).

Structure and oligomeric state of SteT and PotE – To ascertain whether SteT and PotE exist in a monomeric or oligomeric form, three independent methods were used: Blue native-polyacrylamide gel electrophoresis (BN-PAGE), freeze-fracture TEM, and negative stain TEM. BN-PAGE results of SteT are summarized in Suppl. Figure 4. DM-solubilized and purified His-tagged SteT migrated as a strong protein band at an apparent molecular weight (M_{obs}) of ~80 kDa. In addition, two faint bands at ~155 kDa and ~205 kDa as well as a stronger one at ~40 kDa were detected. The latter band had a bright blue colour and corresponds to detergent/Coomassie brilliant blue G-250 micelles as demonstrated previously (38). In Western blot analysis the ~80 kDa, ~155 kDa and ~205 kDa bands reacted with HisProbe-HRP and therefore corresponded to SteT (data not shown). BN-PAGE with purified PotE was not successful because of its strong tendency to precipitate during gel electrophoresis.

In a second approach, freeze-fracture TEM

analysis was applied to estimate the size of functional His-tagged SteT and PotE reconstituted in proteoliposomes (SteT-PLs and PotE-PLs). Aliquots (5 µL) of proteoliposomes, as used for the transport assays, at a protein-to-lipid ratio of 1:100 and for these studies also at 1:40, were prepared for freeze-fracture TEM. SteT- and PotE-PLs at protein-to-lipid ratios of 1:100 and 1:40 showed identical transport activity when corrected by µg of protein (data not shown). In the replicas after freeze-fracture, proteoliposomes appeared as concave or convex surfaces with diameter of 150-200 nm (Figure 7). The majority (90%) of the proteoliposomes fracture faces contained 0-2 and 0-6 intramembrane particles in SteT-PLs (1:100 and 1:40 protein-to-lipid ratios) (Figure 7b-c). A similar number of particles (0-2 and 0-7) were present in PotE-PLs fracture faces (Figure 7d). These particles were absent from control liposomes (Figure 7a). This finding demonstrates that the observed particles represented reconstituted SteT or PotE proteins. At high magnification SteT and PotE particles had a round shape (Figure 7c, d; insets). Measurements on 1:40 protein-to-lipid ratio proteoliposomes indicated single populations of particles with diameters of 7.4 ± 0.1 nm (n= 119) and 8.0 ± 0.1 nm (n= 109) for SteT and PotE, respectively. As a control for our measurements we used lactose permease from *E. coli* and we found a similar particle size. Thus, freeze-fracture TEM analysis applied to the C154G lactose permease mutant (i.e., the same mutant protein used for the crystal structure resolution; ref. 39) reconstituted in liposomes (1:100 protein-to-lipid ratio) showed a particle diameter of 6.0 ± 0.1 nm (n= 72) (data not shown).

Finally, DM-solubilized His-tagged SteT protein was negatively stained and examined by TEM. The homogeneity of purified SteT is documented in Figure 8a. Single SteT proteins were distinguished and displayed an elliptic "donut-like" shape with diameters of ~6 nm by ~7 nm and a central stain-filled indentation (see gallery of well preserved SteT top views in Figure 8b). As already indicated from BN-PAGE, PotE had a tendency to aggregate when solubilized in detergent. This behaviour was further supported by negative stain TEM which showed a heterogeneous population of particles (data not shown).

4.6 DISCUSSION

In this work we have provided, for the first time, information on the functional and structural features of SteT from *B. subtilis*. The first part of our extensive analysis shows that SteT is an excellent prokaryotic model for structural and functional studies of eukaryotic amino acid exchangers, in particular members of the LAT family: (i) Phylogenetic analysis clustered the SteT amino acid sequence together with LAT family members within the APC superfamily (Figure 1a); (ii) The putative membrane topology of SteT (Figure 1b and Suppl. Figure 1) is compatible with that experimentally determined for eukaryotic LAT transporters, i.e. the catalytic subunit of HATs (17); (iii) SteT showed obligate amino acid exchange activity (Figs. 3b, 4 and 6b), which is characteristic of the catalytic subunits of HATs (14); (iv) Furthermore, kinetic analysis supports a sequential mechanism of exchange for SteT (Figure 5), which has also been described for the HAT system $b^{0,+}$ in chicken small intestine (15). Thus, SteT is the first identified and characterized prokaryotic member of the LAT family of amino acid transporters.

In agreement with SteT amino acid exchange activity, and in support of our functional characterization protocol, the SteT ortholog in *B. cereus* (GenBank accession AAP11885, gi 29898613, annotated as hypothetical protein in Ref. 40) showed the same L-serine/ L-threonine exchange activity when purified and reconstituted in proteoliposomes (data not shown). Other than the clear amino acid exchange activity detected here for both orthologues, to our knowledge a SteT-related transport activity still remains to be described in cells or membrane vesicles of *Bacillus sp*.

Analysis of SteT in PLs by freeze-fracture TEM indicated that the protein is integrated in the lipid bilayer of unilamellar vesicles. Reconstituted SteT resides in two orientations in the bilayer, as suggested by freeze fracture TEM images of SteT-PLs showing particle retention on concave as well as convex faces (Figure 7b, c). Distinct substrate affinities may be expected for the two orientations (up to 3 orders of higher apparent affinity for the external than the cytoplasmic side), as has been demonstrated for the reconstituted catalytic subunit ($b^{0,+}AT$) of system $b^{0,+}$ (16) and for other HATs expressed in oocytes (41). In contrast, curve fitting of kinetic data to Michaelis-Menten equations for two distinct affinities for L-serine outside SteT-PLs contraindicates a second K_m for L-serine and gives a unique external apparent K_m of 1.2 mM for L-serine (Figure 5). This suggests that SteT has similar apparent affinities at both sides of the membrane.

The blue native-gel of purified SteT indicates a strong and major band with an M_{obs} of ~80 kDa. This finding suggests that the majority of SteT molecules exist in monomeric form, given that the polypeptide has a molecular mass of ~49 kDa (based on its amino acid sequence) and migrates in BN-PAGE in association with a detergent/ Coomassie brilliant blue G-250 micelle. The use of the conversion factor (1.8) determined by Heuberger *et al*. (42) to estimate the mass of membrane proteins from the M_{obs} also supports the existence of SteT in monomeric form. In addition, experiments performed with a wide range of amino- (DSP, DST and DMS) and sulfhydryl-specific ($BM(PEO)_4$) reagents did not yield cross-linking products (data not shown) in line with the results from BN-PAGE.

Freeze-fracture TEM of functional SteT and PotE in PLs showed homogeneous round-shaped particles with diameters of ~7.4 and ~8.0 nm, respectively. These values are only slightly higher than the diameter estimated for the lactose permease monomer (~6 nm, see results) and in agreement with previous studies (43). Lactose permease has 12 transmembrane segments and similar molecular mass to SteT and PotE, and its monomeric state is well established (39). This suggests that functional SteT and PotE are monomers. The thickness of the platinum-carbon film deposited to produce our replicas may have led to overestimation of the true particle diameter by ~2 nm, as described for freeze-fracture TEM analysis of other transporters inserted in proteoliposomes (44). Thus, the diameter of the SteT and PotE particles can be assumed to be in the range of 5 to 6 nm. TEM of negatively stained SteT proteins revealed an elliptical donut-like structure with diameters of ~7 nm by ~6 nm. Similar to freeze-fracture TEM, the measured dimensions are overestimated. With detergent-solubilized membrane proteins the additional mass arising from the detergent belt and endogenous lipid bound to the protein has to be considered. According to Dekker *et al.* (45) the true protein volume is about 20% smaller than the volume estimated from the negatively stained detergent-solubilized membrane protein. Thus, the corrected dimensions of SteT from negative stain TEM are in agreement with those from freeze fracture TEM. The “red permease” is a lactose permease fusion protein which consists of 12 transmembrane helices, has a comparable molecular mass to SteT, and forms monomers and trimers (46). The reported dimensions for the “red permease” monomer reconstituted into lipid membranes, as determined by negative stain TEM are about 5 x 4 nm (46) and therefore similar to the dimensions of SteT after correction for the detergent bound to the protein (see above and Ref. (45)). Besides the size, both proteins possess a similar overall structure and a pronounced central indentation. As control experiment and additional support for the monomeric nature of negatively stained detergent-solubilized SteT (Figure 8), its dimensions were compared with those of negatively stained detergent-solubilized glutamate transporter (GltP). GltP forms trimers and the molecular mass of the monomer is similar to that of SteT. The measured dimensions of the GltP monomer within the trimer are comparable to those of SteT reinforcing the conclusion that SteT is monomeric (data not shown).In summary,

our data suggest SteT and PotE to be functional in the monomeric form.

Little is known about the structure of the light subunits of HATs (LAT family). The elucidation of their structure is mainly limited by the low HAT expression in native cells and tissues, and the difficulty to overexpress these proteins in heterologues systems. An excellent alternative to understand the molecular mechanism of HATs represents the elucidation of the atomic structure of a prokaryotic transporter with high sequence identity to eukaryotic light subunits of HATs. To this end and for the first time we have identified, cloned, overexpressed and purified the SteT protein, a prokaryotic member of the LAT family. Functional and structural characterization of SteT has provided insights into the transport mechanism, kinetics, structure and oligomeric state of this important family of amino acid transporters. However, a high resolution structure is needed to fully understand the molecular mechanism of SteT and to relate the previous functional data. The first step towards the growth of 2D and 3D protein crystals suitable for high-resolution structure analysis has been achieved here by the successful overexpression and purification of high amounts of functional SteT.

4.7 ACKNOWLEDGEMENTS

We thank Ron H. Kaback for the gift of C154G lactose permease mutant protein and Milton H. Saier Jr. for relevant discussions. We thank Carmen López for excellent technical assistance for freeze-fracture experiments. This study was supported in part by the Spanish Ministry of Science and Technology (SAF2003-08940 and BFU2006-14600 to M.P.), by the EC project grant 502802 (EUGINDAT, to D.F. and M.P.) and by the Instituto de Salud Carlos III (networks C3/08P and G03/054 to M.P.). C.D.R. and M.R. are recipients of a predoctoral fellowship from the Spanish Ministerio de Educación y Ciencia. D.F. acknowledges support from the Swiss National Research Foundation, the M.E. Müller Foundation and the Swiss National Center of Competence in Research (NCCR) 'Structural Biology'. S.A.B. acknowledges the support of the Biotechnology and Biological Sciences Research Council (BBSRC) UK.

4.8 REFERENCES

1. Jack DL., Paulsen IT. and Saier Jr. MH. (2000) Microbiology **146** 1797-1814
2. Palacin M., and Kanai Y. (2004) *Pflugers Arch* **447** 490-494
3. Verrey F., Closs EI., Wagner CA., Palacin M., Endou H., and Kanai Y. (2004) *Pflugers Arch* **447** 532-542
4. Palacin M., Nunes V., Font-Llitjos M., Jimenez-Vidal M., Fort J., Gasol E., Pineda M., Feliubadalo L., Chillaron J. and Zorzano A. (2005) *Physiology (Bethesda)* **20** 112-124
5. Calonge MJ., Gasparini P., Chillaron J., Chillon M., Gallucci M., Rousaud F., Zelante L., Testar X., Dallapiccola B., Di Silverio F., Barcelo P., Estivill X., Zorzano A., Nunes V., and Palacin M. (1994) *Nat Genet* **6** 420-425
6. Feliubadalo L., Font M., Purroy J., Rousaud F., Estivill X., Nunes V., Golomb E., Centola M., Aksentijevich I., Kreiss Y., Goldman B., Pras M., Kastner DL., Pras E., Gasparini P., Bisceglia L., Beccia E., Gallucci M. and *et al.* (1999) *Nat Genet* **23** 52-57
7. Torrents D., Mykkanen J., Pineda M., Feliubadalo L., Estevez R., de Cid R., Sanjurjo P., Zorzano A., Nunes V., Huoponen K., Reinikainen A., Simell O., Savontaus M L., Aula P. and Palacin M. (1999) *Nat Genet* **21** 293-296
8. Borsani G., Bassi MT., Sperandeo MP., De Grandi A., Buoninconti A., Riboni M., Manzoni M., Incerti B., Pepe A., Andria G., Ballabio A. and Sebastio G. (1999) *Nat Genet* **21** 297-301
9. Sato H., Tamba M., Ishii T., and Bannai S. (1999) *J Biol Chem* **274** 11455-11458
10. Bassi MT., Gasol E., Manzoni M., Pineda M., Riboni M., Martin R., Zorzano A., Borsani G. and Palacin M. (2001) *Pflugers Arch* **442** 286-296
11. Kaleeba JA. and Berger EA. (2006) *Science* **311** 1921-1924
12. Baker DA., McFarland K., Lake RW., Shen H., Tang XC., Toda S., and Kalivas PW. (2003) *Nat Neurosci* **6** 743-749
13. Sato H., Shiiya A., Kimata M., Maebara K., Tamba M., Sakakura Y., Makino N., Sugiyama F., Yagami K., Moriguchi T., Takahashi S. and Bannai S. (2005) *J Biol Chem* **280** 37423-37429
14. Chillaron J., Roca R., Valencia A., Zorzano A. and Palacin M. (2001) *Am J Physiol Renal Physiol* **281** F995-1018
15. Torras-Llort M., Torrents D., Soriano-Garcia JF., Gelpi JL., Estevez R., Ferrer R., Palacin M and Moreto M. (2001) *J Membr Biol* **180** 213-220
16. Reig N., Chillaron J., Bartoccioni P., Fernandez E., Bendahan A., Zorzano A., Kanner B., Palacin M. and Bertran J. (2002) *Embo J* **21** 4906-4914
17. Gasol E., Jimenez-Vidal M., Chillaron J., Zorzano A., and Palacin M. (2004) *J Biol Chem* **279** 31228-31236
18. Jimenez-Vidal M., Gasol E., Zorzano A., Nunes V., Palacin M. and Chillaron J. (2004) *J Biol Chem* **279** 11214-11221
19. Kashiwagi K., Kuraishi A., Tomitori H., Igarashi A., Nishimura K., Shirahata A. and Igarashi K. (2000) *J Biol Chem* **275** 36007-36012
20. Hu LA. and King SC. (1999) *Biochem J* **339** 649-655
21. Habermeier A., Wolf S., Martine U., Graf P. and Closs EI. (2003) *J Biol Chem* **278**19492-19499
22. Altschul SF., Gish W., Miller W., Myers E. W. and Lipman DJ. (1990) *J Mol Biol* **215** 403-410
23. Chang AB., Lin R., Keith Studley W., Tran, CV. and Saier MH. Jr. (2004) *Mol Membr Biol* **21** 171-181
24. Do C.B., Mahabhashyam MS., Brudno M. and Batzoglou S. (2005) *Genome Res* **15** 330-340
25. Castresana J. (2000) *Mol Biol Evol* **17** 540-552
26. Chenna R., Sugawara H., Koike T., Lopez R., Gibson TJ., Higgins DG. and Thompson JD. (2003) *Nucleic Acids Res* **31** 3497-3500
27. Xie H., Patching SG., Gallagher MP.,

Litherland GJ., Brough AR., Venter H., Yao SY., Ng AM., Young JD., Herbert RB., Henderson PJ. and Baldwin SA. (2004) *Mol Membr Biol* **21** 323-336
28. Chillaron J., Estevez R., Mora C., Wagner CA., Suessbrich H., Lang F., Gelpi JL., Testar X., Busch AE., Zorzano A. and Palacin M. (1996) *J Biol Chem* **271** 17761-17770
29. Eleno N., Deves R. and Boyd CA. (1994) *J Physiol* **479** 291-300
30. Lopez O., de la Maza A., Coderch L., Lopez-Iglesias C., Wehrli E. and Parra JL. (1998) *FEBS Lett* **426** 314-318
31. Schagger H., Aquila H. and von Jagow G. (1988) *Anal Biochem* **173** 201-205
32. Kunst F., Ogasawara N., Moszer I., Albertini AM., Alloni G., Azevedo V., Bertero MG., Bessieres P., Bolotin A., Borchert S., Borriss R., Boursier L., Brans A., Braun M., Brignell SC., Bron S., Brouillet S., Bruschi CV. and *et al.* (1997) *Nature* **390** 249-256
33. Riley M., Abe T., Arnaud MB., Berlyn MK., Blattner FR., Chaudhuri RR., Glasner JD., Horiuchi T., Keseler IM., Kosuge T., Mori H., Perna NT., Plunkett G., Rudd KE., Serres MH., Thomas GH., Thomson NR., Wishart D. and Wanner BL. (2006) *Nucleic Acids Res* **34** 1-9
34. Kashiwagi K., Miyamoto S., Suzuki F., Kobayashi H. and Igarashi K. (1992) *Proc Natl Acad Sci U S A* **89** 4529-4533
35. Torrents D., Estevez R., Pineda M., Fernandez E., Lloberas J., Shi YB., Zorzano A. and Palacin M. (1998) *J Biol Chem* **273** 32437-32445
36. Fernandez E., Carrascal M., Rousaud F., Abian J., Zorzano A., Palacin M. and Chillaron J. (2002) *Am J Physiol Renal Physiol* **283** F540-548
37. Dierks T., Riemer E. and Kramer R. (1988) *Biochim Biophys Acta* **943** 231-244
38. Suda K., Filipek S., Palczewski K., Engel A. and Fotiadis D. (2004) *Mol Membr Biol* **21** 435-446
39. Abramson J., Smirnova I., Kasho V., Verner G., Kaback HR., Iwata S. (2003) *Science* **301** 610-615.
40. Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B., Kapatral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A., Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T., Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich SD., Overbeek R. and Kyrpides N. (2003) *Nature* **423** 87-91
41. Meier C., Ristic Z., Klauser S. and Verrey F. (2002) *Embo J* **21** 580-589
42. Heuberger EH., Veenhoff LM., Duurkens RH., Friesen RH. and Poolman B. (2002) *J Mol Biol* **317** 591-600
43. Costello MJ., Escaig J., Matsushita K., Viitanen PV., Menick DR., Kaback HR. (1987) *J Biol Chem* **262** 17072-17082
44. Turk E., Kim O., le Coutre J., Whitelegge JP., Eskandari S., Lam JT., Kreman M., Zampighi G., Faull KF. and Wright EM. (2000) *J Biol Chem* **275** 25711-25716
45. Dekker JP., Boekema EJ., Witt HT. and Rögner M. (1988) *Biochim Biophys Acta* **936** 307-318
46. Zhuang J., Prive GG., Werner GE., Ringler, P., Kaback HR. and Engel A. (1999) *J Struct Biol* **125** 63-75
47. Fernandez E., Torrents D., Zorzano A., Palacin M. and Chillaron J. (2005) *J Biol Chem* **280** 19364-19372
48. Krogh A., Larsson B., von Heijne G. and Sonnhammer, E. L. (2001) *J Mol Biol* **305** 567-580

5. Projection Structure of YbgH, a Prokaryotic Member of the Peptide Transporter Family

Fabio Casagrande[1], Daniel Harder[2], Peter Tittmann[3], Andreas Engel[1], Dietmar Weitz[2,4], Heinz Gross[3], Hannelore Daniel[2] and Dimitrios Fotiadis[5]

[1] M.E. Müller Institute for Structural Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland

[2] Department of Food and Nutrition, Molecular Nutrition Unit, Technical University of Munich, Am Forum 5, D-85350 Freising, Germany

[3] Electron Microscopy ETH Zürich (EMEZ), Swiss Federal Institute of Technology, CH-8093, Zürich, Switzerland

[4] Sanofi, Industriepark Hoechst, D-65926 Frankfurt, Germany

[5] Institute of Biochemistry and Molecular Medicine, University of Berne, Bühlstrasse 28, CH-3012, Berne, Switzerland

5.1 ABBREVIATIONS

IPTG, isopropyl 1-thio-β-D-galactopyranoside; **DTT**, dithiotreitol; **EDTA**, ethylenedinitrilotetraacetic acid; **BN-PAGE**, blue native-polyacrylamide gel electrophoresis; **DM**, n-decyl-β-D-maltoside; **TEM**, transmission electron microscopy; **TM**, transmembrane regions; **Ni-NTA**, nickel-nitrilotriacetic acide; $\mathbf{M}_{\mathbf{obs}}$, apparent molecular mass.

