Original published in Journal of Biological Chemistry 281, 7413-7420 (2006)

HIGH RESOLUTION CRYSTAL STRUCTURES AND MOLECULAR DYNAMICS STUDIES REVEAL SUBSTRATE BINDING IN THE PORIN OMP32*

Ulrich Zachariae‡, Thomas Klühspies‡, Sharmila De‡§, Harald Engelhardt ‡, and Kornelius Zeth**

Departments of ‡Molecular Structural Biology and **Membrane Biochemistry, Max Planck Institute of Biochemistry, Am Klopferspitz 18, D-82152 Martinsried, Germany;

§Physics Department, Dinabandhu Mahavidyalaya, Bongaon, West Bengal, India

Address correspondence to: Harald Engelhardt, Max Planck Institute of Biochemistry Dept. of Molecular Structural Biology, Am Klopferspitz 18, D-82152 Martinsried, Germany Tel. +49 89 8578 2650; Fax. +49 89 8578 2641; E-Mail: engelhar@biochem.mpg.de

The porin Omp32 is the major outer membrane protein of the bacterium *Delftia acidovorans*. The crystal structures of the strongly anion-selective porin alone and in complex with the substrate malate were solved at 1.5 and 1.45 Å resolution, respectively, and revealed a malate-binding motif adjacent to the channel constriction zone. Binding is mediated by interaction with a cluster of two arginine residues and two threonines. This binding site is specific for Omp32 and reflects the physiological adaptation of the organism to organic acids. Structural studies are combined with a 7-ns unbiased molecular dynamics simulation of the trimeric channel in a model membrane. Molecular dynamics trajectories show how malate ions are efficiently captured from the surrounding bulk solution by the electrostatic potential of the channel, translocated to the binding site region, and immobilized in the constriction zone. In accordance with these results, conductance measurements with Omp32 inserted in planar lipid membranes revealed binding of malate. The anion-selective channel Omp32 is the first reported example of a porin with a 16-stranded β-barrel and proven substrate specificity. This finding suggests a new view on the correlation of porin structure with substrate binding in specific channels.

Porins are channel proteins that facilitate the exchange of ions and small hydrophilic substrate molecules across the outer membrane present in Gram-negative bacteria, mitochondria, and chloroplasts (1, 2). A number of bacterial porins have been studied structurally. They all possess β-barrels as the common structural principle (3). The trimeric bacterial porins are usually classified according to their number of barrel-forming β-strands (3, 4): 18-stranded porins function as substrate-specific channels and provide periplasmic binding proteins and uptake systems of the inner membrane with their specific substrates (5, 6); 16-stranded porins are reported to merely act as unspecific ("general") diffusion pores, however typically with a distinct ion selectivity (7-11).

The current classification is mainly based on the crystal structures of 3 different 18-stranded and 7 16-stranded porins from *Proteobacteria*. Structural studies of substrate binding to trimeric porins were, thus, restricted to the sugar-specific maltoporins and to ScrY so far (5, 6). The nucleoside-specific pore protein Tsx and the fatty-acid binding transporter FadL are monomeric 12- and 14-stranded β-barrel proteins, respectively (17, 18). They are not regarded as porins proper, however, they also belong to the class of passive outer membrane channel proteins. Thus, in a more general view, substrate binding is not a special feature of 18-stranded outer membrane proteins. Several theories were developed to describe channel-substrate interactions which apply to diffusion pores in general (19). A decisive step of substrate translocation was modelled for maltoporin by conjugate peak refinement (20). Non-equilibrium molecular dynamics (MD) and docking methods were applied to investigate the translocation of sugar molecules and antibiotics through OmpF (21-23).

The 16-stranded porin Omp32 is the major protein constituent of the outer membrane from the bacterium *Delftia acidovorans* (formerly *Comamonas acidovorans*; 12, 13). This species, being facultatively pathogenic (14), is closely related to *Neisseriaceae,* which causes serious

infectious diseases in humans. The predominant carbon sources of *D. acidovorans* are organic acids (15), among which malic acid is one of the characteristic and physiologically important substrates (16). The ability of the bacterium to grow on organic acids as a single carbon source suggests that proteins in the inner and outer membranes are specialized for the influx of these molecules.