5.2 ABSTRACT

Cellular peptide uptake has been extensively studied in the past. In contrast, structural information on peptide transporters is sparse. Here, we have cloned, overexpressed, purified and biochemically characterized the YbgH protein from Escherichia coli, a prokaryotic member of the peptide transporter (PTR) family. Two-dimensional (2D) crystallization of YbgH yielded first tubular crystals that allowed the determination of a projection structure at 27 Å resolution. This low-resolution structure of YbgH represents the first structural view of a member of the PTR family at the given resolution. Furthermore, the presented work sets the basis for future structural works with the peptide transporter YbgH using 2D crystals and electron crystallography.

5.3 INTRODUCTION

Peptide molecules are mostly disabled to cross lipid membranes by simple diffusion, because of size and polarity. Integral peptide transporters mediate the translocation of peptides of two to six residues through cell membranes in an energy dependent manner (1-5). The reflow of disposed peptides plays a crucial role in the energy balance of bacteria, fungi, plants and animals. Peptide uptake provides amino acids, nitrogen and carbon for the synthesis of new proteins and metabolites. In addition, peptide transport is involved in a wide variety of biological functions such as signalling process, regulation of gene expression, biological warfare and metabolic adaptation (6).

Based on function and sequence similarity peptide transporters can be classified in three main families: The ATP binding cassette (ABC) family (7), the oligopeptide transporter (OPT) family (8, 9) and the peptide transporter (PTR) family (10; 11). Peptide transporters of the ABC family are composed of multiple proteins with one or more ABC components. Bacterial ABC peptide transporters require additionally a substrate binding protein that is involved in the process of substrate shuttling (12-14). Proteins of the OPT family are found in yeast, fungi and plants only (15-17). Recently nine OPT orthologs were identified in *Arabidopsis* (*AtOPT1* to *AtOPT9*) providing first evidences of multiple peptide transport systems in plants (18). The PTR family was uncovered by multiple sequence analysis, because a number of discovered peptide transporters distinct from the already known group of ABC transporters. Surprisingly, the sequence alignment of these transporters demonstrated a high number of similarity and identity. Two conserved motives were identified exclusively within this group of proteins which was designated the peptide transporter (PTR) family (11). The high conservation of these signature sequences suggests that they may play an important structural, functional or biogenic role. Subsequently members of the PTR family were found in other species from bacteria, yeast, plants, and invertebrates to vertebrates. Functionally well characterized peptide transporters of the PTR family are DtpT from *Lactococcus lactis* (19, 20), Ptr2 from *Saccharomyces cerevisiae* (21) and the mammalian transporters PEPT1 and PEPT2 (11, 22). In *E. coli* four PTR members are reported that contain the specific family motives: YdgR (TppB, DtpA), YhiP (DtpB), YjdL and YbgH (5, 23, 24). DtpA and DtpB from *E. coli* were functionally characterized as proton dependent oligopeptide transporters with substrate recognition patterns similar to mammalian PEPT1 (25, 26). Because all of the characterized members are proton dependent transporters driven by the proton motive force, the PTR family is also referred as POT or proton dependent oligopeptide transporter family (10). All PTR transporters share a molecular mass of about 450 to 600 amino acid residues with 12 putative α-helical transmembrane segments (TMSs) predicted. In the case of the prokaryotic peptide transporter DtpT from *Lactococcus lactis*, the 12 TMS structure with cytoplasmic located N- and C- termini has experimentally been proven by a gene fusion approach (27).

Functional data characterizing peptide transport systems is continuously increasing, while no structural information is available. The structure of membrane proteins is generally difficult to analyze using X-ray diffraction or NMR techniques, due to problems with crystallization or solubility. Cryo-transmission electron microscopy combined with electron crystallography of two dimensional (2D) crystals represents an attractive alternative to determine membrane protein structures (28). The advantage of 2D crystals is that the membrane protein is analyzed in its native environment, the lipid bilayer.

Here we report the cloning, expression, purification and the first low-resolution structure of YbgH, a PTR family member. Negative-stain TEM and blue native-polyacrylamide gel electrophoresis (BN-PAGE) of detergent-solubilized YbgH revealed a monomeric state. Crystallization experiments yielded well-ordered tubular 2D crystals that were analyzed by transmission electron microscopy (TEM) techniques. Image processing of negatively stained crystals enabled us to derive a map of the structure in projection at 27 Å resolution. This work presents a solid step towards the structural analysis of YbgH at high-resolution using 2D crystals and electron crystallography. Importantly, prokaryotic YbgH presents an excellent model for understanding the molecular architecture of transporters from the PTR family.

5.4 MATERIALS AND METHODS

Cloning − Genomic DNA from *E. coli* strain O157:H7 was prepared with the Qiagen DNeasy Kit and *ybgH* was obtained by PCR using the primers AAAAAGCTTATGAATAAACACGCATCACAG and AAACTCGAGAGACTCCAGCGCCAGCGC that add Hind III and Xho I for cloning. PCR products were digested with Hind III and Xho I and ligated into pET-21 vector (T7 promotor, C-terminal hexahistidine-tag, Novagen/Merck, Darmstadt, Germany). A ribosomal binding site (rbs) (AAGGAG) was added 7 bases 5' of the coding region to improve translation of the protein. The DNA construct was verified by sequencing.

Overexpression in E. coli − Expression experiments were carried out with *E. coli* BL21(DE3) pLysS (Novagen/Merck, Darmstadt, Germany) from freshly streaked glycerin stocks. Cultures were grown in LB Medium supplemented with 100 µg/ml ampicillin to OD600 of about 0.8 and protein production was induced by adding 0.1 mM IPTG (final concentration). Cells were harvested by centrifugation (8000 x g, 8 min) after 3 h shaking incubation at 37 °C.

Membrane preparation − Cell pellets from 500 ml culture were resuspended in 25 ml lysis buffer (10 mM Hepes/Tris pH 7.4, 1 mM DTT, 0.5 mM EDTA) and broken by sonication (10 cycles of 30 s). After a short low speed centrifugation to separate unbroken cells (4,500 x g, 8 min) and a short high speed centrifugation to pellet the outer membrane (120,000 x g, 5 min) the inner membranes from the supernatant were collected by centrifugation at 120,000 x g for 1.5 h (all at 4 °C). The pellets were resuspended in 1 ml buffer (20% glycerol, 10 mM Hepes/Tris pH 7.4, 0.5 mM Tris-2-carboxyethyl-phosphine (TCEP), frozen in liquid nitrogen and stored at -80 °C.

Purification of YbgH − Frozen *E .coli* membranes with overexpressed YbgH were solubilized in 2% DM (n-decyl-ß-D-maltopyranoside), 20 mM Tris-HCl pH 8, 300 mM NaCl, 10% glycerol and 0.01% NaN_3 for 1 hours at 4 °C under gentle agitation. The protein concentration during solubilization was 2 mg/ml. After ultracentrifugation (100,000 x g, 50 min at 4°C) the supernatant was diluted two-fold with 20 mM Tris-HCl pH 8, 300 mM NaCl, 0.4% DM, 5 mM histidine, 10% glycerol, 0.01% NaN_3 (washing buffer) and bound to Ni-NTA Superflow beads (Qiagen) for 2 h at 4 °C. The beads were then loaded onto a spin column (Promega), washed with washing buffer and

eluted with the same buffer containing 200 mM histidine. Concentration of purified YbgH was spectrophotometrically determined measuring the absorbance at 280 nm and using a molar extinction coefficient of 103,750 M^{-1} cm^{-1}. This value was calculated from the amino acid sequence of recombinant YbgH using the ProtParam tool from the ExPASy proteomics server (http://www.expasy.org/tools/protparam.html).

Blue native gel electrophoresis − Linear 5–12% gradient gels for blue native polyacrylamide gel electrophoresis (BN-PAGE) were prepared and run as previously described by Schägger *et al.* 1991. Thyroglobulin (669 kDa), ferritin (440 kDa), lactate dehydrogenase (140 kDa), and bovine serum albumin (66 kDa) were used as standard proteins.

Negative stain TEM − DM-solubilized YbgH protein at ~3 µg/ml was adsorbed for ~10 sec to parlodion carbon-coated copper grids rendered hydrophilic by glow discharge at low pressure in air. Grids were washed with three drops of double-distilled water and stained with 2 drops of 0.75% uranyl formate. Grids of YbgH 2D crystallization trials were prepared similarly, but included a longer adsorption time of ~30 sec. Electron micrographs were recorded at a magnification of x50'000 on Eastman Kodak Co. S0-163 sheet films with a Hitachi H-7000 electron microscope operated at 100 kV.

2D crystallization of YbgH − Purified YbgH protein (3 mg/ml) solubilized in DM was mixed with lipids solubilized in the same detergent: 5 mg/ml lipid in 1% detergent in 20 mM Tris-HCl pH 8, 150 mM NaCl, 10% glycerol, 0.01% NaN_3 at lipid-to-protein ratios (LPRs) of 0.2–0.8 (w/w). The final protein concentration varied between 1.5 and 3 mg/ml depending on the LPR. The samples were dialyzed against detergent-free buffer for five days at room temperature. The tested crystallization conditions are listed in Table 1.

Freeze-drying and heavy metal shadowing − Tubular 2D crystals of YbgH were adsorbed to glow-discharged carbon coated 400 mesh grids, washed twice with double-distilled water, blotted on filter paper, and immediately frozen in liquid nitrogen. Submerged in liquid nitrogen at all times, the grids were transferred to the pre-cooled cold stage of the MIDILAB, a complete freeze-drying and shadowing unit that is mounted permanently onto the column of a Philips CM12 TEM (29). Freeze-drying, unidirectional heavy-metal shadowing with Ta/W at an elevation angle of 45°, operation of the TEM and image acquisition were performed as described (30).

Image processing − Electron micrographs of negatively stained YbgH crystals were evaluated by optical diffraction. Micrographs of well diffracting tubular crystals were digitized with a Heidelberg Primescan D 7100 at 4 Å per pixel resolution. Image processing, i.e. measurement of lattice vectors, calculation of power spectra and correlation averages, was performed with the SEMPER image processing system (31).

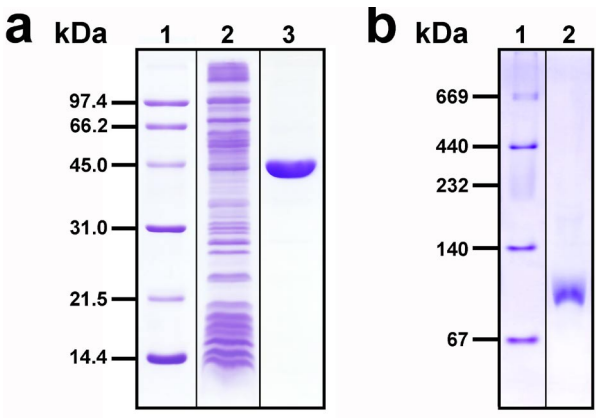

Figure 1. SDS-PAGE and BN-PAGE of pure YbgH protein. (a) Purification of His-tagged YbgH by nickel affinity chromatography: low molecular mass standard (lane 1), DM-solubilized *E. coli* total membranes containing overexpressed YbgH (lane 2) and purified YbgH protein (lane 3). Denatured YbgH runs as a prominent band at ~45 kDa by SDS-PAGE. (b) BN-PAGE of YbgH. Protein standard (lane 1) and purified YbgH (lane 2). DM-solubilized YbgH protein migrates at ~85 kDa in the 5-12% linear gradient polyacrylamide gel. Applied amount of protein per lane: ~10 µg (lane 2, panel (a)), ~5 µg (lane 3, panel (a)), ~2 µg (lane 2, panel (b)).

5.5 RESULTS

Cloning, expression and purification of YbgH − The *ybgH* gene from *E. coli* was amplified from genomic DNA by PCR using gene specific primers and cloned into the pET-21 vector fused with a C-terminal hexahistidine-tag (pET-21-*ybgH*-His_6) as previously described for *dtbA* (*ybgH*) (25). For homologous over-expression of YbgH the vector was transformed into *E. coli* BL21(DE3)pLysS. Isolated *E. coli* inner membranes were solubilized with DM and His-tagged YbgH was purified in a single step by nickel affinity chromatography. The yield was in the range of 2-3 milligrams of pure YbgH protein per litre of *E. coli* cell culture. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of DM-solubilized inner membranes and of purified YbgH protein is shown in Figure 1a, lanes 2 and 3. Denatured YbgH protein migrated as a single band with an apparent mass of ~45 kDa in contrast to its molecular mass of 55.2 kDa. The assignment of this band to YbgH was confirmed by Western blot analysis using an anti-His antibody (data not shown).

Characterization of YbgH by BN-PAGE and negative-stain TEM − BN-PAGE was performed to unveil the oligomeric state of purified YbgH protein. The blue native-gel in Figure 1b revealed one single band with an apparent mass of ~85 kDa for DM-solubilized YbgH. Furthermore, we performed negative-stain TEM to assess protein quality, shape, dimensions and low-resolution structure of YbgH. The homogeneity of purified YbgH is documented in Figure 2a. Well preserved top-view particles exhibited a donut-like shape with a central, shallow depression (see Figure 2b, gallery) and had a diameter of 8.0±0.5 nm (n=100).

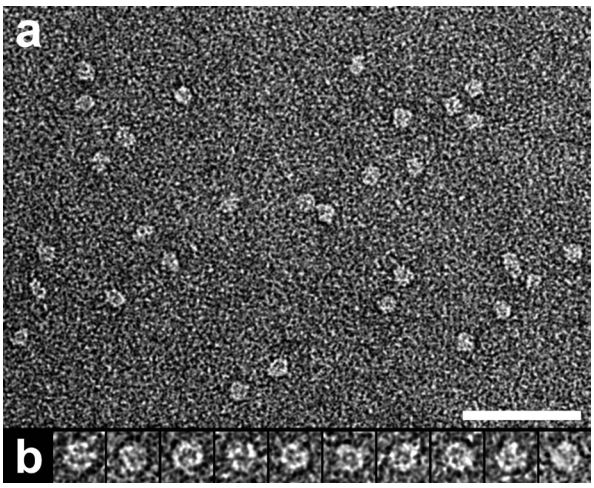

Figure 2. TEM of single purified YbgH proteins. (a) TEM of negatively stained YbgH proteins. The homogeneity of the purified protein is reflected in the electron micrograph.
(b) Gallery of well preserved top-view particles. YbgH particles are donut-shaped with a central depression and have a diameter of 8.0±0.5 nm (n=100). The scale bar represents 50 nm, and the frame size of the magnified particles in the gallery is 13 nm.

2D crystallization of YbgH − Purified, DM-solubilized YbgH was subjected to 2D crystallization in the presence of different lipids and additives, at different pHs, ionic strengths, temperatures and lipid-to-protein ratios (LPRs; see Table 1 for a summary). Most of the tested lipids were able to reconstitute YbgH into tubular crystals at LPRs between 0.3 and 0.4. LPRs higher than 0.4 produced densely packed vesicles while LPRs below 0.3 resulted in protein aggregation (data not shown). Interestingly, YbgH had the intrinsic tendency to form tubular crystals instead of planar 2D crystals. The crystal quality and length of the tubes was optimized by varying the crystallization parameters listed in Tabel 1. Furthermore, the stability of the YbgH protein could be improved by the presence of 10% glycerol and 250 mM mannitol in the dialysis buffer. From all tested parameters the best crystals were grown with *E. coli* polar lipids (LPRs 0.3-0.4) in 300 mM NaCl, 250 mM mannitol, 10% glycerol and 20 mM Tris-HCl pH 8. Unfortunately, no crystallization condition was found to induce the growth of large vesicular or sheet-like 2D crystals suitable for high-resolution studies by TEM and electron crystallography.

Negative stain TEM of tubular YbgH crystals − Negative stain TEM of on parlodion carbon-coated electron microscopy grids adsorbed and flattened YbgH tubes revealed well-ordered 2D crystals (Figure 3a). The width of the tubes was 100-110 nm and quite constant. In contrast to the latter, the length of the tubes

Table 1. Summary of 2D-crystallization conditions

Parameter	Range tested	Best condition
Lipid type	*E. coli* polar and soy bean lipids, DLPC, DMPC, DMPS, DMPG, DOPC, DOPA,DOPE, DOPG, DOTAP, POPC, POPS, POPE, POPA	*E. coli*
LPR	0.2–0.8 (0.05 increments)	0.3-0.4
pH	5–10 (0.5 increments)	8
NaCl	10, 150, 300, 500 and 1000 mM	300 mM
KCl	50 and 150 mM	none
$MgCl_2$	1, 5, 25 mM	none
Stabilizers	250 mM betaine, 250 mM glucose, 250 mM mannitol	250 mM mannitol, 10% glycerol
Antioxidant	2-5 mM DTT	none
Tempearture	4°C, RT, 37°C	RT

Lipid abbreviations: DLPC, 1,2-dilauroyl-sn-glycero-3-phosphocholine; DMPC, 1,2-dimyristoyl-sn-glycero-3 phosphocholine; DMPS, 1,2-dimyristoyl-sn-glycero-3-phospho-L-serine (sodium salt); DMPG, 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (sodium salt); DOPC, 1,2-dioleoyl-*sn*-glycero-3-phosphocholine; DOPA, 1,2-dioleoyl-*sn*-glycero-3-phosphate (monosodium salt); DOPE, 1,2-dilinoleoyl-*sn*-glycero-3-phosphoethanolamine; DOPG, 1,2-dioleoyl-*sn*-glycero-3-(phospho-*rac*-(1-glycerol)) (sodium salt); DOTAP, 1,2-dioleoyl-3-trimethylammonium-propane (chloride salt); POPC, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; POPS, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-(phospho-L-serine] (sodium salt); POPE, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine; POPA, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphate (monosodium salt**).** Used buffers: pH 5-5.5 (cirtic acide-NaOH), pH 6-6.5 (Mes-NaOH), pH 7-9 (Tris-HCl), pH 9.5-10 (CAPS-HCl)

varied from nano- to micrometers. Power spectra calculated from such flattened tubular crystals displayed a distinct periodic pattern arising from two superimposed collapsed layers (Figure 3b). Such power spectra represented therefore the superposition of two separate diffraction patterns. To distinguish between the two sets of diffraction spots, tubular crystals of YbgH were freeze-dried and unidirectionally shadowed with Ta/W (Figure 3c). This method revealed only the top layer allowing the lattice parameters of the tubular crystals to be defined unambiguously from the diffraction pattern shown in Figure 3d. The unit cell had a rhombic shape and lattice parameters of a=110 Å; b=105 Å; γ=97°.

From several dozens of micrographs of negatively stained YbgH tubes the best four images were selected by optical diffraction for the calculation of the correlation average displayed in Figure 4. Top and bottom layers of collapsed tubes were processed separately. Interestingly, one of the two layers very often indicated better diffraction, suggesting better conservation after adsorption. The four best correlation averages were merged together to yield a projection map at 27 Å (Figure 4). The latter featured alternating horizontal rows of stronger and weaker intensities.

5.6 DISCUSSION

Peptide uptake in *E. coli* has been extensively studied during the last two decades. In contrast, structural information on peptide transporters is sparse. To our knowledge, only a low-resolution structure based on negative stain TEM of single DtpA (YdgR) peptide transporter particles has been reported so far (25). Here, we have presented the first projection structure of a PTR family member at 27 Å resolution determined by negative stain TEM of tubular YbgH protein crystals.

On our aim to elucidate the structure of peptide transport proteins, we have cloned, overexpressed, purified and biochemically characterized YbgH. Overexpression of YbgH yielded 2-3 milligrams

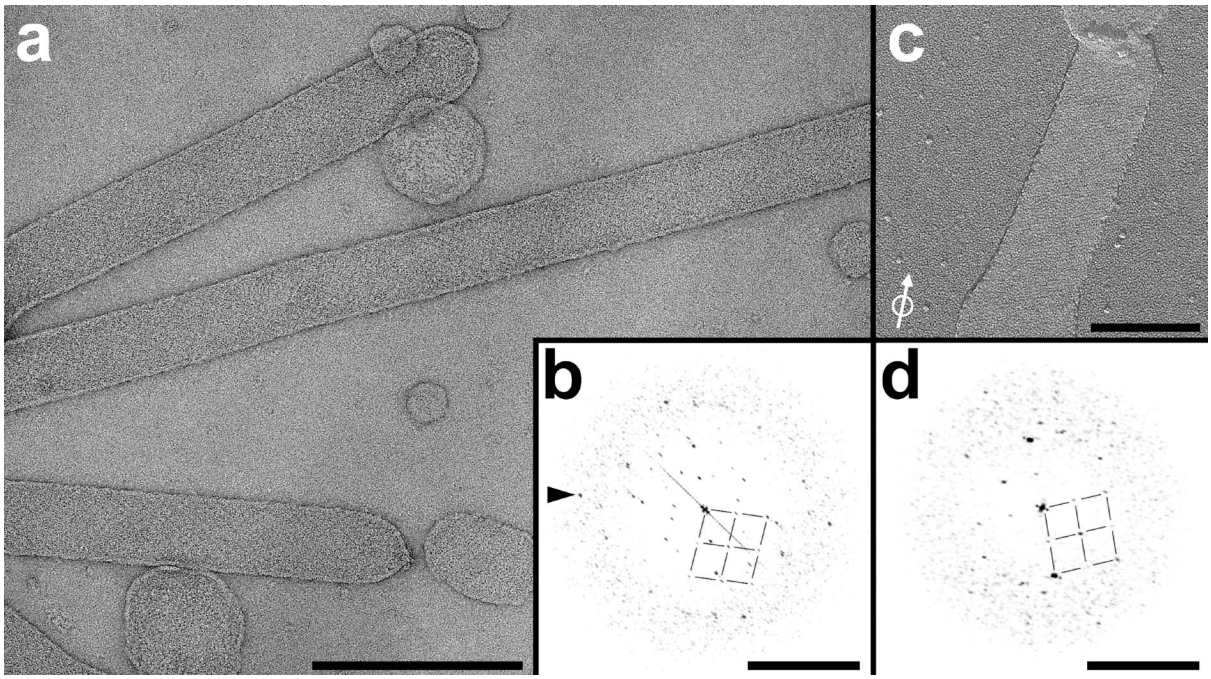

Figure 3. Negatively stained tubular crystals of purified YbgH. (a) Overview electron micrograph of crystalline YbgH Tubes. (b) The power spectrum calculated from a flattened tubular YbgH crystal indicates crystallinity and diffraction up to 27 Å resolution (arrowhead). Two separate sets of diffraction spots are distinct (one indicated by lattice lines) arising from the two superimposed collapsed layers. (c) Electron micrograph of a freeze-dried and unidirectionally shadowed YbgH tube. The white arrow indicates the shadowing direction. (d) The power spectrum calculated from (c) shows the diffractions spots derived from the top layer (indicated by lattice lines) of the collapsed tube. Scale bars 300 nm (a), $(3\ nm)^{-1}$ (b), 150 nm (c), $(3\ nm)^{-1}$ (d).

of pure protein per liter of bacterial cell culture setting the stage for 2D crystallization trials. SDS-PAGE of purified YbgH indicated an apparent mass of ~45 kDa in contrast to its molecular mass of ~55 kDa. Such SDS-PAGE running anomalies of membrane proteins, where the abnormal migration behaviour causes an underestimation of the true protein mass, are frequently reported (32, 33). Interestingly, also two other PTR family members from *E. coli*, i.e. DtpA and DtpB, showed similar running behaviours by SDS-PAGE (25, 26).

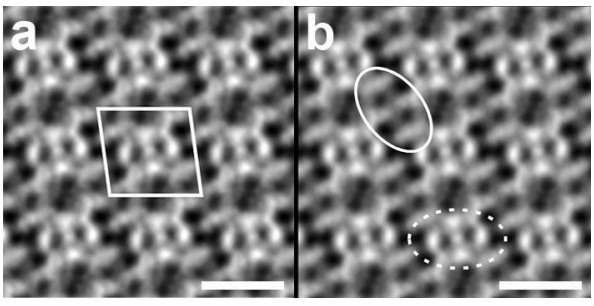

Figure 4. Projection structure of YbgH at 27 Å resolution. (a) The unit cell is marked. The lattice parameters are a=110Å; b=105 Å; γ=97°. (b) The p1 projection map features alternating horizontal rows of stronger and weaker intensities. Eyeglasses-like structures, each containing two central depressions are discerned: see solid and broken ellipses. Scale bars represent 100 nm (panels a and b).

To ascertain whether YbgH exists in a monomeric or oligomeric form when solubilized in DM, we performed BN-PAGE. On BN gels, DM-solubilized YbgH migrated as a single band at an apparent molecular mass of ~85 kDa. Assuming an average mass of ~55 kDa for recombinant YbgH and a DM/coomassie brilliant blue G-250 micelle of ~30 kDa (34) bound to the proteins during BN-PAGE, YbgH exists as monomer when solubilized in DM. The latter was further corroborated by TEM of negatively stained YbgH proteins that exhibited a donut-like shape with a diameter of ~8.0 nm, similar to monomeric detergent-solubilized DtpA (YdgR) proteins (25).

The high amounts of pure and homogeneous YbgH protein formed an excellent basis for its

2D crystallization. As shown in Figure 3, YbgH had the propensity to form tubular crystals. Correlation averaging of flattened tubular crystals revealed alternating horizontal rows of stronger and weaker intensities (Figure 4). This suggests an upside down and downside up orientation of the YbgH proteins inside the lipid bilayer exposing cytoplasmic and periplasmic face. The unit cell of the YbgH crystal was rhombic with an area of ~11464 Å^2. To estimate the number of YbgH monomers in the unit cell, we calculated the area per AdiC monomer in the previously reported 2D crystals (35): AdiC is the L-arginine/agmantine antiporter from *E. coli* and member of the amino/polyamine/organocation superfamily, has a comparable mass to YbgH, i.e. ~49 kDa, and is also predicted to consist of twelve α-helical TMS. Thus in 2D AdiC crystals, one monomer occupied an area of 2737 Å^2. Division of the total unit cell area in YbgH crystals (~11464 Å^2) by the area occupied by one AdiC protein (2737 Å^2) results in ~4.2 YbgH monomers per unit cell. Considering that the molecular mass of AdiC is slightly below that of YbgH, four YbgH monomers are housed per unit cell.

Projection maps featured alternating horizontal rows of stronger and weaker intensities (Figure 4). Assuming uneven staining of the flattened 2D crystals, this feature indicates rows of up and down oriented YbgH proteins, i.e. proteins exposing the cytoplasmic and periplasmic face, respectively. Such rows indicated eyeglasses-like structures, each containing two central depressions: see solid and broken ellipses in Figure 4b. One unit cell contained two such eyeglasses-like structures, one with strong and one with weak intensities. This suggests that one unit cell contains four YbgH proteins in total, two exposing the cytoplasmic and two the periplasmic side. Furthermore, alternating horizontal rows of stronger and weaker intensities suggest that the peptide mass with respect to the membrane plane is asymmetric in the YbgH molecule, e.g. more peptide mass protruding out of the membrane on the cytoplasmic than on the periplasmic side, or vice versa.

This low-resolution projection map of YbgH has provided a first structural view of a member of the PTR family at 27 Å resolution. Obtaining a high-resolution structure of YbgH is crucial for the understanding of the function and mechanism of peptide transport proteins. Our present work with YbgH sets the basis for future structural analysis of a peptide transporter using 2D crystals and electron crystallography.

5.7 REFERENCES

1. Becker JM. and Naider F. (1980) *In Microorganisms and Nitrogen Sources* 257-279
2. Higgins CF. and Payne JW. (1980) *In Microorganisms and Nitrogen Sources* 211-258
3. Matthews DM. (1987) *Contr Infusion TherClin Nutr* 6-53
4. Payne JW. and Smith MW. (1994) *Adv Micro Physiol* **36** 1-80
5. Meredith D. and Boyd CA. (2000) *Cell Mol Life Sci* **57** 754-78
6. Detmers FJ., Lanfermeijer FC. and Poolman B. (2001) *Res Microbiol* **152** 245-258
7. Higgins CF. (1992) *Annu Rev Ceil Biol* **8** 67-113
8. Lubkowitz MA., Hauser L., Breslav M., Naider F. and Becker JM. (1997) *Microbiology* **143** 387-96
9. Hauser M., Narita V., Donhardt AM., Naider F. and Becker JM. (2001) *Mol Membr Biol* **18** 105-12
10. Paulsen IT. and Skurray RA. (1994) *Trends Biochem Sci* **18** 404
11. Steiner HY., Naider F. and Becker JM. (1995) *Mol Microbiol* **16** 825-34
12. Abouhamad WN., Manson M., Gibson MM. and Higgins CF. (1991) *Mol Microbiol* **5** 1035-1047
13. Higgins CF. (2001) *Res Microbiol* **152** 205-210
14. Hiles ID., Gallagher MP., Jamieson DJ. and Higgins CF. (1987) *J Mol Biol* **195** 125-142
15. Lubkowitz MA., Barnes D., Breslav M., Burchfield A., Naider F. and Becker JM. (1998) *Mol Microbiol* **28** 729-41
16. Hauser M., Donhardt AM., Barnes D., Naider F. and Becker JM. (2000) *J Biol Chem* **275** 3037-41
17. Stacey G., Koh S., Granger C. and Becker JM. (2002) *Trends Plant Sci* **7** 257-63
18. Koh S., Wiles AM., Sharp JS., Naider FR., Becker JM. and Stacey G. (2002) *Plant Physiol* **128** 21-9
19. Hagting A., Kunji ERS., Leenhouts KJ., Poolman B. and Konings WN (1994) *J Sfo/Chem* **269** 11391-11399
20. Nakajima H., Hagting A., Kunji ER., Poolman B. and Konings WN. (1997) *Appl Environ Microbiol* **63** 2213-7
21. Perry JR., Basrai MA., Steiner HY., Naider F. and Becker JM.(1994) *Mol Cell Biol* **14** 104-15.
22. Daniel H., Spanier B., Kottra G. and Weitz D. (2006) *Physiology* (Bethesda) **21** 93-102
23. Smith MW., Tyreman DR., Payne GM., Marshall NJ. and Payne JW. (1999) *Microbiology* **145** 2891-901
24. Goh EB., Siino DF. and Igo MM. (2004) *J Bacteriol* **186** 4019-24
25. Weitz D., Harder D., Casagrande F., Fotiadis D., Obrdlik P., Kelety B. and Daniel H. (2007) *J Biol Chem* **282** 2832-9
26. Harder D., Stolz J., Casagrande F., Obrdlik P., Weitz D., Fotiadis D. and Daniel H. (2008) *FEBS J* (in press)
27. Hagting A., vd Velde J., Poolman B., Konings WN. (1997) *Biochemistry* **36** 6777-85
28. Stahlberg H, Fotiadis D, Scheuring S, Rémigy H, Braun T, Mitsuoka K, Fujiyoshi Y, Engel A. (2001) *FEBS Lett* **504** 166-72
29. Gross H., Krusche K. and Tittmann P. (1990) In XIIth International Congress of Electron Microscopy, San Francisco Press Inc., Seattle.
30. Walz T., Tittmann P., Fuchs KH., Müller DJ., Smith BL., Agre P., Gross H. and Engel A. (1996) *J Mol Biol* **264** 907-918
31. Saxton WO. (1996) *J Struct Biol* **116** 230-61996
32. Raunser S., Haase W., Bostina M., Parcej DN. and Kühlbrandt W (2005) *J Mol Biol* **351** 598-613.
33. Vinothkumar KR., Raunser S., Jung H.

and Kühlbrandt W. (2006) *J Biol Chem* **281** 4795-801
34. Reig N., del Rio C., Casagrande F., Ratera M., Gelpí JL., Torrents D., Henderson PJ., Xie H., Baldwin SA., Zorzano A., Fotiadis D. and Palacín M. (2007) *J Biol Chem* **282** 13270-81
35. Casagrande F. *(to be submitted)*

6. Functional and Structural Characterization of a Prokaryotic Peptide Transporter with Features Similar to Mammalian PEPT1

Dietmar Weitz*, Daniel Harder*, Fabio Casagrande§, Dimitrios Fotiadis§, Petr Obrdlik&, Bela Kelety& and Hannelore Daniel*

* Molecular Nutrition Unit, Department of Food and Nutrition, Technical University of Munich, 85350 Freising, Germany

§ M.E. Müller Institute for Structural Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland

& Iongate Biosciences GmbH, D528 Industriepark Höchst, D-65926 Frankfurt, Germany

6.1 ABBREVIATIONS

IPTG, isopropyl 1-thio-β-D-galactopyranoside; **CCCP**, carbonyl cyanide *p*-chlorophenylhydrazone; **TCEP**, Tris-2-carboxyethyl-phosphine; **BN-PAGE**, blue native-polyacrylamide gel electrophoresis; **DDM**, n-dodecyl-β-D-maltoside; **TEM**, transmission electron microscopy; **TM**, transmembrane regions; **Ni-NTA**, nickel-nitrilotriacetic acide; **rbs**, ribosomal binding site; **TBS**, Tris-buffered saline; **AMCA**, N_{ε}-7-amino-4-methyl-coumarin-3-acetic acid; **MES**, 4-morpholineethanesulfonic acid; $\mathbf{M_{obs}}$, apparent molecular mass.

6.2 ABSTRACT

The *ydgR* gene of *Escherichia coli* encodes a protein of the proton-dependent oligopeptide transporter (POT) family. We cloned YdgR and overexpressed the His-tagged fusion protein in *E. coli* BL21 cells. Bacterial growth inhibition in the presence of the toxic phosphonopeptide alafosfalin established YgdR functionality. Transport was abolished in the presence of the proton ionophore carbonyl cyanide p-chlorophenylhydrazone, suggesting a proton-coupled transport mechanism. YdgR transports selectively only di- and tripeptides and structurally related peptidomimetics (such as aminocephalosporins) with a substrate recognition pattern almost identical to the mammalian peptide transporter PEPT1. The YdgR protein was purified to homogeneity from *E. coli* membranes. Blue native-polyacrylamide gel electrophoresis and transmission electron microscopy of detergent-solubilized YdgR suggest that it exists in monomeric form. Transmission electron microscopy revealed a crown-like structure with a diameter of ~8 nm and a central density. These are the first structural data obtained from a proton-dependent peptide transporter and the YgdR proteins seems an excellent model for studies on substrate and inhibitor interactions as well as on the molecular architecture of cell membrane peptide transporters.

6.3 INTRODUCTION

Peptide transporters are integral membrane proteins that mediate the cellular uptake of di- and tripeptides and a variety of peptidomimetics (for review, see Refs. 1-4). They are found in bacteria, yeast, plants, invertebrates and vertebrates. In vertebrates, the two peptide transporter proteins PEPT1 (SLC15A1) and PEPT2 (SLC15A2) are expressed predominantly in brush border membranes of small intestine (PEPT1), kidney (PEPT1 and PEPT2) and lung (PEPT2). In these transport proteins substrate flux is coupled to proton movement down an electrochemical proton gradient with the membrane potential as the main driving force. PEPT1 and PEPT2 accept essentially all 400 possible dipeptides and 8000 possible tripeptides composed of L-α amino acids as substrates. Moreover, they also transport a large spectrum of therapeutic drugs like β-lactam antibiotics, selected angiotensin-converting enzyme inhibitors and peptidase inhibitors and thereby determine their bioavailability and pharmacokinetics. Certain drugs with an intrinsic low oral bioavailability like L-DOPA and acyclovir have by coupling to an amino acid (L-DOPA-Phe and Val-acyclovir) turned into substrates of peptide transporters with markedly improved availability (5,6). Peptide transporters are therefore considered as important and potent drug delivery systems. Although functionally characterized in extenso, very little is known about the structure of peptide transporter proteins. Twelve trans-membrane domains are predicted (7) and amino acid residues critical for transport activity have been identified in particular in transmembrane domains 2, 3, 4 and 10 by functional analysis of mutants (8-12).

Mammalian peptide transporters are part of the PTR2 family of membrane transporters characterised by two signatures that are conserved in all family members (13). The first is a region that begins at the end of the second putative transmembrane domain, including the following first cytoplasmic loop as well as the third transmembrane domain. The second motif corresponds to the core region of the fifth transmembrane region. Besides the mammalian PEPT1 and PEPT2 proteins, the PTR2 family includes the yeast peptide transporter PTR2, DtpT from *Lactococcus lactis* and numerous "orphan" transporters for which function is not known yet. Most orphan transporters are found in prokaryotic organism, *e.g.* the four members *ybgH*, *ydgR*, *yhiP* and *yjdL* in *E. coli*. Although these gene sequences belong to the same family, the function of the corresponding proteins may be quite different.

Here we describe the cloning of the *ydgR* gene, which was identified by sequence analysis as the *E. coli* homologue of *tppB* from *Salmonella thyphimurium* (14). Based on growth experiments with mutant bacterial strains *tppB* was identified already in 1984 as a tripeptide permease (15,16). However, neither TppB nor YdgR have been characterised biochemically nor with respect to mode of function. We overexpressed YdgR as a fusion protein with the IPTG-inducible pET-expression system. YdgR encodes a proton-dependent peptide transporter with a broad substrate specificity ranging from di- and tripeptides to a variety of related peptidomimetics like β-lactam antibiotics. The YdgR protein was purified to homogeneity and functionally reconstituted into proteoliposomes. Analysis by blue native polyacrylamide gel electrophoresis (BN-Page) and transmission electron microscopy (TEM) of detergent-solubilized YdgR suggests a monomeric state of the protein. In addition, TEM provided the first structural data on a proton-dependent di- and tripeptide transporter.

6.4 MATERIALS AND METHODS

Cloning and expression of the YdgR transport protein in E. coli - Genomic DNA from *E. coli* strain O157:H7 was prepared with the Qiagen DNeasy Kit. The *ydgR* gene was cloned from genomic DNA by PCR with primers (5'-3') AGCTTATGTCCACTGCAAACCAAAAAC and AAACTCGAGCGCTACGGCTGCTTTC-GC. PCR products were digested with HindIII/XhoI and ligated into pET-21 vector (T7 promotor, C-terminal hexahistidine tag, Novagen). A ribosomal binding site (rbs) (AAGGAG) was added 7 bases 5' of the coding region to improve translation of the protein. DNA construct (pET-21-rbs-YdgR-His) was verified by sequencing. Expression experiments were carried out with freshly transformed *E. coli* BL21(DE3)pLysS. Cultures were grown in LB Medium supplemented with 100 µg/ml ampicillin to A_{600} of ~1 and protein expression was induced by adding 0.1 mM IPTG (final concentration). After 3 h of incubation at 37 °C cells were harvested for biochemical or functional analysis. A second expression vector (pET-21b-rbs-T7-YdgR-His) was constructed by cutting pET-21-rbs-YdgR-His with HindIII/XhoI and ligating the insert into pET-21b vector.

Western blot analysis - For Western blot analysis, cells from 1 ml culture were pelleted and resuspended in lysis buffer (10 mM Hepes-NaOH pH 7.4, 0.5 mM EDTA, 1 mM DTT and protease inhibitor cocktail 1:500 (Sigma)) with lysozyme (1 µg) added. After 1 h of incubation on ice, bacterial DNA was degraded for 30 min by 1 unit benzonase in the presence of 3 mM $MgCl_2$. Membranes were pelleted by centrifugation at 18000 x *g* for 30 min and solubilized in a buffer containing 10 mM Hepes-NaOH pH 7.4, 150 mM NaCl, 0.5 mM EDTA, 1 mM DTT and 1% *n*-dodecyl-β-D-maltoside (DDM). After 15 min on ice unsolubilized material was removed by centrifugation at 40000 x *g* for 45 min. Proteins were separated by SDS-PAGE and blotted onto polyvinylidene difluorid membranes (Millipore). Filters were blocked by incubation for 1 h with 1% (w/v) milk powder in Tris-buffered saline (TBS; 137 mM NaCl, 3 mM KCl and 25 mM Tris-Cl pH 7.5) followed by incubation for 60 min with anti-His antibody (1:2000 dilution, Novagen) in TBS-T (TBS, 0.05% Tween-20). Filters were washed twice in the same buffer and then incubated for 30 min with secondary antibody (1:5000 dilution, goat anti-mouse-HRP, Santa Cruz). Filters were first washed twice in TBS-T and twice with TBS (10 min per wash). Labelled proteins were detected using the ECL system (Pharmacia).