The Omp32 channel strongly discriminates between anions and cations, with a selectivity factor of ~20 (24). Accordingly, its structure displays a large number of basic residues inside the pore and an unusually narrow constriction zone located approximately halfway along the pore axis (25). Recent MD studies showed that even the small chloride ions experience intimate interactions with charged amino acids arranged as a 'basic ladder' and located in the channel constriction (26), a situation that is less obvious but also present in porins exhibiting a larger pore cross section such as OmpF from *Escherichia coli* (27, 28). The affinity of *D. acidovorans* toward organic acids as carbon source, together with the strong anion selectivity and the small channel diameter of Omp32 suggest that a certain level of substrate specificity may also be present in a porin that would be attributed as unspecific by its structure according to the current classification (3). We combined X-ray crystallography and unbiased MD simulations of a complete membrane system and studied substrate-protein interactions in Omp32 in more detail. The simulation revealed that malate ions are spontaneously attracted by the channels from the surrounding solution, translocated into the pore and bound near the constriction zone. The X-ray structure at 1.45 Å resolution of the porin-malate complex gives a detailed view of protein-substrate interactions, and functional measurements confirm substrate binding in the course of malate translocation.

EXPERIMENTAL PROCEDURES

Crystallization of Omp32 – The preparation of outer membranes from *D. acidovorans* strain JL0 and the purification of porin Omp32 were adapted from (25; see Supplemental Data). Crystallization trials were performed using purified porin at 15 mg/ml in buffer B (20 mM Tris at pH 8.6 and 2% β-D-octylglucoside). Crystallization was tested against crystal screens I and II from Hampton Research by mixing 1 µl protein and 0.5 µl reservoir solution to yield hanging drops which were incubated at 20° C. In order to prepare crystals of Omp32 including organic acids, protein crystals were soaked for 12 h by the addition of 100-200 mM L-malate (MLT structure) to the crystallization drops. All crystals were directly frozen in liquid nitrogen. Data were collected at beamline ID14-EH4 of the synchrotron radiation source ESRF (European Synchrotron Radiation Facility, Grenoble, France) at 100 K and a wavelength of 1.0053 Å. Diffraction patterns were recorded on a ADSC Q4R mosaic CCD detector system and all data indexed, integrated and scaled using programs XDS and XSCALE, respectively (29).

X-Ray Crystallographic Analysis, Data Collection, Phasing and Refinement – Molecular replacement trials using the room temperature structure of Omp32 (PDB-entry 1E54; 25) were performed by means of the program Molrep (30) from the CCP4 program package. The solution observed was further refined by the molecular refinement program Refmac (31). Side chains were successively rebuilt according to $2F_O$-F_C and F_O-F_C electron density maps followed by iterative cycles of model rebuilding with O (32). At this stage the malate ligand was placed into the electron density and refined under the restraints included in the Refmac program libraries. Solvent waters were then placed into 3σ peaks of F_O–F_C-maps using the automated ArpWarp routine (33) and all atoms were refined with independent anisotropic B-factors. The final refined protein model at 1.45 Å of MLT comprises residues 3 to 332 (1.5 Å and residues 1-332 for NAT, the Omp32 without malate) according to the full-length protein sequence. The model geometry was checked with PROCHECK (34), and the solvent content V_M of the crystals was calculated using programs included in the CCP4 program package (35). The structural superposition of MLT with the 1E54 model of Omp32 and the unsoaked structure was carried out using the program TOP (http://gamma.mbb.ki.se/~guoguang/top.html).

Molecular Dynamics – The hexagonal simulation system was built using the original crystal structure of Omp32 (PDB 1E54). A model membrane, consisting of a bilayer of hydrated DMPC lipids, was constructed according to Woolf and Roux (36, 37) and Im and Roux (38, 39). The simulation system comprised an Omp32 trimer embedded in the lipid bilayer. Two water layers were placed above and below the membrane layer, and the

pores were filled with water. The system was equilibrated following the methods given in (38). 12 chloride anions were used to neutralize the system during the equilibration step. Before the production run was started, six chloride ions were removed and six were substituted by malate ions, corresponding to a malate concentration of ~ 25 mM. Three malate ions were placed into the exoplasmic (MAL1-MAL3), three into the periplasmic region (MAL4-MAL6), and equilibrated for further 10 ps. Three of the malate ions possessed L-stereochemistry (MAL1, MAL3, MAL5), the others were D-enantiomers (MAL2, MAL4, MAL6) because *D. acidovorans* consumes both types (Engelhardt, unpublished results). Altogether, the system contained 996 amino acid residues, three protein-bound Ca^{2+} ions, 162 DMPC lipid molecules, 14323 TIP3P water molecules (40), and six malate ions, yielding 76737 atoms in total.