Transport assays - Transport assays were performed *in vivo* with cells 3 h after induction with IPTG (see above) with the fluorescent dipeptide β-Ala-Lys-N_ε-7-amino-4-methyl-coumarin-3-acetic acid (β-Ala-Lys-AMCA) (custom-synthesis by Biotrend, Cologne, Germany). β-Ala-Lys-AMCA was previously established as a reporter substrate for peptide transport (17,18). Around 5 x 10^9 cells were harvested by centrifugation and resuspended in 1.5 ml modified Krebs-Buffer

(25 mM Hepes/Tris 7.4, 140 mM NaCl, 5.4 mM KCl, 1.8 mM $CaCl_2$, 0.8 mM $MgSO_4$ and 5 mM glucose). The assay volume of 100 µl was made up with 40 µl bacteria cells (1.3 x 10^8 cells), 10 µl of a 500 µM β-Ala-Lys-AMCA stock solution (final concentration 50 µM), and 50 µl of Krebs-Buffer (control) or a competitor solution. Uptake was performed for 15 min at 37 °C and stopped by washing the cells twice with ice-cold Krebs-Buffer by centrifugation. Uptake of β-Ala-Lys-AMCA was quantified by fluorescence (excitation at 340 nm and emission at 460 nm, Thermo Varioscan).

Purification of the YdgR protein - Cell pellets from 600 ml culture were resuspended in 25 ml lysis buffer (10 mM Hepes/Tris pH 7.4, 1 mM DTT, 0.5 mM EDTA and a Sigma protease inhibitor cocktail (1:500) and broken by sonification (10 cycles of 30 s). After a short low speed centrifugation to separate unbroken cells (4500 x *g*, 8 min) and a short high speed centrifugation to pellet the outer membrane (120000 x *g*, 5 min) the inner membranes from the supernatant were collected by centrifugation at 120000 x *g* for 1.5 h (all at 4 °C). The pellets were resuspended in buffer (20% glycerol, 10 mM Hepes/Tris pH 7.4, 0.5 mM TCEP, frozen in liquid nitrogen and stored at -80 °C. YdgR membranes were solubilized (60 min, 4 °C) in 20 mM Tris-HCl pH 8, 300 mM NaCl, 1% DDM, 10% glycerol, 0.01% NaN_3 at protein concentrations of 1-3 mg/ml. After centrifugation (100000 x *g*, 45 min) the supernatant was diluted two-fold with 20 mM Tris-HCl pH 8, 300 mM NaCl, 0.04% DDM, 2 mM histidine, 10% glycerol, 0.01% NaN_3 (wash buffer) and bound to Ni-NTA Superflow beads (2 h, 4 °C; Qiagen). The beads were then loaded onto a spin column (Promega), washed with wash buffer and eluted with buffer containing 200 mM histidine.

Reconstitution into proteoliposomes - For functional reconstitution YdgR containing membranes were solubilized in buffer (10 mM Hepes/Tris pH 7.4, 150 mM NaCl, 30 mM imidazole, 1% DDM, 5% glycerol, 0.1 mM TCEP) at a protein concentration of 1 mg/ml for 30 min on ice. After centrifugation (40000 x *g*, 30 min) the supernatant was loaded onto a Ni-NTA column (HisTrap FF, Amersham) and washed with running buffer (10 mM Hepes/Tris pH 7.4, 150 mM NaCl, 30 mM imidazole, 0.06% DDM, 5% glycerol, 0.1 mM TCEP). Protein was eluted with a gradient from 30 to 250 mM imidazole in running buffer. YdgR elutes in a sharp peak at about 150 mM imidazole. For reconstitution, *E. coli* lipids (500 µl, Avanti Polar Lipids) at a concentration of 20 mg/ml in $CHCl_3$ were dried under vacuum and resuspended in 1 ml buffer (25 mM Hepes/Na pH 7.4, 150 mM NaCl). Liposomes were destabilized with 10 mg DDM and sonicated for 30 min. 400 µl purified YdgR protein as eluted from the Ni-NTA column was added at a concentration of 250µg/ml. After 10 min incubation on ice, detergent was removed by adding 500 mg Bio-Beads SM-2 Adsorbants (BioRad) overnight at 4 °C. The detergent removal step was repeated with 200 mg Bio-Beads for 4 h at 4 °C.

Electrical measurements of reconstituted YdgR with the SURFE^{2}R^{one} setup - Electrical measurements of YdgR transport were based on the solid-supported membrane technology, which allows detection of capacitively coupled currents (19). YdgR-loaded sensors were prepared as described by Zuber *et al*. (20) using YdgR proteoliposomes (see above) and SURFE^{2}R^{one} gold electrodes from IonGate BioSciences GmbH. The measurements were performed with the commercially available surface electrogenic event reader SURFE^{2}R^{one} setup (IonGate Biosciences). The measurements of transporter-related currents on the chip are based on the shift of electrical charges as the transporters go through the transport cycle and the shift can originate from the movement of charged substrates or of protein moieties carrying (partial) charges. YdgR-mediated transport was activated via rapid solution exchange from a so-called non-activating (off) (40 mM KCl, 50 mM Hepes, 50 mM MES, 2 mM $MgCl_2$, pH 6.7) to an "activating" (on) solution (containing 30 mM glycine and 20 mM glycyl-

glycine). In "activating" solutions with lower glycyl-glycine concentrations, the total osmolarity of the organic solutes glycine and glycyl-glycine together was adjusted to 50 mM. After a rapid fluid exchange to a peptide-containing solution the charging of the proteoliposomes on the sensor driven by the H^+/peptide symport is measured. Previous comparisons of the characteristics of rheogenic transporters employing this new cell-free electrophysiological technique with findings from patch clamp studies revealed a very good correlation in all features (21).

Blue native gel electrophoresis - Linear 5-12% gradient gels for BN-PAGE were prepared and run as previously described by Schägger and von Jagow (22). Thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), lactate dehydrogenase (140 kDa), and bovine serum albumin (66 kDa) were used as standard proteins.

Transmission electron microscopy - DDM-solubilized YdgR protein as eluted from the Ni-NTA column was adsorbed for 10 s to parlodion carbon-coated copper grids rendered hydrophilic by glow discharge at low pressure in air. Grids were washed with four drops of double-distilled water and stained with 2 drops of 0.75% uranyl formate. Images were recorded on Eastman Kodak Co. SO-163 sheet films with a Hitachi H-7000 electron microscope operated at 100 kV.

Data analysis - All experiments were performed for the indicated number of observations (*n*). IC_{50} values were obtained by nonlinear regression and the value is given ± standard error. The K_i were calculated from the IC_{50} using the equation from Cheng and Prusoff (23).

6.5 RESULTS

Cloning and functional expression of the YdgR protein from E. coli - The *ydgR* gene from *E. coli* was amplified from genomic DNA by PCR using gene specific primers. The gene

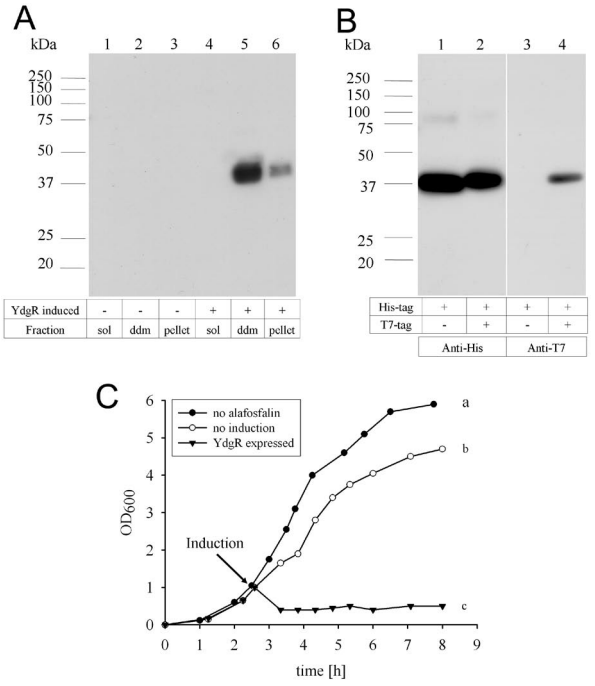

FIGURE 1. Functional expression of YdgR. (A) Western Blot analysis of YdgR protein expressed in *E. coli* BL21(DE3)pLysS with pET-21-rbs-YdgR-His. Soluble and membrane fractions of uninduced (*lane 1-3*) and IPTG-induced (*lane 4-6*) *E. coli* cells were separated on a 12.5% SDS-polyacrylamide gel, subsequently transferred on PVDF membrane and probed with an anti-His antibody (Novagen). Each lane contains protein from an equivalent of 40 µl bacterial culture: soluble proteins (*lanes 1 and 4*), DDM solubilized membrane proteins (*lanes 2 and 5*) and non-solubilized proteins (*lanes 3 and 6*). (B) Western Blot analysis of YdgR protein expressed with a C-terminal hexahistidine-tag compared with YdgR exhibiting a C-terminal hexahistidine-tag and an N-terminal T7-tag. Membrane proteins of IPTG-induced *E. coli* cells were separated on a 12.5% SDS-polyacrylamide gel, subsequently transferred on PVDF membrane and probed with an anti-His antibody (Novagen, *lanes 1* and *2*) or an anti-T7 antibody (Novagen, *lanes 3* and *4*). Each lane contains protein from an equivalent of 100 µl bacterial culture. (C) Growth curve in the presence of 200 µg/ml alafosfalin. *E. coli* BL21 transformed with pET-21-rbs-YdgR-His were grown in the absence (*a*) and presence (*b* and *c*) of alafosfalin. At OD_{600}= 1 expression of YdgR was induced by 0.1 mM IPTG (*a* and *c*).

was cloned into the pET-21 vector fused with a C-terminal hexahistidine tag (pET-21-ydgR-His). The expression of the YdgR protein in *E. coli* BL21(DE3)pLysS was tested by Western blot analysis with an antibody derived against the hexahistine tag (Figure 1A). Proteins from non-induced (control, Figure 1A, *lanes 1–3*) and induced *E. coli* (Figure 1A, *lanes 4-6*) were fractionated in (i) soluble proteins, (ii) membrane proteins solubilized in 1% DDM and (iii) insoluble proteins, and separated by SDS-PAGE. The anti-His antibody detected a protein in the

membrane protein fraction of induced cells (*lane 5*) with an apparent molecular weight of 39 kDa. Small amounts of the protein were detected in the insoluble protein fraction (*lane 6*) indicating some protein not properly solubilized or from inclusion bodies. Since no band was detected from extracts obtained from control cells, the 39 kDa protein band represents the YdgR protein. The discrepancy between the expected molecular weight (55 kDa) as deduced from the amino acid sequence of YdgR and the apparent molecular weight as determined by SDS-PAGE might be due to the known abnormal migration behaviour of several membrane proteins. This often causes an underestimation of the true molecular weight. To test this hypothesis and to verify that the protein is not degraded we constructed a second expression vector with an N-terminal T7-tag (eleven amino terminal amino acids of the T7 gene 10 plus ten amino acids spacer) in addition to the C-terminal hexahistidine tag (pET-21b-rbs-T7-ydgR-His). The apparent molecular weight of the two proteins "YdgR-His" and "T7-YdgR-His" were compared by Western blot analysis with the anti-His antibody (Figure 1B, *lanes 1 and 2*) and the anti-T7 antibody (Figure 1B, *lanes 3 and 4*). As expected, both proteins were detected by the anti-His antibody (Figure 1B, *lanes 1 and 2*), while the anti-T7 antibody detected only "T7-YdgR-His" (Figure 1B, *lane 4*). Since both antibodies detected the same protein ("T7-YdgR-His") and the apparent molecular weight of this protein is only slightly icreased compared to "YdgR-His" we conclude that the overall deduced low apparent molecular weight is due to an abnormal migration behaviour, but not proteolysis. All further experiments were performed with the pET-21-rbs-ydgR-His construct.

To assess the functionality of the expressed YdgR protein, simple growth experiments were conducted (Figure 1C). In *Salmonella typhimurium* the toxic phosphonopeptide alafosfalin is taken up by the tripeptide permease TppB and mutants in TppB show resistance to alafosfalin (16). We therefore grew *E. coli* carrying the vector (pET-21-rbs-YdgR-His) in the presence of a non-toxic

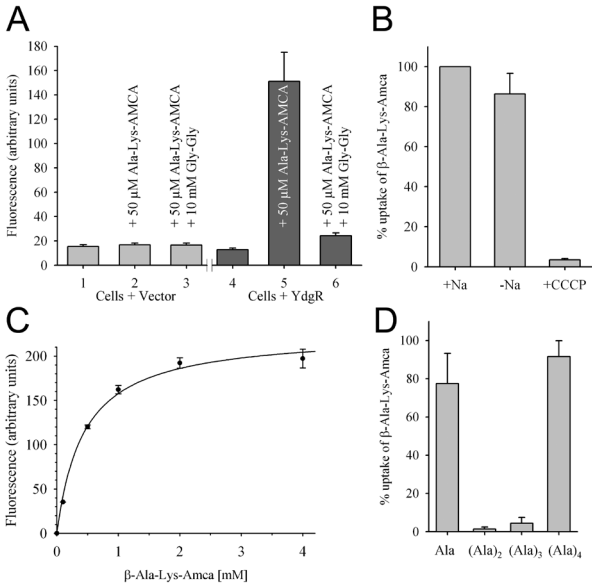

FIGURE 2. Transport function of the YdgR protein with the fluorescent dipeptide β-Ala-Lys-AMCA serving as a substrate. (A) Control cells transformed with pET-21 vector (bars 1-3) and cells expressing the YdgR protein (bars 4-6) were incubated in Krebs-Buffer alone (*bars 1* and *4*) or Krebs-Buffer containing 50 µM β-Ala-Lys-AMCA (*bars 2* and *5*) or 50 µM β-Ala-Lys-AMCA together with 10 mM of the competitor Gly-Gln (*bars 3* and *6*) (*n=6*). (B) The Na^+ and H^+ dependence of β-Ala-Lys-AMCA uptake by YdgR was assessed by incubating *E. coli* expressing YdgR with 50 µM β-Ala-Lys-AMCA in the presence (*left-most*) and in the absence (*center bar*) of Na^+ (by replacing sodium with choline) or by exposing cells (*n=4*) to 10 µM of the selective proton-ionophore CCCP (*right-most bar*). (C) Uptake of β-Ala-Lys-AMCA as a function of substrate concentrations showed saturation kinetics with an apparent K_t of 0.44 ± 0.05 mM. (D) Inhibition of β-Ala-Lys-AMCA uptake by the amino acid L-Ala and the corresponding di-, tri- and tetrapeptides of L-Ala. Uptake of β-Ala-Lys-AMCA (50 µM, final concentration) was determined in the presence of 10 mM L-Ala (bar 1), L-Ala-L-Ala (*bar 2*), L-Ala-L-Ala-L-Ala (*bar 3*), L-Ala-L-Ala-L-Ala-L-Ala (*bar 4*) (*n=6*).

concentration of alafosfalin (200 µg/ml, *curve b*). As compared to E. coli grown in the absence of alafosfalin (*curve a*) growth rates were only slightly reduced. When expression of YdgR was induced at an A_{600} of ~1 (*curve c*) the peptide now caused cell death. This indicates that YdgR was functionally expressed in the membrane, loading the cells with the toxic agent. Transport function of YdgR was also determined using the fluorescent dipeptide reporter β-Ala-Lys-AMCA as a substrate (Figure 2A). Cells expressing YdgR showed only minor autofluorescence in the absence of β-Ala-Lys-AMCA (*bar 4*), but fluorescence increased significantly after incubation of cells with 50 µM β-Ala-Lys-AMCA (*bar 5*). Fluorescence was abolished by addition of an excess amount of the

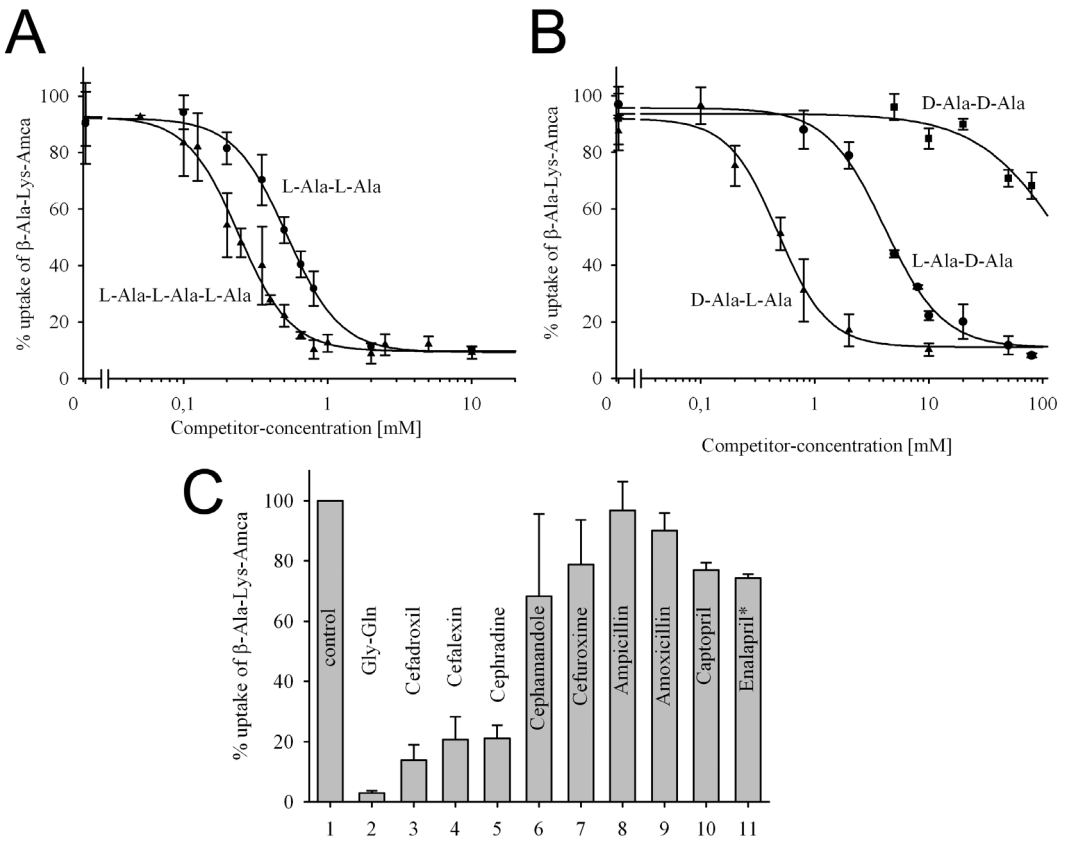

FIGURE 3. Substrate specificity of the YdgR protein. (A) Inhibition of β-Ala-Lys-AMCA uptake by the di- and tripeptides of L-Ala. Uptake of β-Ala-Lys-AMCA (50 μM) was determined in the presence of increasing concentrations (0.05 - 10 mM) of either L-Ala-L-Ala (circles) with an IC_{50} of 0.52 ± 0.03 mM, or L-Ala-L-Ala-L-Ala (triangles) with an IC_{50} of 0.24 ± 0.01 mM; (B) Stereoselectivity of transport was determined by inhibition of β-Ala-Lys-AMCA (50 μM) uptake by 0.1 - 10 mM of D-Ala-L-Ala (*triangles*) with an IC_{50} of 0.48 ± 0.05 mM, or 0.8 - 80 mM of L-Ala-D-Ala (*circles*) with an IC_{50} of 4.10 ± 0.33 mM, or 5 - 80 mM of D-Ala-D-Ala (*squares*). (C) Interaction of peptidomimetics with YdgR was tested by uptake (*n*=4; **n*=2) of 50 μM β-Ala-Lys-AMCA in the absence (*bar 1*) and in the presence of 10 mM of the following competitors: Gly-Gln (*bar 2*), cefadroxil (*bar 3*), cefalexin (*bar 4*), cephradine (*bar 5*), cephamandole (*bar 6*), cefuroxime (*bar 7*), ampicillin (*bar 8*), amoxicillin (*bar 9*), captopril (*bar 10*) or enalapril (*bar 11*).

non-labelled dipeptide Gly-Gln (*bar 6*). When β-Ala-Lys-AMCA uptake was determined as a function of its concentration, transport was found to be saturable with an apparent K_t of 0.44 ± 0.05 mM (Figure 2C). Control experiments with pET-21 vector transformed *E. coli* showed no uptake of β-Ala-Lys-AMCA (*bars 1-3*) indicating that no other endogenous transport system in *E. coli* mediates β-Ala-Lys-AMCA uptake. For further characterization of transport, we studied the requirements of Na^+ and H^+ as cotransport-ions (Figure 2B). The replacement of Na^+ by choline had no significant effect on the uptake of β-Ala-Lys-AMCA, but the presence of CCCP, a proton ionophore, caused complete inhibition of transport. Thus, peptide uptake via YdgR depends on the proton-motive force and is most likely as described for the other family members mediated by a coupled proton-substrate cotransport.

Studies on substrate specificity of YdgR - To determine the transporter's substrate specificity, *E. coli* cells overexpressing YdgR were incubated with 50 μM β-Ala-Lys-AMCA in the presence of an excess (200-fold) of L-alanine either as free amino acid or as a di-, tri- and tetrapeptides (Figure 2D). Alanine and tetra-alanine did not inhibit uptake of β-Ala-Lys-AMCA, while di- and tri-alanine completely inhibited β-Ala-Lys-AMCA uptake. Inhibition of transport by increasing concentrations of di- and tri-alanine allowed apparent affinities (IC_{50} values) of 0.52 ± 0.03 mM and 0.24 ± 0.01 mM to be determined (Figure 3A). Stereospecificity of YdgR-mediated flux was assessed by determining IC_{50} values of dipeptides carrying L- or D-alanine residues

(Figure 3B). Substitution of the N-terminal L-Ala for the D-isomer did not alter affinity significantly, but when the C-terminal L-Ala was replaced by a D-isomer affinity of the competitor was reduced about 8-fold (Figure 3B). D-Ala-D-Ala did not show detectable affinity for interaction with YdgR (Figure 3B).

Transport characteristics of differently charged dipeptide substrates - To further characterise substrate specificity of YdgR, we determined IC_{50} values of various other compounds ranging from differently charged dipeptides consisting of L-amino acids, to ω-amino-fatty-acid and a known inhibitor of mammalian peptide transporters (see Table I). Neutral dipeptides represented by Gly-Gln or cationic peptides such as Lys-Gly showed relatively high affinities with IC_{50} values of 0.51 ± 0.06 mM and 0.43 ± 0.02 mM, respectively. Introducing the positively charged amino acid at the C-terminal position (Gly-Lys) resulted in a 7-fold reduction of affinity represented by an IC_{50} of 2.83 ± 0.27 mM as compared to 0.43 mM for Lys-Gly. This indicates an asymmetric substrate binding site in YdgR similar to that described for the mammalian peptide transporters (24). This observation is strengthened by experiments with the anionic dipeptide Asp-Gly. When Asp instead of Lys is placed in the first position affinity decreases by 10-fold (Asp-Gly, IC_{50} = 4.46 ± 0.45 mM) whereas when placed in the second position affinity increases again when compared to a Lys residue (Gly-Asp, IC_{50} = 1.92 ± 0.13 mM). Modifying the peptide bond nitrogen by a CH_3 group like in glycyl-sarcosine, yielded a moderate affinity (IC_{50} = 1.16 ± 0.09 mM). The highest affinity of all test compounds displayed the inhibitor of mammalian peptide transporters Lys-Z-Nitro-Pro with an IC_{50} value of 0.033 ± 0.006 mM. Mammalian peptide transporters do not require a peptide bond for recognition of a substrate (25). To test whether this holds true also for the YdgR protein, we used 5-aminolevulinic acid as a substrate which carries only the two oppositely charged head groups separated by four carbon units and a backbone carbonyl. Its apparent affinity was 1.69 ± 0.14 mM and therefore higher

TABLE 1
Summary of IC_{50} values (mM) and K_i values

Numbers in parentheses indicate the number of independent determinations. K_i values were determined by the equation $K_i = IC_{50}/(1 + [substrate]/K_D)$ (Cheng and Prusoff (23)). All amino acids are L isoforms unless otherwise indicated.

	IC_{50}	K_i
Gly-Gln	0.51 ± 0.06 (7)	0.46 ± 0.05
Lys-Gly	0.43 ± 0.02 (4)	0.39 ± 0.02
Gly-Lys	2.83 ± 0.27 (5)	2.54 ± 0.24
Asp-Gly	4.46 ± 0.45 (6)	4.00 ± 0.40
Gly-Asp	1.92 ± 0.13 (12)	1.72 ± 0.12
Lys-Ala	0.77 ± 0.05 (4)	0.69 ± 0.04
Leu-Ala	1.27 ± 0.10 (4)	1.14 ± 0.09
Ala-Ala	0.52 ± 0.03 (6)	0.47 ± 0.03
Ala-D-Ala	4.10 ± 0.33 (4)	3.68 ± 0.30
D-Ala-Ala	0.48 ± 0.05 (4)	0.43 ± 0.04
Ala-Ala-Ala	0.24 ± 0.01 (8)	0.22 ± 0.01
Gly-Sar	1.16 ± 0.09 (10)	1.04 ± 0.08
Alafosfalin	0.28 ± 0.03 (4)	0.25 ± 0.03
Lys-Z-Nitro-Pro	0.033 ± 0.006 (4)	0.03 ± 0.01
5-Aminolevulinic acid	1.69 ± 0.14 (6)	1.52 ± 0.13

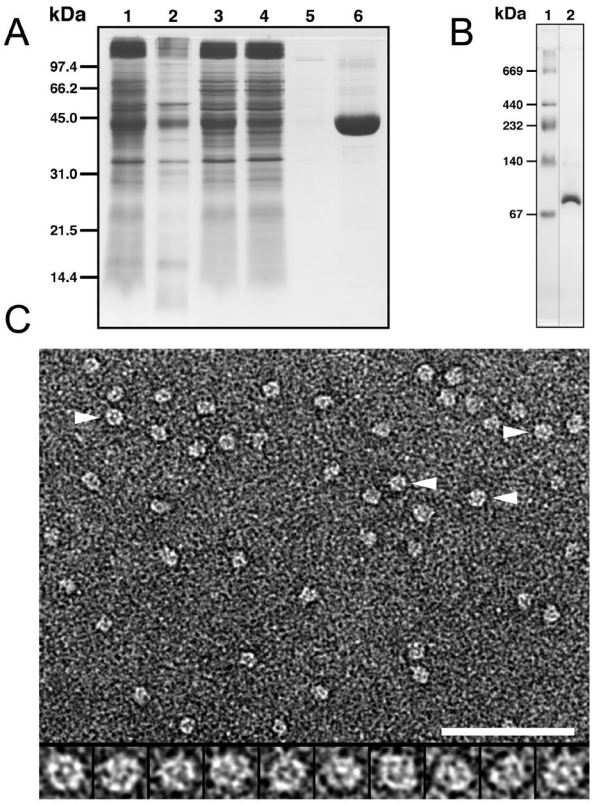

FIGURE 4. Purification, BN-PAGE and TEM of the YdgR protein. (A) Purification of YdgR. SDS-PAGE of the different purification steps: DDM-solubilized membranes (lane 1), pellet (lane 2) and supernatant (lane 3) after ultracentrifugation of the DDM-solubilized membranes, Ni^{2+}-NTA flowthrough (lane 4), Ni^{2+}-NTA wash (lane 5), Ni^{2+}-NTA elution (*lane 6*, 5 µg loaded). (B) BN-PAGE of purified YdgR. Protein standard (*lane 1*) and purified YdgR protein (*lane 2*, 9 µg loaded). The YdgR protein migrates as a ~85 kDa band in the 5-12% linear gradient gel. (C) TEM of negatively stained purified YdgR proteins. The homogeneity of the purified YdgR proteins is reflected in the electron micrograph. The selected particles marked by *arrowheads* were magnified and are displayed in the *gallery*. The *scale bar* represents 60 nm and the frame size of the magnified particles in the gallery is 10.8 nm.

than that of most charged dipeptides.

Transport of peptidomimetics - Beside di- and tripeptides mammalian peptide transporters accept a broad spectrum of peptidomimetics like β-lactam antibiotics and ACE-inhibitors as substrates. We have tested selected peptidomimetics as known substrates of mammalian peptide transporters for analysis of their interaction with the YdgR protein in competition assays (Figure 3C). Compared to the dipeptide Gly-Gln (*bar 2*), which inhibited β-Ala-Lys-AMCA uptake by 97%, similar inhibition rates of 86%, 79% and 79%, respectively were observed for the three aminocephalosporins cefadroxil (*bar 3*), cefalexin (*bar 4*) and cephradine (*bar 5*) at 10 mM concentration. Cefuroxime (bar 6) and cefamandole (*bar 7*) showed markedly reduced affinities with modest inhibition of only 21% and 32%, respectively, which results from the lack of an α-amino group important for high affinity. Whereas the two aminopenicillins ampicillin (*bar 8*) and amoxicillin (*bar 9*, 5 mM concentration) seemed not to serve as substrates, the ACE-inhibitors captopril (*bar 10*) and enalapril (*bar 11*) also showed only low affinity type inhibition of β-Ala-Lys-AMCA uptake.

Purification and functional reconstitution of YdgR - Inner membranes of *E. coli* were solubilized with DDM and YdgR was purified by nickel affinity chromatography. The SDS polyacrylamide gel displayed in Figure 4A summarizes the different purification steps. The purified YdgR protein migrated as a single band with an apparent mass of ~39 kDa that is identical to that observed by Western blot analysis for the non-purified protein (see Figure 1A). Typically 2–4 mg of YdgR protein could be purified from 1 liter of bacterial culture. The functionality of purified YdgR was assessed after its reconstitution into liposomes using the commercially available SURFE^{2}R^{one} setup (IonGate Biosciences). The SURFE^{2}R^{one} setup is based on the solid-supported membrane technology and allows detection of capacity-coupled currents induced by movement of charged molecules across lipid bilayers (19,20). To measure YdgR-specific transport, proteoliposomes containing YdgR were adsorbed to SURFE^{2}R^{one} sensors and transport was activated by rapid exchange of a "non-activating" solution containing glycine against an "activating solution" containing the dipeptide glycyl-glycine (Gly-Gly). Figure 5A demonstrates that Gly-Gly induced significant currents in YdgR-proteoliposomes with increasing currents at increasing substrate concentrations (1 to 20 mM). No Gly-Gly-induced currents were observed when YdgR-free liposomes were loaded onto the sensor (Figure 5B *trace d*). Additon of 5 mM Gly-Gly in both, the "non-activating" and the "activating solutions", partially inhibited the YdgR-response to the 20 mM Gly-Gly concentration jump (Figure 5B, *trace b*), whereas 10 mM Gly in both solutions did not have any significant effect on the YdgR response (Figure 5B, *trace c*). This demonstrates that only dipeptide Gly-Gly but not the free glycine causes the currents. Since the transport studies were performed with Na^+-free buffers of pH 6.2 the observed currents must originate from proton movement coupled to dipeptide translocation in a symport mechanism as demonstrated for the mammalian peptide transporters PEPT1 and PEPT2.

Characterization of the purified YdgR protein by BN-PAGE and TEM - To determine whether YdgR exists in a monomeric or oligomeric state, the purified protein was subjected to BN-PAGE. The results are summarized in Figure 4B and provide a strong protein band at an apparent molecular weight (M_{obs}) of ~85 kDa. In addition, DDM-solubilized YdgR protein was negatively stained and examined by TEM. The homogeneity of the purified YdgR is documented in Figure 4C. Single YdgR proteins are distinguished and display a crown-like structure with a typical diameter of ~8 nm. In addition, a plug-like density was discerned in the centre of the protein (see gallery of well preserved YdgR top views in Figure 4C, *lower panel*).

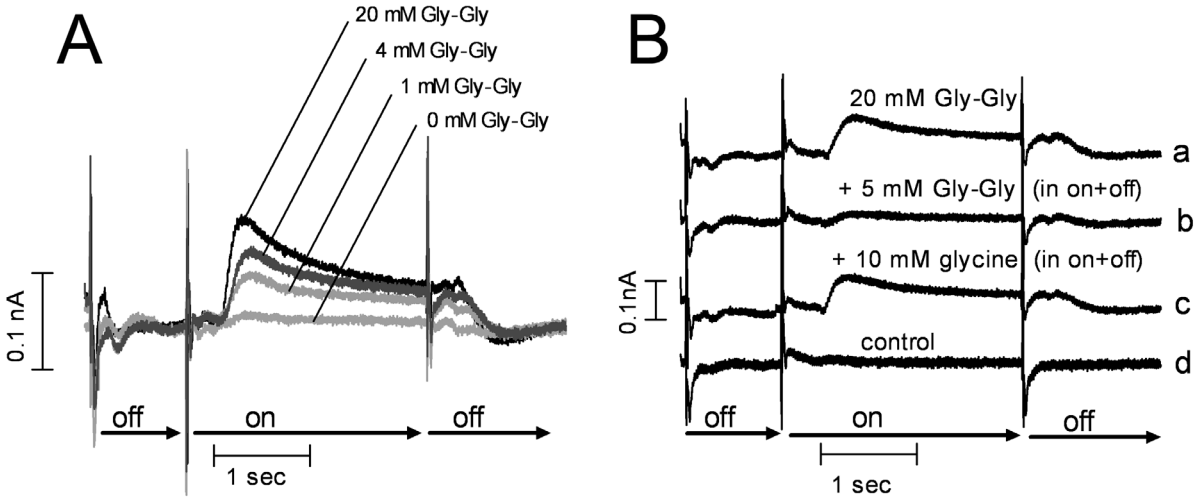

FIGURE 5. Functional reconstitution of YdgR. (A) Electrical response of YdgR reconstituted into liposomes to a dipeptide solution. YdgR-containing proteoliposomes were adsorbed to SURFE^{2}R^{one} sensors and perfused with Na^+-free "activating" solution (*on*) containing different concentrations of the dipeptide Gly-Gly (*n=3*). (B) The response of YdgR-containing liposomes is inhibited by Gly-Gly but not by free glycine. For traces *a-d*, the "activating solution" (*on*) contained 20 mM Gly-Gly. The electrical response of the YdgR-containing liposomes on the sensor to a 20 mM Gly-Gly concentration jump is shown in *trace a*. Addition of 5 mM Gly-Gly to the "non-activating" (*off*) and "activating" (*on*) solutions blunts the response to 20 mM Gly-Gly (*trace b*) whereas the addition of 10 mM glycine to the "non-activating" and "activating" solutions does not change the YgdR-response (*trace c*). Trace d shows the recording from a sensor loaded with YdgR-free liposomes serving as a negative control.

6.6 DISCUSSION

The genome of *E. coli* contains four yet uncharacterized members of the family of proton-dependent oligopeptide transporter (POT-family) named *ydgR*, *ybgH*, *yhiP* and *yjdL* as identified by sequence analysis. Because of lack of functional data they are still classified as hypothetical proteins. We have cloned the *ydgR* gene from genomic DNA and overexpressed the protein in *E. coli* BL21 cells. Coomassie stained SDS-PAGE and Western blot analysis identified the YdgR protein, and uptake experiments with the fluorescent dipeptide β-Ala-Lys-AMCA in bacterial cells demonstrated its function as a dipeptide transporter with features similar to mammalian peptide transporters. Moreover, employing the SURFE^{2}R^{one} sensor technology we demonstrate for the first time rheogenic transport by a prokaryotic peptide transporter.

The bacterial peptide transport system tppB (tripeptide permease) was genetically characterised in mutants of *Salmonella typhimurium*. Deficiency in the locus for tppB conferred resistance to the toxic phosphonopeptide alafosfalin and this was classified as a characteristic tppB activity (16). By locus analysis, the *ydgR* gene of *E. coli* was identified as a similar tripeptide permease (tppB) (14) but had not been functionally characterised. By growth experiments with the toxic phosphonopeptide alafosfalin we could experimentally confirm that YdgR represents the *E. coli* homolog of tppB. However, our functional data on substrate specificity of YdgR reveal that not only tripeptides but also dipeptides serve as substrates. Moreover, transport was proven to be Na^+-independent and completely abolished in the presence of the proton ionophore CCCP suggesting a proton-coupling of substrate transport in analogy to other members of the family such as PEPT1 or PEPT2. These two proteins operate as electrogenic proton-coupled symporters with a variable flux-coupling stoichiometry for proton to substrate cotransport and the main driving force being provided by membrane voltage. Via the SURFE^{2}R^{one} measurements of YdgR reconstituted into proteoliposomes we could experimentally verify that dipeptide transport by YdgR is also of electrogenic nature and occurs in the absence

of Na^+ ions, most likely in analogy to PEPT1 and PEPT2 as proton-peptide symport.

The substrate recognition pattern of YdgR shows a remarkable and unexpected similarity to the mammalian PEPT1 in various aspects. All known substrates of PEPT1 that were tested with YdgR also interact with the substrate binding site of the bacterial protein with very similar affinities. Moreover, the observed stereoselectivity of transport and differences in affinities of charged peptides with identical side chains but in different spatial position (N- *versus* C-terminal) are also characteristic for PEPT1. Finally, even peptidomimetics such as the aminocephalosporins interact with the YdgR substrate binding site with similar affinities as determined for PEPT1. YgdR represents in its substrate recognition pattern in all aspects a mammalian PEPT1-phenotype. This suggests that YdgR possesses a similar architecture in the substrate binding domain.