After complete equilibration, a 7 ns MD production run was carried out in a CPTA ensemble with hexagonal periodic boundary conditions. The integration time step was 1.6 fs. The trajectory was generated at a constant normal pressure of 1 atm, a constant temperature of 330 K, and a constant cross-section surface area. The temperature was kept constant using a Nosé-Hoover thermostat. The temperature of 330 K ensured a fluid liquid-crystalline Lα-membrane phase (38). All calculations were performed using CHARMM version 28b2 (41) together with the PARAM22 all-atom parameter set for proteins and phospholipids (42). The partial charges on the malate ions were determined by MNDO in conjunction with Mulliken population analysis using the program InsightII (Accelrys Software, Inc., San Diego, USA). The porin trimers were separated by 99.88 Å in the xy-plane, and the initial distance between the membrane centers was 87 Å in the z-dimension. Bond lengths involving hydrogen atoms were kept constant using the SHAKE algorithm (43). Long-range electrostatic interactions were treated using the Particle-Mesh-Ewald method (44) on a 100x100x72 grid for fast Fourier transforms, $\kappa = 0.32$ Å^{-1}, and a sixth-order B-spline interpolation. The calculation of the van der Waals interactions was switched off between 8-10 Å. Global translation of the trimeric Omp32 was counteracted by the application of a cylindrical harmonic restraint to its center of mass in the xy-plane. The equilibration and the following 7 ns MD trajectory were calculated on 8 processors of a Beowulf linux cluster.

Electrophysiological Experiments – Conductance measurements were performed in sodium L-malate solutions at room temperature, using a two-compartment chamber separated by a septum containing a hole of 90 μm in diameter. The hole was covered by a lipid bilayer made of diphytanoyl phosphatidylcholine (Avanti Polar Lipids, Alabaster, AL). Application of porin, data acquisition, and the digital setup were described previously (24). Measurements were performed at +20 mV with respect to the compartment containing the porin. The recombinant Omp32 used had the same conductance characteristics in KCl solutions as porin purified from outer membranes (Roth and Engelhardt, unpublished results). The current steps of porin insertion events were analyzed by means of our STEPSIZE procedure that is based on a correlation approach and was written for IGOR (WaveMetrics, Lake Oswega, OR).

RESULTS AND DISCUSSION

Crystallization, Structure Determination, and Refinement – The crystallization attempts were aimed at obtaining high resolution structures of the membrane protein Omp32 in order to allow us to identify ordered organic acids which were incubated with the crystals. The protein was initially tested for crystallization in the presence of 2% β-D-octylglucoside under conditions similar to our previous experiments (25) at high sulfate concentrations (>1.5 M). To allow organic acid molecules to compete for possible binding sites we screened for conditions depleted in divalent anions. However, a certain amount of sulfate ions appears to be indispensable for the formation of highly ordered crystalline arrays. Small needle-shaped crystals that grew over a two weeks period were observed with 25% PEG4000, 0.1 M sodium acetate and 0.2 M ammonium sulfate. These conditions were further refined yielding crystals of 0.3 x 0.05 x 0.05 mm. Crystals of appropriate size were directly used or soaked with L-malate before data collection. Soaking experiments with succinate and fumarate did not yield ordered substrate positions admittedly done in the presence of higher amounts of sulfate ions.

Although the room temperature structure (PDB 1E54) was determined at 2.1 Å in rhombohedral space group R3, the data of the newly discovered crystal form were collected at 100 K. The collection statistics of all data are summarized in Table 1 in the supplemental data.