There is a moderate degree of sequence homology between YdgR and mammalian PEPT´s. We have projected the YdgR amino acid sequence onto the proposed topology model of PEPT1 (Figure 6) on the basis of a ClustalW analysis (26). Highest sequence identity can be found in the first half of the protein especially in the transmembrane regions (TM) 1-6 and the first extracellular loop. A modest sequence identity is also found in TM10–12 and as suggested by functional analysis of mammalian PEPTs based on chimeric proteins and by side directed mutagenesis; these regions are involved in substrate binding and transport (8-12,27). Some of the amino acid residues in these regions were identified as essential for transport by PEPT1 and are well conserved in the YdgR protein including Trp^{295} in TM7 and Glu^{595} in TM10. The amino acid His^{121} in TM4 is not conserved, but there is an alternative His residue nearby in the primary structure of YdgR. Surprisingly the His^{57}

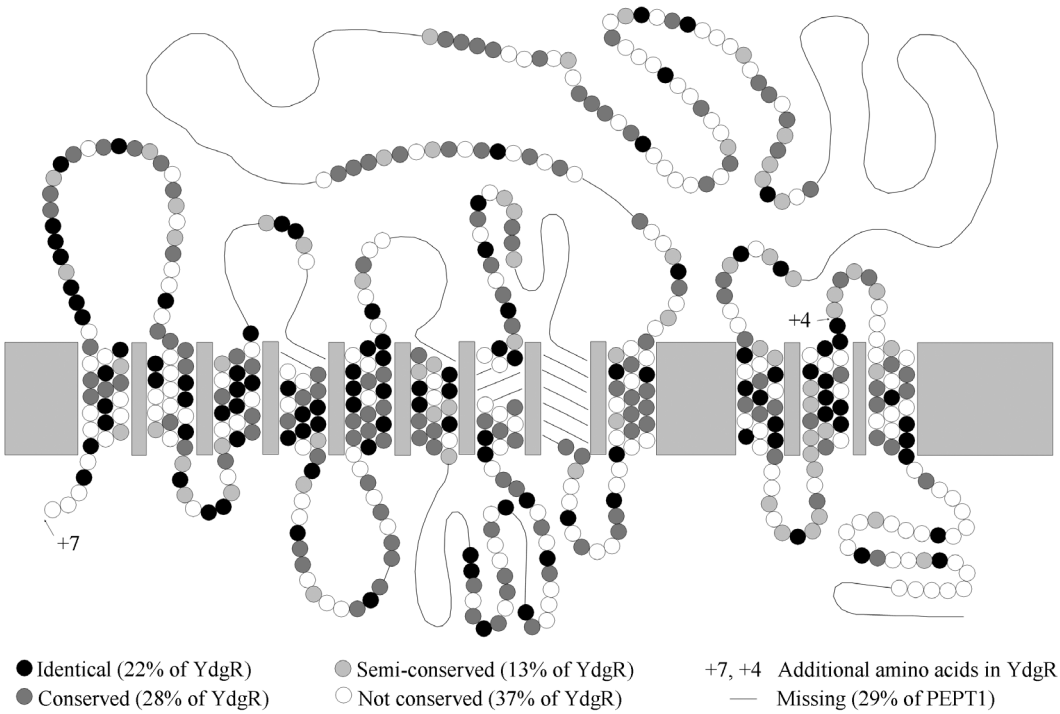

FIGURE 6. Homology model of YdgR to human PEPT1. The proposed topology model of human PEPT1 was modified to show the homology to YdgR calculated by Clustal W. From the 500 amino acids of YdgR 22% are identical in PEPT1, 28% are conserved and 13% are semi-conserved. 29% of PEPT1 is missing in YdgR.

residue that is described as essential for proper function in the mammalian peptide transporters is not conserved in YdgR and is replaced by a serine residue. The lowest homology found between the mammalian proteins and YdgR is in transmembrane regions 7–9 together with the loops in between. These regions are thought to contribute to the different kinetic phenotypes of PEPT1 and PEPT2 (28). Most strikingly, YdgR does not possess the large extracellular loop between TM 9 and TM 10 found in PEPT1 and PEPT2 which suggests that the loop domain is not important at all for the transport process.

Nothing is known about the 3D structure of any peptide transporter protein of the PTR-family. Protein expression is low in native cells and tissues, and heterologous expression of mammalian PEPT1 in *Pichia pastoris* did not yield enough protein for structural analysis (29). The discovery of a bacterial homologue with essentially identical transport characteristics to the mammalian proteins that can be expressed and purified with high yield represents therefore an important step towards structural analysis of the protein and insights into the transport mechanism. We were able to obtain 2–4 mg of pure and stable YdgR membrane protein from 1 liter of bacterial culture. BN-PAGE of purified YdgR displayed a strong band with a M_{obs} of ~85 kDa (Figure 4B). This M_{obs} is composed of the molecular mass of YdgR (~55 kDa) and the mass of the DDM-Coomassie brilliant blue G-250 micelle and lipids attached to the protein. Thus the M_{obs} indicates that YdgR may exist as a monomer. The use of the conversion factor determined by Heuberger *et al*. (30) to estimate the mass of membrane proteins from the M_{obs} also supports the monomeric nature of YdgR. Chemical crosslinking experiments performed with the amino-specific reagents disuccinimidyl suberate and bis-sulfosuccinimidyl-suberate did not yield specific crosslinking products (data not shown) in line with the results from BN-PAGE.

TEM of negatively stained DDM-solubilized YdgR proteins revealed a crown-like structure with a central density. The measured diameter of the YdgR particles was ~8 nm. Assuming a boundary layer of ~ 1.5 nm (31) for DDM attached to the hydrophobic part of the protein the resulting diameter is similar to that of the red permease monomer when embedded into the lipid bilayer; Red permease is a lactose permease fusion protein which forms monomers and trimers. It consists of twelve transmembrane helices and has a similar molecular weight to YgdR (32).

In summary we have cloned, overexpressed and characterized biochemically, functionally and structurally the prokaryotic proton-dependent peptide transporter YdgR. The substrate recognition pattern of YdgR shows a remarkable similarity to the mammalian PEPT1 protein and therefore YdgR could serve as a paradigm for the mammalian peptide transporters. To understand the biophysical principles underlying transmembrane peptide transport, a high-resolution structure is indispensable. Our present work with YdgR sets the basis for further structural analysis using 2D and 3D crystals, and electron and X-ray crystallography studies of the first peptide transporter protein. Obtaining a high-resolution structure of YdgR is crucial for any exploitation of the architecture of this class of transport proteins that are unique amongst the solute carriers with respect to their ability to transport literally thousands of substrate differing in size, polarity and charge.

6.7 ACKNOWLEDGEMENTS

We would like to thank Mrs Daniela Kolmeder for excellent technical assistance. This work was supported by the Sixth Framework Programme of the European Union (EUGINDAT, European Genomics Initiative on Disorders of Plasma Membrane Amino Acid Transporters).

6.8 REFERENCES

1. Brandsch M., Knutter I. and Leibach FH. (2004) *Eur J Pharm Sci* **21** 53-60
2. Daniel H. (2004) *Annu Rev Physiol* **66** 361-384
3. Daniel H. and Kottra G. (2004) *Pflugers Arch* **447** 610-618
4. Terada T. and Inui K. (2004) *Curr Drug Metab* **5** 85-94
5. Ganapathy ME., Huang W., Wang H., Ganapathy V. and Leibach FH. (1998) *Biochem Biophys Res Commun* **246** 470-475
6. Tamai I., Nakanishi T., Nakahara H., Sai Y., Ganapathy V., Leibach FH. and Tsuji A. (1998) *J Pharm Sci* **87** 1542-1546
7. Covitz KM., Amidon GL. and Sadee W. (1998) *Biochemistry* **37** 15214-15221
8. Bolger MB., Haworth IS., Yeung AK., Ann D., von Grafenstein H., Hamm-Alvarez S., Okamoto CT., Kim KJ., Basu SK., Wu S. and Lee VH. (1998) *J Pharm Sci* **87** 1286-1291
9. Chen XZ., Steel A. and Hediger MA. (2000) *Biochem Biophys Res Commun* **272** 726-730
10. Fei YJ., Liu W., Prasad PD., Kekuda R., Oblak TG., Ganapathy V. and Leibach FH. (1997) *Biochemistry* **36** 452-460
11. Terada T., Saito H., Mukai M. and Inui KI. (1996) *FEBS Lett* **394** 196-200
12. Yeung AK., Basu SK., Wu SK., Chu C., Okamoto CT., Hamm-Alvarez SF., von Grafenstein H., Shen WC., Kim KJ., Bolger MB., Haworth IS., Ann DK. and Lee VH. (1998) *Biochem Biophys Res Commun* **250** 103-107
13. Steiner HY., Naider F. and Becker JM (1995) *Mol Microbiol* **16** 825-834
14. Goh EB., Siino DF. and Igo MM. (2004) *J Bacteriol* **186** 4019-4024
15. Higgins CF. and Gibson MM. (1986) *Methods Enzymol* **125** 365-377
16. Gibson MM., Price M. and Higgins CF. (1984) *J Bacteriol* **160** 122-130
17. Dieck ST., Heuer H., Ehrchen J., Otto C. and Bauer K. (1999) *Glia* **25** 10-20
18. Otto C., tom Dieck S. and Bauer K. (1996) *Am J Physiol* **271** 210-217
19. Meyer-Lipp K., Ganea C., Pourcher T., Leblanc G. and Fendler K. (2004) *Biochemistry* **43** 12606-12613
20. Zuber D., Krause R., Venturi M., Padan E., Bamberg E. and Fendler K. (2005) *Biochim Biophys Acta* **1709** 240-250
21. Geibel S., Flores-Herr N., Licher T. and Vollert H. (2006) *J Biomol Screen* **11** 262-268
22. Schagger H. and von Jagow G. (1991) *Anal Biochem* **199** 223-231
23. Cheng Y. and Prusoff WH. (1973) *Biochem Pharmacol* **22** 3099-3108
24. Theis S., Hartrodt B., Kottra G., Neubert K. and Daniel H. (2002) *Mol Pharmacol* **61** 214-221
25. Doring F., Will J., Amasheh S., Clauss W., Ahlbrecht H. and Daniel H. (1998) *J Biol Chem* **273** 23211-23218
26. Chenna R., Sugawara H., Koike T., Lopez R., Gibson TJ., Higgins DG. and Thompson JD. (2003) *Nucleic Acids Res* **31** 3497-3500
27. Doring F., Dorn D., Bachfischer U., Amasheh S., Herget M. and Daniel H. (1996) *J Physiol* **497** 773-779
28. Fei YJ., Liu JC., Fujita T., Liang R., Ganapathy V. and Leibach FH. (1998) *Biochem Biophys Res Commun* **246** 39-44
29. Theis S., Doring F. and Daniel H. (2001) *Protein Expr Purif* **22** 436-442
30. Heuberger EH., Veenhoff LM., Duurkens RH., Friesen RH. and Poolman B. (2002) *J Mol Biol* **317** 591-600
31. Dekker JP., Boekema EJ., Witt HT. and Rogner M. (1988) *Biochimica et Biophysica Acta (BBA)* - **936** 307-318
32. Zhuang J., Prive GG., Werner GE., Ringler P., Kaback HR. and Engel A. (1999) *J Struct Biol* **125** 63-75

7. DtpB (YhiP) and DtpA (TppB, YdgR) are Prototypical Proton-Dependent Peptide Transporters of E. coli

Daniel Harder[1], Jürgen Stolz[1], Fabio Casagrande[2],
Petr Obrdlik[3], Dietmar Weitz[1], Dimitrios Fotiadis[2] and Hannelore Daniel[1]

[1] Molecular Nutrition Unit, Department of Food and Nutrition, Technical University of Munich, 85350 Freising, Germany

[2] M.E. Müller Institute for Structural Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland

[3] Iongate Biosciences GmbH, D528 Industriepark Höchst, D-65926 Frankfurt, Germany

7.1 ABBREVIATIONS

IPTG, isopropyl 1-thio-β-D-galactopyranoside; **CCCP**, carbonyl cyanide *p*-chlorophenylhydrazone; **TCEP**, Tris-2-carboxyethyl-phosphine; **DDM**, n-dodecyl-β-D-maltoside; **TEM**, transmission electron microscopy; **TM**, transmembrane regions; **Ni-NTA**, nickel-nitrilotriacetic acide; **rbs**, ribosomal binding site; **TBS**, Tris-buffered saline; **AMCA**, N_{ε}-7-amino-4-methyl-coumarin-3-acetic acid; **MES**, 4-morpholineethanesulfonic acid; $\mathbf{M_{obs}}$, apparent molecular mass.

7.2 ABSTRACT

The genome of *E. coli* contains four genes assigned to the PTR (peptide transporter) family. Only one gene was identified as a tripeptide permease (TppB, YdgR), whereas the functions of the other three genes (*yhiP*, *ybgH* and *yjdL*) remain unknown. Here we describe the over-expression of *yhiP* as a His-tagged fusion protein in *E. coli*. With glycyl-sarcosine (Gly-Sar) as a tracer substrate an apparent affinity constant of 6.5 mM for saturable influx was obtained and over-expression of the protein increased markedly the susceptibility of cells to the toxic phosphonopeptide alafosfalin. Transport activity in cells was strongly decreased in the presence of a protonophore but unaffected by sodium depletion, suggesting H^+-dependence. Competition experiments established the protein to transport di- and tripeptides with a slightly different specificity than TppB. Western blot analysis revealed YhiP to posses an apparent mass of ~40 kDa. Negative-stain transmission electron microscopy of purified YhiP revealed a donut-like shape of ~8 nm in diameter with a distinct central depression. Functional reconstitution of purified YhiP and TppB into proteoliposomes allowed their transport to be studied and compared *in vitro*. In the presence of an artificial proton gradient both transporter proteins showed Gly-Sar influx and when immobilized on a chip-based sensor also generated transport currents. Taken together, *yhiP* encodes a protein that mediates a proton-dependent electrogenic transport process for di- and tripeptides with similarities to the mammalian PEPT1-proteins. Based on our results we propose to rename YhiP to DtpB (di- and tripeptide permease B) and TppB to DtpA.

7.3 INTRODUCTION

Uptake of peptides into *E. coli* cells is thought to be mediated by three different transport systems represented by the dipeptide permease Dpp, the tripeptide permease TppB and the oligopeptide permease Opp (1-3). Despite the fact that these proteins seem to discriminate according to the backbone length of the peptides they do show some overlapping substrate specificity. Their prime physiological role lies in the uptake of amino acids in peptide form as an economic process to provide energy substrates and building blocks for cellular metabolism, but peptide uptake seems to also be involved in signalling events and metabolic adaptation (4). Whereas Opp and Dpp are ATP binding cassette (ABC) transporters with periplasmatic binding proteins (5-7) TppB belongs to the family of proton-dependent peptide symporters that utilize the proton gradient as driving force and lack cognate binding proteins (8). After identification of the corresponding gene *ydgR* (8,9) it was genetically classified as member of the peptide transport (PTR) family (also called proton-dependent oligopeptide transporter or POT family). These transporters are found essentially in all living organisms from bacteria to man, with examples such as DtpT from *Lactococcus lactis*, Ptr2 from *Saccharomyces cerevisiae* or the mammalian transporters PEPT1 and PEPT2 (10,11). First attempts to functionally characterize TppB in *Salmonella typhimurium* (12,13) and in *E. coli* employing deletion mutants (14,15) revealed some information on substrate preferences and demonstrated up-regulation of transport upon anaerobiosis and by leucine (13,16). Although TppB, like Dpp, was shown to transport di- and tripeptides, it seemed to prefer tripeptides and in particular those of hydrophobic nature. The transport mode of TppB, however, was poorly characterized.

We recently cloned *ydgR* encoding TppB and over-expressed the transport protein for a detailed functional analysis and initial structural characterization. We observed a striking functional similarity to the mammalian PEPT1 protein

despite the fact that the *E. coli* and mammalian proteins share only very low sequence homology (17). Bacterial PTR transporters are not only interesting for defining their functional features and physiological importance but also as they are more easily purified in large amounts for crystallisation approaches to derive structural models. These might then be applied also to the human proteins, which are of prime importance for example for delivery of peptidomimetic drugs in the intestine.

Three of the four *E. coli* members of the PTR family (*yhiP*, *ybgH* and *yjdL*) have not been studied yet. Here we report the cloning, purification, and structural and functional characterization of the *yhiP* gene product. Negative-stain transmission electron microscopy (TEM) of detergent-solubilized YhiP revealed its low-resolution structure suggesting a monomeric state of the protein. Furthermore, we compared the YhiP activity with that of TppB to improve our understanding of the multitude of peptide transport processes in prokaryots. We establish that *ydgR* and *yhiP* code for prototypical H^+-coupled symporter proteins that are specific for di- and tripeptides. In analogy to similar transporters in other prokaryotes we propose to name them Dtp (di- and tripeptide permeases) with the first one identified DtpA (former TppB, YdgR) and DtpB (YhiP) described here as the second member of the PTR family in *E. coli*.

7.4 MATERIALS AND METHODS

Cloning and expression of the transport proteins in E. coli – *dtpB (yhiP)* was cloned into the pET-21 vector (T7 promotor, C-terminal hexahistidine-tag, Novagen) using the primers AAAAGCTTATGAATACAACAACACCCATG and AAACTCGAGATGGCTTTCCGGCGTC-GC as described before for *dtpA* (*ydgR*) (17). Expression was performed as described (17).

Western blot analysis – *E. coli* inner membranes were prepared as described for purification in (17). From a sample corresponding to 250 µl culture the proteins were separated by SDS-PAGE and blotted onto PVDF membranes (Millipore) following detection as described (17).

Transport assays in intact cells – Transport assays were performed 3 h after induction with IPTG with radioactively labelled glycylsarcosine ([^{14}C]Gly-Sar, custom-synthesis by Biotrend, Cologne, Germany). Around 3×10^9 cells were harvested by centrifugation and resuspended in 1 ml buffer (25 mM Hepes/Tris 7.4, 150 mM NaCl, 5 mM glucose). The assay volume of 50 µl was made up with 20 µl cells (6×10^7 cells), 5 µl of a 10 mM Gly-Sar (containing 0.1 µCi [^{14}C]Gly-Sar, final concentration 1 mM, 2 mCi/mmol), and 25 µl of buffer (control) or a 20 mM competitor solution (final concentration 10 mM). The cells were incubated with the substrate for 30 s at 37 °C, followed by filtration (0.45 µm mixed cellulose esters, ME25 Whatman) and washing twice with ice-cold buffer. Uptake of [^{14}C] Gly-Sar was quantified by liquid scintillation counting. Significance ($p < 0.05$) was determined by t-test using SigmaPlot. For determination of apparent K_t values uptake rates at different substrate concentrations were determined by linear regression of time course experiments (10, 20, 40, 60 s). Then K_t values and maximal uptake rate were determined by nonlinear regression of the rates versus concentration using SigmaPlot.

Growth experiments – Cells were grown in LB to OD_{600}=1 and transferred to a 12 well plate (1 ml/well) and IPTG (0.1 mM) and alafosfalin (500 µg/ml) were added. Growth was continuously recorded by measuring the OD_{600} while shaking at 37 °C in a plate reader (Varioscan, Thermo).

Purification of the His-tagged-proteins – Cell pellets from 500 ml culture were resuspended in 20 ml lysis buffer (10 mM Hepes/Tris pH 7.4, 1 mM DTT, 0.5 mM EDTA) and broken by sonification (10 cycles of 30 s). Membranes were prepared as described for purification in (17) and

solubilized (60 min, 4 °C) in 10 mM Hepes/Tris pH 7.4, 150 mM NaCl, 1% DDM, 5% glycerol, 0.1 mM TCEP, 30 mM imidazole at a protein concentrations of 1-3 mg/ml. After centrifugation (40000 × g, 20 min) the supernatant was loaded to a Ni^{2+}-NTA column (HisTrap FF, GE) with a FPLC (ÄKTA, Amersham) and washed with running buffer (10 mM Hepes/Tris pH 7.4, 150 mM NaCl, 30 mM imidazole, 0.06% DDM, 5% glycerol, 0.1 mM TCEP). Protein was eluted with a gradient from 30 to 250 mM imidazole in running buffer. Elution yielded a sharp peak at about 150 mM imidazole, which was shown by SDS-PAGE and Western Blotting to contain the purified transporters.

Negative-stain transmission electron microscopy – Purified DtpB protein at ~6 µg/ml was adsorbed for 10 seconds to parlodion carbon-coated copper grids rendered hydrophilic by glow discharge at low pressures in air. Grids were then washed with three drops of double-distilled water and stained with two drops of 0.75% uranyl formate. Electron micrographs of purified DtpB were recorded with a Hitachi H-7000 TEM operated at 100 kV.

Reconstitution into proteoliposomes – For reconstitution, 250 µl *E. coli* lipids at a concentration of 20 mg/ml in $CHCl_3$ (total extract, Avanti Polar Lipids) were dried to a thin film under a stream of N_2 and further incubated for 30 min in vacuum. The lipids were resuspended in 500 ml buffer (50 mM KPO_4 pH 6.3), sonicated until clearing (3 × 3 s, probe type sonifier) and destabilized with 0.5% DDM. 100 µg (about 100 µl) purified protein as eluted from the Ni^{2+}-NTA column was added. After 10 min on ice, the detergent was removed by adding 200 mg Bio-Beads SM-2 (BioRad) and incubation for 4 h at 4 °C. The detergent removal step was repeated 3 times with new Bio-Beads and incubation at 4 °C for a total of 24 h.