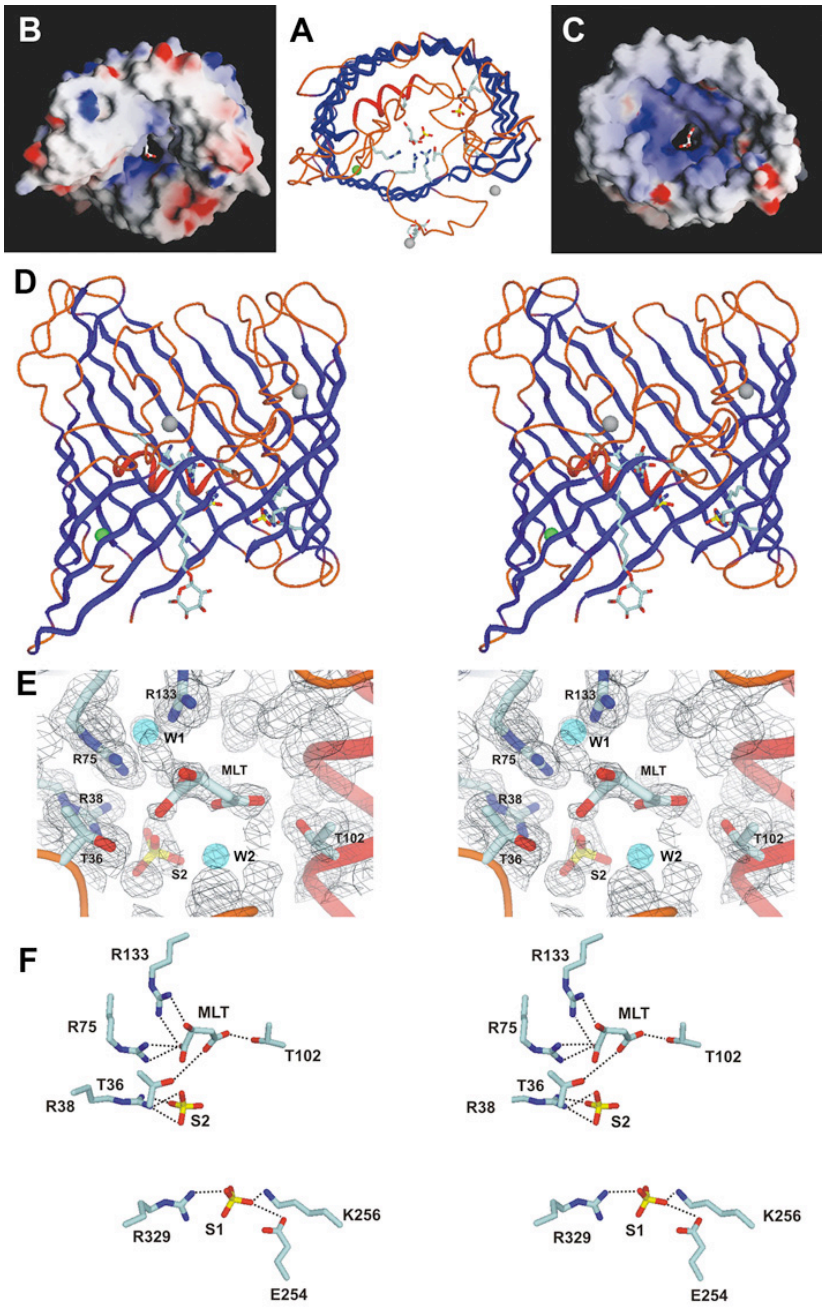


FIGURE 1. Crystal structure of porin Omp32 from *D. acidovorans* and ligand binding sites. (A) The ion channel in a top view representation (from the expoplasmic side) and side view (stereo) as seen from the trimer center (D) with β-strands in blue, α-helices in red and loops in orange (extracellular loops on top). The β-octyl glycoside molecule facing the periplasmic side is shown in white, calcium ions are in grey, chloride ion in green, sulfate ions in yellow, malate is white and residues involved in ionic interactions are drawn. (B, C) Electrostatic surface representation of Omp32 seen from the exoplasmic and periplasmic sides, respectively, the malate ligand is shown in white. (E) Close-up of the substrate binding site and the simulated annealing omit map. The important residues and water molecules (W1, W2) involved in substrate binding are specifically marked. (F) Side view onto the ionic ladder. The residues R133, R75, T102 and T36 are involved in malate (MLT) binding. Additional sulfate ions S1 and S2 as well as their ligating residues are marked. The direction of ion flux from the extracellular to the periplasmic space is from top to bottom. Figs. A,D-F were made with DINO.

In brief, Omp32 diffraction datasets NAT and MLT merge in the hexagonal space group P63 and diffracted to 1.5 and 1.45 Å resolution, respectively.

Initial molecular replacement trials using the room temperature model as a basis resulted in clear solutions for the NAT and MLT data in rotation and translation function, and a crystallographically useful model arrangement with one monomer per asymmetric unit was achieved. The NAT protein structure was refined leading to a complete porin model (residues 1 - 332) and eight ordered residues of the periplasmic porin-associated protein PAP. PAP is highly flexible in crystals so that its complete structure could not be determined yet. The protein homologue of PAP in *Neisseria meningitidis* putatively binds to the peptidoglycan (45) and is assumed to stabilize outer membrane proteins (46). Additional cofactors, two SO_4^{2-}, two Ca^{2+}, and one Cl^- ions were added to both structures (NAT and MLT) according to high difference peaks in F_O-F_C maps. Surprisingly, at the same stage of the MLT model building it became obvious that crystals soaked either with malate or acetate (data not shown) lost their ability to bind the PAP. Concomitant with the loss of PAP

a significant destabilization of the N-terminus was observed. While the MLT structure was refined under identical crystallographic restraints as NAT the resulting electron density allowed tracing of residues 3 - 332 only and the refined structure showed significantly increased B-values for the PAP-binding residues 3 - 6. The displacement of PAP yielded access for the integration of a single ordered detergent molecule into the hydrophobic pocket (Fig. 1A, D). Two minor parts of the structure that deviate between the NAT and MLT structures are located within loop L3 (residues 98 to 110), both are directly involved in malate binding (see below). Further deviations were observed for the region at turn 2 (residues 46 to 51),

More significant F_O-F_C density features within the pore interior not present in the unsoaked crystal structure were attributed to the malate binding site and were included in the further refinement (Figs. 1A, D). The final well-defined model of MLT including 377 (409 for the NAT structure) automatically picked waters has been refined to *R/Rfree*-values of 15.0/17.0% (15.0/17.9 for NAT), average B-factors of 14.8 (16.4 for NAT) for the main chain atoms and an excellent overall stereochemistry (Table I, Supplemental Data). The large number of water molecules observed may originate from the surface charges, leading to the appearance of a densely covered protein channel.

Channel Architecture and Substrate Location – The principal structural features of Omp32 corroborated a 2.1 Å model of the substrate-free porin based on data collected at room temperature (25). The porin is a homotrimer formed by three 16-stranded β-barrels with antiparallel β-strands (β1 – β16). Similar to other structurally known porins Omp32 possesses 8 exoplasmic loops (L1 – L8) and 7 periplasmic β-turns of variable length. The N- and C-termini are oriented toward the periplasmic side and are located in the contact region between neighboring monomers. The longest loop L3 folds into the central pore and narrows the cross section of the channel jointly with L8 and the small protrusion located on β2 (residues 32 – 37). While L3 also constricts the channel in other porins with 16 β-strands the participation of L8 and the existence of a β-strand protrusion is unique for Omp32 (25). Its channel constriction is therefore particularly narrow. The accessible cross section for a molecule of 3 Å in diameter measures only 25 $Å^2$ near the center in the crystal structure (26; Fig. 1).

It is interesting to compare the strategies of pore constrictions in the substrate-specific porins. In maltoporins, loops L1, L3, and L6 fold into the channel (5, 6). Loop 1 is essentially a large 'protrusion' of β-strand 2 that is stabilized by a cystine bridge (47). The loop is located at a position corresponding to that of L8 in Omp32. Loops L3 and L6 jointly form the constriction at the opposite channel wall in maltoporins. L3 in Omp32 possesses an additional protrusion (25) and is much longer, apparently 'replacing' the contribution of L6 in the sugar-specific porins. In principle, we observe a congruence in the building plan of the pore constriction in Omp32, the maltoporins and ScrY, despite the distinct differences in the number of β-strands and in structural detail.