Uptake with cytochome c oxidase energized vesicles – To generate a proton gradient, membrane vesicles or proteoliposomes were fused with proteoliposomes containing cytochome *c* oxidase (18,19). Cytochrome *c* oxidase (1 mg, purified from bovine heart (20)) was reconstituted with 40 mg *E. coli* lipids and 12 mg octyl-glucoside in 2 ml buffer (50 mM KPO_4 pH 6.3) and detergent was removed by dialysis. These proteoliposomes were mixed with *E. coli* membranes (usually 2 mg lipid + 100 µg membrane protein) or with other liposomes containing purified transporter (usually 27 µg protein) in a final volume of 250 µl. After addition of 1 mM $MgSO_4$ a cycle of freezing (liquid nitrogen) and thawing (room temperature) was performed, which was followed by sonification for 3 s with a probe type sonifier. For the uptake assay, 2.5 µl [^{14}C]Glycylsarcosine (2.5 µCi, 175 µM final concentration, 57 mCi/mmol) was added to the proteoliposomes, which were stirred at 37 °C. At indicated times, 25 µl aliquots were withdrawn, diluted with 2 ml 100 mM LiCl, and filtered (0.45 µm, mixed cellulose esters: ME25, Whatman). The filters were washed with 2 ml 100 mM LiCl followed by scintillation counting. Proton gradient generation was started by the addition of 20 mM ascorbate, 200 µM TMPD (N,N,N',N',Tetramethyl-p-phenylendiamine) and 20 µM cytochrome *c* (Sigma). Fresh ascorbate was added (1:50) when the assay solution had turned blue due to oxidized TMPD. CCCP (Carbonyl cyanide 3-chlorophenylhydrazone, Sigma) was used at a final concentration of 10 µM to destroy the proton gradient.

Electrical measurements of reconstituted transporters with the SURFE^{2}R^{one} setup – Basis is the solid-supported membrane technology, which allows detection of capacitively coupled currents (21). They are caused by the shift of electrical charges as the transporters go through the transport cycle and the shift can originate from the movement of charged substrates, co-transported ions or of protein moieties carrying partial charges. Experiments were conducted as described in (17) with the following modification: The transport was activated via rapid exchange of a buffer without substrate (30 mM glycine, 140 mM KCl, 25 mM Hepes, 25 mM MES, 2 mM

$MgCl_2$, pH 7.0) to the same buffer containing the substrate glycyl-glycine (30 mM) instead of glycine. After a rapid fluid exchange to the Gly-Gly containing solution, the charging of the proteoliposomes on the sensor driven by the H^+/peptide symport was measured.

7.5 RESULTS

Functional over-expression of DtpB in E. coli – In a similar approach as previously reported for *dtpA* (*tppB, ydgR)* (17) we amplified the *dtpB* (*yhiP*) gene of *E. coli* from genomic DNA by PCR using gene specific primers. The gene was cloned into the pET-21 vector fused with a C-terminal hexahistidine-tag (pET-21-dtpB-His). The expression of the DtpB protein in *E. coli* BL21(DE3)pLysS was tested by Western blot analysis with an antibody against the hexahistidine-tag (Figure 1A). The anti-His antibody detected a protein with an apparent molecular weight of about 40 kDa in membrane preparations of DtpB over-expressing cells. Since no signal was obtained with IPTG-induced cells carrying an empty vector (17) this band clearly represents DtpB. The expected molecular weight based on the amino acid sequence of DtpB is 55 kDa. A similar increased mobility in SDS-PAGE was observed for DtpA, where we provided evidence that the protein was complete by probing it with a second antibody that recognized an added N-terminal tag (17).

To assess the function of the over-expressed DtpB protein we performed uptake experiments using radio-labelled [^{14}C]glycylsarcosine (Gly-Sar), which is a commonly used reporter substrate for mammalian peptide transporters (Figure 1B). Cells expressing DtpB showed a higher uptake of Gly-Sar than control cells. This was also the case for DtpA, indicating that both proteins are mediating efficient Gly-Sar uptake. As a second approach to assess the functionality of the over-expressed transport protein we performed growth experiments (Figure 1C) with the toxic phosphonopeptide alafosfalin, a known substrate of DtpA. *E. coli* cells carrying the expression vector for DtpB or the empty vector were pre-grown in the presence of a non-lethal concentration of alafosfalin (500 µg/ml) and protein expression was induced by addition of IPTG. Cells over-expressing DtpB showed a strong growth inhibition one hour after induction. Since non-induced cells or cells carrying the empty vector showed no growth inhibition, this demonstrates DtpB-mediated uptake of alafosfalin. In similar experiments with DtpA, lower alafosfalin concentrations (200 µg/ml) were needed for growth inhibition (17) suggesting that DtpB may have a lower affinity to alafosfalin than DtpA. To compare the relative substrate affinities of DtpB and DtpA, we determined the uptake rates in cells over-expressing the proteins as a function of Gly-Sar concentration (Figure 1D). Transport was found to be saturable at a v_{max} of 0.21 nmol/s with an apparent K_t of 6.5 mM in case of DtpB and 0.17 nmol/s and an apparent K_t of 1 mM in

TABLE 1
Substrate specificity of DtpA and DtpB

competitior	[^{14}C]Gly-Sar uptake [%]	
	DtpA	DtpB
none	100 ± 19	100 ± 11
Gly-Sar	27 ± 11*	71 ± 12*
Gly-Sar (100 mM)	n.d.	32 ± 6*
Ala	76 ± 2	124 ± 20
Ala-Ala	1 ± 0*	17 ± 3*
Ala-Ala-Ala	1 ± 0*	17 ± 3*
Ala-Ala-Ala-Ala	78 ± 22	110 ± 20
L-Ala-D-Ala	63 ± 11*	102 ± 3
D-Ala-L-Ala	5 ± 5*	88 ± 3
D-Ala-D-Ala	84 ± 32	102 ± 3
Gly-Lys	90 ± 40	99 ± 2
Lys-Gly	2 ± 1*	57 ± 12*
Gly-Asp	26 ± 6*	17 ± 1*
Asp-Gly	54 ± 16*	97 ± 5
Cefamandole	145 ± 50	101 ± 18
Cefuroxime	100 ± 2	99 ± 19
Cephradine	29 ± 5*	57 ± 4*
Cefalexin	35 ± 7*	78 ± 2*
Cefadroxil	19 ± 1*	90 ± 23

Uptake of Gly-Sar in % ± standard deviation (n ≥ 2). Competitor was 10 mM (10-fold excess) except otherwise indicated. 30 s uptake time. *: significant ($p < 0.05$) reduction relative to assay containing no competitor. 100% for DtpA = 40 pmol*s^{-1}*0.2 OD^{-1}, 100% for DtpB = 10 pmol*s^{-1}*0.2 OD^{-1}

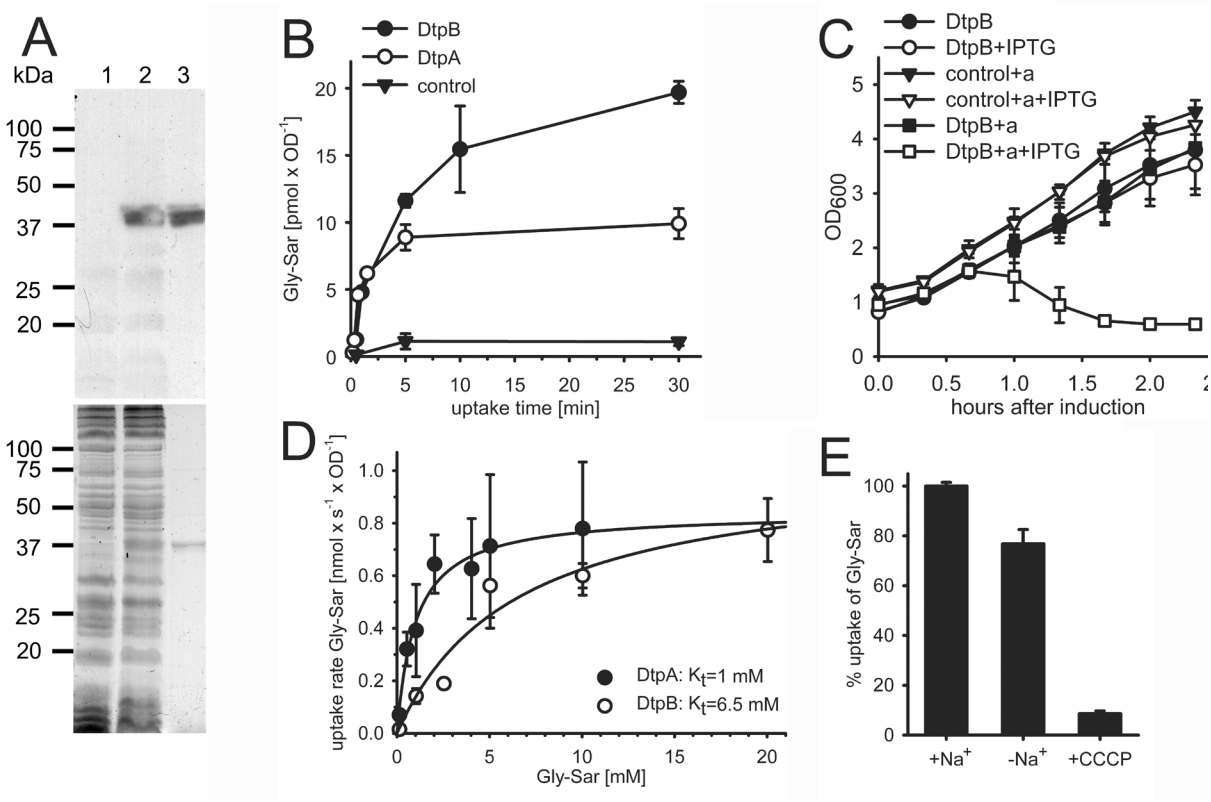

Figure 1. Over-expression of dtpB in E. coli. (A) Anti his-tag Western blot (upper panel) and Coomassie stained (lower panel) SDS-gel (12.5% acrylamide) of E. coli membranes of control cells (lane 1, 50 µg total protein) or membranes containing over-produced DtpB (lane 2, 50 µg total protein) or purified YhiP (0.5 µg upper panel, 1.5 µg lower panel). (B) Uptake of [14C]Gly-Sar (1 mM) by E. coli cells carrying the expression plasmid for dtpA, dtpB or the empty vector 3 hours after induction with IPTG. Over-expression of either gene causes increased Gly-Sar transport relative to the control. n=2, ±SD. (C) Growth inhibition by alafosfalin. Induced (+IPTG) or not induced cells, carrying either the expression plasmid for dtpB or an empty plasmid, were grown in the presence of 500 µg/ml (2.55 mM) alafosfalin (+a) or without alafosfalin. n=2, ±SD. (D) Saturation kinetics for DtpA and DtpB. Uptake assays were performed with E. coli cells over-expressing dtpA or dtpB, in the presence of various concentrations of Gly-Sar. The uptake velocities were determined using four time points (10, 20, 40, 60 s). Kt-values were determined by nonlinear regression analysis of the presented data. n≥2, ±SD. (E) Sodium and proton dependence of Gly-Sar uptake. For E. coli cells over-expressing dtpB, Gly-Sar (50 µM) uptake rates were determined from a 10-min time course. The rates were determined in buffers containing 150 mM NaCl (+Na+), cholinechloride (150 mM) or in the presence of NaCl and the protonophore CCCP (10 µM). 100% corresponds to 3.3 pmol x s-1 x OD-1. n=2, ±SD.

case of DtpA. Despite similar maximal transport rates there are marked difference in affinities with DtpB representing a system of lower affinity. We next asked whether transport requires Na^+ or H^+ (Figure 1E). When we replaced Na^+ in the buffer by choline, this had only a minor effect on uptake of Gly-Sar, whereas the presence of the proton ionophore CCCP caused a complete inhibition of transport. This suggests that uptake via DtpB depends on the proton-motive force and that DtpB may function as a proton-peptide cotransporter.

Substrate specificity – To determinate the substrates transported by DtpB (and for comparision with DtpA), we performed competition experiments with 1 mM Gly-Sar as a substrate and 10 mM of the competitors in *E. coli* cells over-expressing either DtpB or DtpA. The residual uptake rates in the presence of the competitor are presented in Table 1. The substrate specificity data presented here for DtpA essentially match those previously determined with β-Ala-Lys(AMCA), a fluorescent dipeptide analogue used as a substrate in our previous study (17). Similar to DtpA, neither the amino acid Ala nor the tetrapeptide Ala-Ala-Ala-Ala caused significant inhibition of DtpB-mediated transport. In the case of DtpB, essentially all tested competitors reduced transport to a lower extent than with DtpA, which reflects its generally lower affinity: Despite this, there were differences between DtpB and DtpA when inhibitory effects of individual

peptides were compared. When for DtpA L-Ala in dipeptides was substituted by D-stereoisomers, the competitor affinity was less impaired when the D-isomer was present in the N-terminal position. Virtually no competition was seen when both residues are D-isomers. In the case of DtpB, D-Ala in N-terminal position resulted in a very low competitor affinity, while L-Ala-D-Ala and D-Ala-D-Ala did not cause any inhibition. We also assessed the effects of charged amino acid residues and their spatial position on the uptake of Gly-Sar. In the case of DtpA, a positively charged amino acid at the N-terminus caused stronger inhibition than when located in the C-terminus. A negatively charged residue was clearly preferred at the C-terminus. This also was observed with DtpB but here inhibition was higher with Gly-Asp than in case of DtpA. Selected β-lactam antibiotics are known substrates of mammalian peptide transporters (peptidomimetics). DtpA showed efficient inhibition when an excess of cefadroxil, cefalexin and cephradine was added, whereas only cephradine significantly reduced Gly-Sar uptake via DtpB.

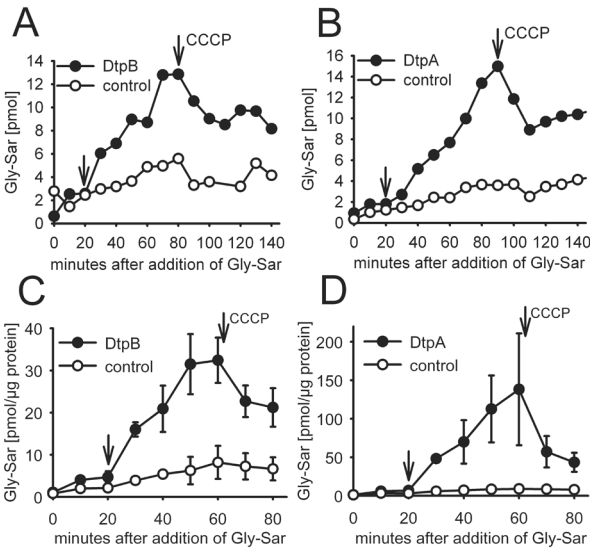

Figure 3. In vitro transport studies of DtpB and DtpA. (A) Uptake of Gly-Sar (175 µM) in membrane vesicles generated by fusion of membranes from DtpB over-expressing or control cells with cytochrome c oxidase containing proteoliposomes. A proton gradient was established by addition of ascorbate/TMPD/cytochrome *c* and destroyed by addition of 10 µM CCCP as indicated. (B) like (A) with DtpA instead of DtpB. (one of two similar experiments). (C) Uptake of Gly-Sar (175 µM) in liposomes containing purified DtpB and cytochrome *c* oxidase. Generation of proton gradient was started by addition of ascorbate/TMPD/cytochrome *c* and destroyed by addition of 10 µM CCCP as indicated. (one of two similar experiments). (D) like (C) with purified DtpA instead of DtpB. (one of two similar experiments)

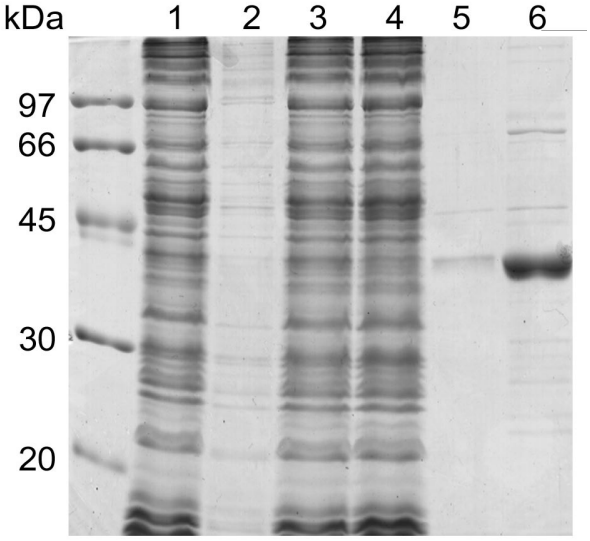

Figure 2. Purification of DtpB (A) Coomassie stained SDS-gel (12.5% acrylamide) of a Ni^{2+}-NTA purification of DtpB.Lanes: 1) DDM solubilized membranes, 2) pellet after solubilization, 3) supernatant after solubilization, 4) Ni^{2+}-NTA flow-through, 5) wash, 6) elution at about 150 mM imidazole.

Assessing the transport mode in reconstitution experiments – Although transport inhibition by the protonophore CCCP (Figure 1E) suggested H^+-dependence, we attempted to characterize the transport mode in more detail by reconstitution in a cell-free system. Membranes of *E. coli* containing the over-expressed transport proteins were prepared and fused with proteoliposomes containing bovine cytochrome *c* oxidase. A transmembrane proton gradient (inside negative) was generated, when cytochrome *c* oxidase was supplied with electrons, which are transferred from reduced ascorbate via TMPD and cytochrome *c*. Before addition of ascorbate, we observed a slow influx of Gly-Sar into all types of vesicles. Addition of ascorbate caused a marked increase in Gly-Sar influx into vesicles containing either DtpA or DtpB, but not into control vesicles (Figure 3A/B). Efflux of Gly-Sar was observed when the proton gradient was dissipated by addition of CCCP, indicating accumulation of

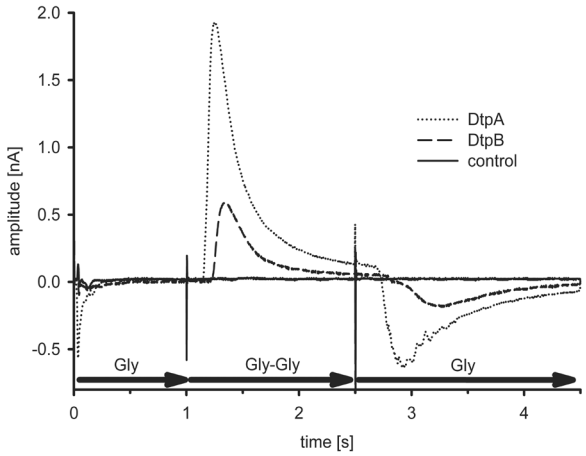

Figure 4. Electrophysiology on proteoliposomes. Electrical response of liposome reconstituted DtpB (dashed line) and DtpA (dotted line) to a change from solution without substrate (Gly) to solution with substrate (Gly-Gly). The solid line shows the recording from a sensor loaded with protein-free liposomes. The current peaks indicate proton co-transport. (one of >10 similar experiments using ≥2 sensors)

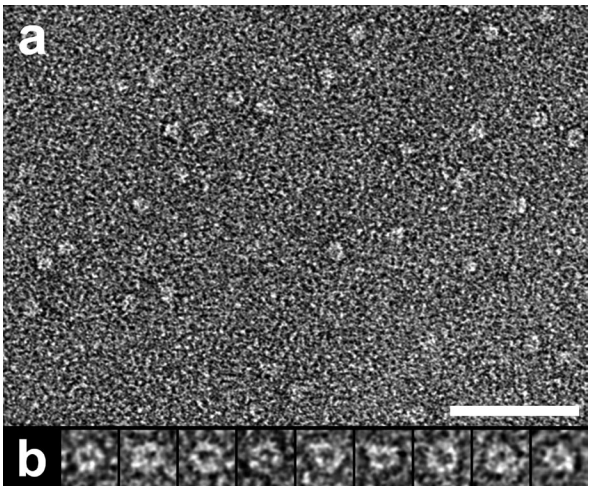

Figure 5. Negative-stain TEM of purified DtpB protein. The homogeneity of the purified DtpB proteins is reflected in the electron micrograph. DtpB proteins are donut-shaped with a central depression and have a diameter of 8.0±0.3 nm (n=120). The scale bar represents 50 nm, and the frame size of the magnified particles in the gallery is 13 nm.

substrate above the extravesicular concentration in the presence of the proton-gradient.