In Omp32, 17 positively charged residues are distributed over both pore openings, 5 arginines are located in the exoplasmic half, while 7 arginines and 5 lysines are located on the periplasmic side. The positive charges in the pore are not compensated by negatively charged groups, which is the case in other porin channels (7-11). This is particularly important for the residues Arg38, Arg75, and Arg133, forming the fully charged arginine ladder in the constriction zone. Together they create a strong electrostatic field inside the channel that attracts anions (Figs. 1B, C; 26, 48). In the MLT structure, five ions (two SO_4^{2-}, one Cl^- and two Ca^{2+} ions) were discovered, some of which were not visible in the 1E54 structure probably due to differing crystallization conditions (i.e. pH), crystal packing, or the significant improvement in resolution (25). In all three structures (1E54, NAT, MLT), a sulfate ion is bound to Arg38 (Fig. 1E, F) in the same position, which is close to the minimum of the free energy a chloride ion experiences inside the channel (26). The second SO_4^{2-} is kept in a clamp-like arrangement between Lys256 (on β13) and Arg329 (on β16) located close to the periplasmic channel exit (Fig. 1D, F). The chloride ion is embedded in a pocket formed by the residues Arg76, Thr78, Ser80, and a water molecule while the two Ca^{2+} are located outside at the monomer-monomer interface.

All these sites were different from the binding site found for L-malate. The position of the divalent acid was identified in a clamp-like arrangement, creating an extended network of salt bridges and H-bonds in close vicinity to the constriction zone (Fig. 1). The carboxy atom O5 of

the malate molecule forms a H-bond to Thr102 of L3 with a distance of 2.8 Å (Fig. 1F). The malate hydroxy atom O3 is involved in a H-bond of 3.1 Å distance to the guanidino group of Arg133. O1 of the second carboxylic group forms two salt bridges of 3.2 and 3.1 Å to Arg75 and Arg133 of the arginine cluster, respectively. O2 interacts with Arg75 and Thr36, a residue of the small protrusion that interrupts β2. Arg75 and Thr36 are also involved in the coordination of Cl^- ions during their passage through the constriction zone according to MD simulations (26). They form, together with Arg38 and Ser108, the most preferred binding site for chloride ions. The center of the Cl^- binding site is, however, different from that for L-malate (Fig. 1). The latter is located closer toward the center of the pore constriction and does not include Arg38 and Ser108 but Arg133 and Thr102. The interacting residues either belong to the arginine cluster or comprise threonines located on the protrusion that is unique for a 16-stranded porin, and on the porin-specific loop L3 at the opposite channel wall. The binding site together with the large number of positive charges, thus, represents a particular structural feature of the porin Omp32. In the sucrose and maltose porins contacts between the sugar substrates and the protein environment are also mediated by arginine residues of the pore interior but in addition also by hydrophobic interactions (5, 6, 49).

Molecular Dynamics – Malate Attraction and Binding – Our computational studies complement the X-ray structure with regard to the pathways and timescale of substrate attraction. A 7 ns trajectory of the full molecular system, including the Omp32 trimer embedded in a phospholipid bilayer and covered by water containing substrate molecules, was obtained by an unbiased MD approach. In the simulation, malate ions were positioned in the bulk aqueous solution. The results of our simulation reached a high level of agreement with the experimentally obtained structural data (Fig. 2; Supplemental Figure). The MD trajectory highlighted the strong interaction of the divalent substrate anions with the positive potential in and near the channels. The malate ions showed largely unaffected diffusive behavior in the bulk solution only during the first few hundred ps of the simulation (Figs. 2A, B). Subsequently, three of the malate anions were captured by the pores through strongly directed diffusion, by which each of the substrate molecules was pulled into a single channel. Occupation of the pores occurred remarkably fast. Malate ion MAL1 (L-malate) was attracted into the exoplasmic vestibule of monomer 1 from the exoplasmic side, moving over a distance of 29 Å in ~1 ns (Fig. 3A), MAL6 (D-malate) was drawn into the periplasmic vestibule of Omp32 monomer 2 within less than 1 ns, covering a distance of ~23 Å (Fig. 3B), while malate ion MAL5 (L-malate) travelled over 25 Å from the periplasmic side into monomer 3 within a period of 2.5 ns. In this way, the central regions of all available pores were occupied by malate ions within 1 – 2.5 ns.