Next we purified DtpB and DtpA from *E. coli* membranes. The membranes were solubilized with the detergent n-dodecyl-β-D-maltoside (DDM) and the proteins were purified by metal-affinity chromatography making use of the C-terminal His-tags. This procedure yielded about 2-3 mg of relatively pure protein per liter of cell culture (Figure 2). For reconstitution into liposomes the detergent was removed by adsorption to BioBeads. Subsequently the liposomes containing the peptide transporters were fused with vesicles containing cytochrome *c* oxidase using a freeze-thaw-sonication cycle. Gly-Sar uptake after initiation of the proton gradient increased almost 10-fold in case of DtpB and 60-fold in case of DtpA (Figure 3C/D) and substrate efflux was observed after addition of CCCP. To unequivocally demonstrate the electrogenic nature of the transport process we next used a chip-based assay system, in which the proteins were adsorbed onto a gold surface of the SURFE^{2}R^{one} setup (IonGate Biosciences GmbH, Germany). This system is based on the solid-supported membrane technology and allows detection of capacity-coupled currents induced by movement of charged molecules across a lipid bilayer (21,22). Transport was initiated by rapid exchange of a buffer containing glycine against a buffer containing the dipeptide substrate glycyl-glycine. Figure 4 demonstrates that Gly-Gly induced significant currents in DtpA- and DtpB containing proteoliposomes but not in control liposomes lacking the transport proteins. Since the transport studies were performed with Na^+-free buffers at pH 7.0, the observed currents must originate from proton movement coupled to dipeptide translocation in a symport mechanism. Like in the cytochrome *c* oxidase energized vesicles, DtpA caused a higher signal than DtpB, suggesting that DtpA has either a higher turn-over rate or higher stability *in vitro*. In intact cells, however, both transporters possessed similar maximal transport rates.

Characterization of DtpB by TEM – To unveil the low-resolution structure of DtpB, DDM-solubilized His-tagged protein was negatively stained and examined by TEM. The homogeneity of purified DtpB is displayed in Figure 5. Single DtpB proteins are discerned as donut-like particles with a diameter of 8.0±0.3 nm (n=120) and a central stain-filled indentation (see gallery of well preserved YhiP top views in Figure 5).

7.6 DISCUSSION

Despite the fact that peptide uptake in *E. coli* has been studied for more than 20 years (1-3), essentially no biochemical characterization has been performed on DtpA. Another prokaryotic member of the PTR family, DtpT from *Lactococcus lactis* has been characterized in more detail and some features of its membrane topology are known (23). With the focus on the four *E. coli* genes encoding proteins that carry the typical PTR-family motif (11) we first cloned, purified and characterized the mode of action of the DtpA (YdgR) protein (17). Here we provide a detailed functional analysis of DtpB as the second yet uncharacterised PTR member in *E. coli*. To assess the differences between DtpA and DtpB and better understand why there are four putative peptide transporters in *E. coli*, we compared the transport features of DtpA and DtpB under identical experimental conditions.

Substrate specificity – DtpB was already cloned and expressed in a global membrane proteome analysis approach of *E. coli*, where by GFP and PhoA fusion constructs the C-terminus was found to be localized in the cytoplasm (24). We here show that DtpB over-expressed in *E. coli* possesses an apparent molecular weight of 40 kDa in SDS-PAGE and can be purified in quantities sufficient for reconstitution experiments with conservation of function (Figure 2). Gly-Sar, a commonly used radiotracer substrate for mammalian peptide transporter proteins, showed saturation kinetics for uptake into cells over-expressing DtpB with a fairly low affinity (6.5 mM) and transport was sensitive to a protonophore suggesting proton-coupling. A previous study with *E. coli* mutants lacking the two major peptide transporters Opp and Dpp but containing DtpA could not detect Gly-Sar uptake (25). Since we clearly show Gly-Sar uptake in cells over-expressing DtpA and in liposomes containing the purified protein (Figs. 1B/D, 3D), this finding may result from low endogenous expression of DtpA in native cells. The apparent affinity of DtpA for Gly-Sar (1 mM) is in the same range as for many other model dipeptides (17). Alafosfalin, a known toxic substrate for bacterial peptide transporters, was described to be specifically transported by DtpA (12). We here show that DtpB also confers alafosfalin-toxicity in *E. coli* cells over-expressing the protein, but higher concentration of alafosfalin are needed than for DtpA (500 instead of 200 μg/ml) to obtain toxicity. This indicates a lower affinity of DtpB for alafosfalin which already was observed for Gly-Sar.

The substrate specificity was addressed by competition experiments with Gly-Sar as substrate (Table 1). The inhibition pattern observed for DtpA confirmed previous data obtained by using the fluorescent substrate β-Ala-Lys(AMCA) (17). Since DtpB failed to transport β-Ala-Lys(AMCA) we here employed Gly-Sar and compared substrate specificity for both proteins under identical experimental conditions. DtpB, like DtpA, has a clear preference for di- and tripeptides composed of L-amino acids. The chain length restriction and the distinct stereoselectivity with a tolerance of D-alanine in N-terminal position matches well with data of the mammalian PEPT1 and PEPT2 proteins (26,27). DtpB, like DtpA, discriminates dipeptides based on the position of charges within the substrate. Peptides containing a positively charged side chain in N-terminal position are good competitors, but ability to compete is lost when the charge is present at the C-terminus. The reverse pattern was observed for negatively charged side chains, which are generally better tolerated at the C-terminal position. A similar charge preference is found in the mammalian peptide transporters (28,29). Of the ß-lactam antibiotics tested, only cephradine and cefalexin showed modest inhibition of Gly-Sar influx via DtpB whereas for DtpA best inhibition was seen with cefadroxil followed by cephradine and cefalexin. The mammalian peptide transporter PEPT1, which is expressed in the intestine, also clearly prefers those substrates but does not transport cefamandole or cefuroxime that are considered to be inactive when administered orally (30). Taken together, DtpB shows essentially the prototypical

substrate characteristics of a di-/tripeptide transporter of the PTR-family. Despite its lower affinity for Gly-Sar and alafosfalin, it is similar to DtpA in the substrate recognition pattern and also shows similarities to the mammalian transporters of the PEPT-subgroup.

Transport mode – Inhibition of substrate influx via DtpB in *E. coli* cells in the presence of CCCP already suggested proton-coupling. This was confirmed by reconstitution of the purified protein into proteoliposomes containing cytochrome *c* oxidase for energization as previously shown for the *Lactococcus lactis* orthologue DtpT (31). When an electrochemical proton gradient was generated, uptake via DtpB as well as DtpA proceeded until the gradient was collapsed by adding a protonophore. Moreover, capacity-coupled currents obtained with immobilized proteoliposomes on the SURFE^{2}R^{one} chip document the electrogenic nature of the DtpB and DtpA mediated transport process, in similarity to on-chip recordings with membranes containing the mammalian PEPT transporters (32). This unequivocally establishes that dipeptide transport by the two bacterial transporters occurs by H^+-symport.

Structure and oligomeric state of purified DtpB – TEM of negatively stained DDM-solubilized DtpB proteins revealed a donut-like structure with a distinct central depression (Figure 5). The measured diameter of the YhiP particles was ~8 nm similar to that of DtpA (17) suggesting a monomeric state of the protein.

Conclusions – DtpB is shown here to represent the second proton-dependent di- and tri-peptide transporter of *E. coli*, with functional features similar to other PTR family members including those from mammals. Although DtpB appears to represent a system of lower affinity there is an obvious redundancy in substrate specificity compared with DtpA. This includes Gly-Sar, for which we here clearly show that it is a DtpA substrate. The question why *E. coli* has two proteins with similar functional features will likely be resolved when their regulation and expression under different environmental conditions is studied.

7.6 ACKNOWLEDGEMENTS

We would like to thank Daniela Kolmeder for excellent technical assistance and IonGate BioSciences GmbH for the possibility to use the SURFE^{2}R^{one} setup. This work was supported by the Sixth Framework Programme of the European Union (*EUGINDAT*, European Genomics Initiative on Disorders of Plasma Membrane Amino Acid Transporters).

7.8 REFERENCES

1. Payne JW. and Smith MW. (1994) *Adv Microb Physiol* **36** 1-80
2. Higgins CF. and Gibson MM. (1986) *Methods Enzymol* **125** 365-377
3. Payne JW. and Marshall NJ. (2001) *Microbial Transport Systems (Winkelmann G., Ed.) Wiley-VCH, Weinheim, Germany*, 139-164
4. Detmers FJ., Lanfermeijer FC. and Poolman B. (2001) *Res Microbiol* **152** 245-258
5. Abouhamad WN., Manson M., Gibson M. M. and Higgins CF. (1991) *Mol Microbiol* **5** 1035-1047
6. Higgins CF. (2001) *Res Microbiol* **152** 205-210
7. Hiles ID., Gallagher MP., Jamieson DJ. and Higgins CF. (1987) *J Mol Biol* **195** 125-142
8. Smith MW., Tyreman DR., Payne GM., Marshall NJ. and Payne JW. (1999) *Microbiology* **145** 2891-2901
9. Goh EB., Siino DF. and Igo MM. (2004) *J Bacteriol* **186** 4019-4024
10. Daniel H., Spanier B., Kottra G. and Weitz D. (2006) *Physiology (Bethesda)* **21** 93-102
11. Steiner HY., Naider F. and Becker JM. (1995) *Mol Microbiol* **16** 825-834
12. Gibson MM., Price M. and Higgins CF. (1984) *J Bacteriol* **160** 122-130
13. Jamieson DJ. and Higgins CF. (1984) *J Bacteriol* **160** 131-136
14. Payne JW., Grail BM. and Marshall NJ. (2000) *Biochem Biophys Res Commun* **267** 283-289
15. Payne JW., Grail BM., Gupta S., Ladbury JE., Marshall NJ., O'Brien R. and Payne GM. (2000) *Arch Biochem Biophys* **384** 9-23
16. Jamieson DJ. and Higgins CF. (1986) *J Bacteriol* **168** 389-397
17. Weitz D., Harder D., Casagrande F., Fotiadis D., Obrdlik P., Kelety B. and Daniel H. (2007) *J Biol Chem* **282** 2832-2839
18. Driessen AJ., de Vrij W. and Konings WN. (1985) *Proc Natl Acad Sci USA* **82** 7555-7559
19. Stolz J., Stadler R., Opekarova M. and Sauer N. (1994) *Plant J* **6** 225-233
20. Yu C., Yu L. and King TE. (1975) *J Biol Chem* **250** 1383-1392
21. Meyer-Lipp K., Ganea C., Pourcher T., Leblanc G., and Fendler K. (2004) *Biochemistry* **43** 12606-12613
22. Zuber D., Krause R., Venturi M., Padan E., Bamberg E. and Fendler K. (2005) *Biochim Biophys Acta* **1709** 240-250
23. Hagting A., vd Velde J., Poolman B. and Konings WN. (1997) *Biochemistry* **36** 6777-6785
24. Daley DO., Rapp M., Granseth E., Melen K., Drew D. and von Heijne G. (2005) *Science* **308** 1321-1323
25. Payne JW., Payne GM., Gupta S., Marshall NJ. and Grail BM. (2001) *Biochim Biophys Acta* **1514** 65-75
26. Terada T., Sawada K., Irie M., Saito H., Hashimoto Y. and Inui K. (2000) *Pflugers Arch* **440** 679-684
27. Biegel A., Knutter I., Hartrodt B., Gebauer S., Theis S., Luckner P., Kottra G., Rastetter M., Zebisch K., Thondorf I., Daniel H., Neubert K. and Brandsch M. (2006) *Amino Acids* **31** 137-156
28. Daniel H., Morse EL. and Adibi SA. (1992) *J Biol Chem* **267** 9565-9573
29. Kottra G., Stamfort A. and Daniel H. (2002) *J Biol Chem* **277** 32683-32691
30. Bretschneider B., Brandsch M. and Neubert R. (1999) *Pharm Res* **16** 55-61
31. Hagting A., Knol J., Hasemeier B., Streutker MR., Fang G., Poolman B. and Konings WN. (1997) *Eur J Biochem* **247** 581-587
32. Kelety B., Diekert K., Tobien J., Watzke N., Dorner W., Obrdlik P. and Fendler K. (2006) *Assay Drug Dev Technol* **4** 575-582

8. General Discussion and Conclusion

Scope of the thesis – The main scope of this thesis was to gain insights into the structure of amino acid and peptide transport proteins. To this aim, we selected bacterial amino acid and peptide transporters with high amino acid sequence homology and identity to the human counterparts. This was AdiC from *E. coli* and SteT from *Bacillus subtilis* as representatives from the APC family, and the *E. coli* proteins YbgH, YdgR (TppB, DtpA) and YhiP (DtpB) as representatives from the PTR family. For each protein an optimized purification protocol was established to obtain high protein amounts, purity and homogeneity suitable for structural analysis. Structure determination was performed by transmission electron (TEM) and atomic force microscopy (AFM). To elucidate the oligomeric state of purified detergent-solubilized amino acid and peptide transporters, blue native-polyacrylamide gel electrophoresis (BN-PAGE) and analysis by TEM of single negatively stained protein particles was performed. BN-PAGE and negative stain TEM revealed for the first time the oligomeric state as well as first structural information for APC and PTR family members. Furthermore, we established protocols for the successful 2D crystallization of AdiC (APC) and YbgH (PTR). The analysis of highly ordered 2D crystals provided the first projection structure of an APC and PTR family member. Besides structure determination using single protein particles and 2D crystals, we developed a new methodological approach for the semi-automated analysis of membrane proteins by high throughput single molecule force spectroscopy (HT-SMFS).

BN-PAGE – To determine whether APC and PTR proteins exists in a monomeric or higher oligomeric state, detergent-solubilized proteins were subjected to BN-PAGE. The power of this technique was nicely demonstrated for the dimeric AdiC (chapter 2), and the monomeric SteT protein (chapter 4), and the monomeric YbgH (chapter 5) and YdgR (chapter 6) proteins. Furthermore, the results from BN-PAGE were corroborated by negative stain TEM of single protein particles, gel filtration and chemical cross-linking.

Single particle analysis – TEM of negatively stained single protein particles represents another well-established method to determine the oligomeric state of detergent-solubilized proteins. Similar to BN-PAGE, results from TEM were in line with other techniques such as BN-PAGE, gel filtration and chemical cross-linking. In addition, TEM of single detergent-solubilized transporters provided dimensions, shape and low-resolution structure of the protein under investigation. Electron micrographs of the monomeric SteT revealed a slightly elliptical particle with the dimensions of ~6 x ~7 nm (chapter 4), while the AdiC dimer exhibited dimensions of ~7.5 x ~12 nm (chapter 2). Also the three monomeric PTR members matched perfectly with the results from the APC proteins, when the slightly higher molecular weight for the PTR proteins was considered. YbgH, YdgR and YhiP had a diameter of ~8 nm (chapter 5, 6 and 7). As stated above, single particles analysis is a powerful technique to reveal structural features. The two ring-like structures of the AdiC dimer in our single particle studies were confirmed by the calculated projection map (chapter 2) from 2D crystals. Furthermore, single particle studies on SteT (chapter 4), YdgR (chapter 6) and YhiP (chapter 7) present the first structural data at low-resolution for these proteins and transport protein families.

2D crystallization – Highly ordered 2D crystals were reported for AdiC and YbgH (chapter 2 and 5). For AdiC, the introduction of the point mutation (i.e., AdiC-W293L) represented the breakthrough that enabled us the growth of highly-ordered crystals and the collection of

high resolution data by cryo-TEM. We suggest that the point mutation induced a stabilization of AdiC protein leading to the adoption of a fixed conformation. For AdiC-W293L, we managed to calculate a projection map at 6.5 Å resolution. The exact arrangement of the transmembrane helices could not be revealed in the projection structure. However, we showed that most of the transmembrane helices must be tilted with the exception of one (chapter 2). Also a member of the PTR family was successfully crystallized. Well-ordered tubular crystals of YbgH were analyzed by negative stain TEM. The calculated YbgH low resolution projection map represents the first structural view of a member of the PTR family (chapter 5).

Single molecule force spectroscopy – AdiC reconstituted in *E. coli* polar lipids was used to demonstrate that single molecule force spectroscopy (SMFS) can be automated to a larger extent. SMFS was used to screen for molecular interactions between AdiC and its substrates. The recorded force spectra of AdiC in the absence and presence of L-arginine, D-arginine and agmatine showed no significant differences. Considering that the unfolding process of AdiC is not affected by ligand binding we proposed that the protein substrate interaction is of weak nature (chapter 3). The establishment of the high throughput SMFS approach will open new avenues for the structural analysis of other membrane proteins.

Conclusion – The AdiC and YbgH proteins represent excellent models for the understanding of the molecular architecture of transporters from the APC and PTR family. Improvement of the present 2D crystals in the near future will allow the determination of the AdiC and YbgH structures at high-resolution. These structures will contribute to our understanding of the function and mechanism of APC and PTR transport proteins and will facilitate the development of novel drugs by structure-based drug design.

9. Acknowledgments

First of all I would like to thank Prof. Andreas Engel for giving me the opportunity to work in his laboratory during my Ph.D. thesis. Andreas offered great confidence and support, creating a perfect environment for my scientific research. In spite of his busy position he always found time for my requests, providing extremely perceptive, helpful, and appropriate comments.

I am especially grateful to Prof. Dimitrios Fotiadis for being an excellent supervisor. I felt his enthusiasm and excitement for structural biology, which inspired my research immensely. Dimitrios arranged an interesting challenge for my thesis by our participation in the European initiative Eugindat. I feel extremely lucky being taught and accompanied by Dimitrios, an outstanding scientist and good friend.

Very special thanks go to Kitaru Suda for having him by my side in the laboratory. It was a great pleasure to work with Kitaru, a great biochemist who provided a lot of practical help and good advices. After more than 30 years in science Kitaru retired one year ago, leaving an irreplaceable position to the lab, and to the entire Biozentrum.

I'm very grateful to Patrick Bosshart for analyzing my samples by the AFM with great engagement. Patrick is a highly talented newcomer in the lab always on the spot for a fruitful discussion or an amusing companionable talk.

Thanks go also to Mohamed Chami for the great help with the cryo-transmission electron microscope. I enormously profited from his excellent skills and knowledge of high resolution imaging.

Furthermore I thank Andreas Schenk for the analysis of 2D crystals by the IPLT software package. His enormous skills and experiences in image processing speed up our structural investigations.

It was a great pleasure to collaborate with Prof. Manuel Palacin, Merce Ratera, Dr. Nuria Reig, Dr. Cesar del Rio, Dr. Jose Luis Vazquez-Ibar from the University of Barcelona, Prof. Hannelore Daniel, Daniel Harder and Dr. Dietmar Weitz from the University of Munich and Prof. Baruch Kanner from the Hebrew University of Jerusalem. The yearly Eugindat meetings provided a pleasant atmosphere for many constructive discussions. The fruits of these excellent collaborations are reflected in my thesis.

I want to express my gratitude to Peter Tittmann and Dr. Heinz Gross, for introducing me to the freeze drying and shadowing technique. It was a pleasure to visit the facility at the ETH Zürich and learn from their sample preparation experiences.

Many people of the group of Prof. Engel assisted to create a pleasurable environment for work. I am especially grateful to Marco Biasini, Sandrine Bingler, Gottfried Eisner, Françoise Erne-Brand, Patrick Frederix, Shirley Müller, Ralf Paul, Ansgar Philippsen, Wanda Kukulski, Hervé Remigy, Philippe Ringler and Paul Werten (arranged alphabetically).

For the Biozentrum support in general I thank Roland Bürki, Robert Häring, Barbara Merz, Gianni Morson Beat Schumacher and Robert Wyss

I wish them all the best for the future.

Finally, I thank my family for their constant support and for having made all this possible.