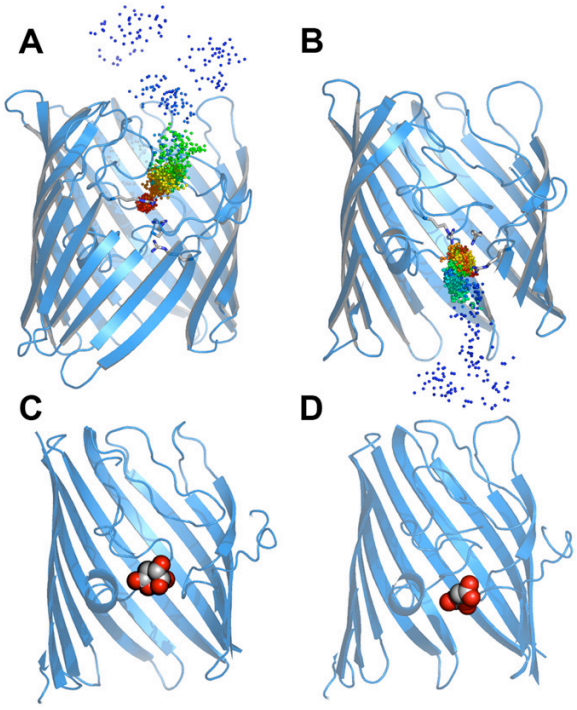


FIGURE 2. (A, B) Trajectories of malate diffusion after 7 ns in Omp32 monomers 1 and 2 (center of mass, progression from dark blue to red), MAL1 starting from the exoplasmic side and MAL6, starting from the periplasmic side. Trajectory points are 5 ps apart. Note that ion capturing from the exoplasmic side is more stringent near the pore constriction than from the periplasmic side. (C, D) Comparison of malate locations derived from the X-ray structure (MLT) and after 6 ns of MD simulation. Water molecules and lipids have been omitted for clarity in all illustrations. Figures were made with Pymol (http://www. sourceforge.net).

Movement into the channels was characterized by direct attraction of malate by the functionally most important part of the channels, the constriction zone (Figs. 1B, 1C, 2), and especially by the cluster of charged arginine side-chains (Arg38, Arg75, Arg133). The trajectory of the exoplasmically attracted malate shows a bend

resulting from an accumulation of basic residues from all monomers near the trimer axis (Fig. 2A). It is important to note that the large number of arginines and lysines present at the exoplasmic and periplasmic vestibules did not significantly slow down or divert the attraction of the substrate molecules. However, the side chains of Arg119/Lys308 and Arg76/Arg329, located close to the constriction zone in the exoplasmic and periplasmic vestibules, respectively, played a role in its desolvation (see below) and may, thus, represent steps of the functional ladder of charged residues.

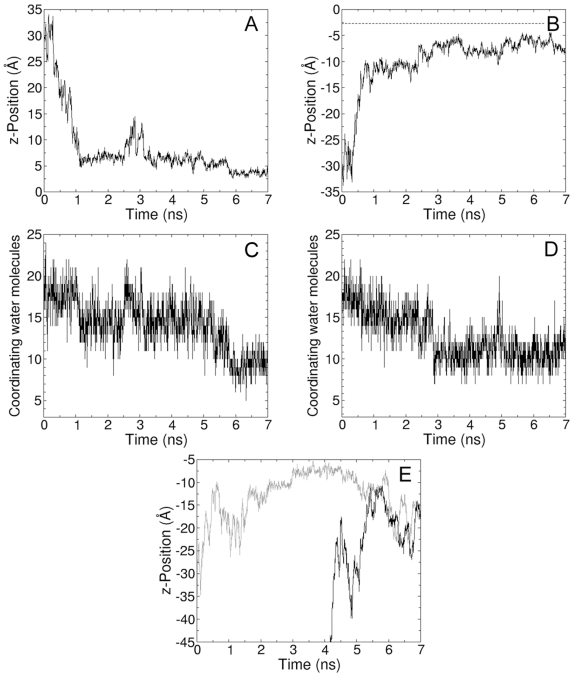


FIGURE 3. Time course of the travel of malate ions MAL1 from the exoplasmic (A, C) and MAL6 from the periplasmic sides (B, D) toward the constriction zone of Omp32 in monomers 1 (A, C), 2 (B, D) and of MAL5 and MAL3 in monomer 3 (E). (A, B) Position of malate along the longitudinal axis of the Omp32 channels (centers of mass) with z = 0 in the constriction zone. The dotted line in (B) indicates the position of L-malate in its binding site derived from the crystal structure. (C, D) Number of water molecules associated with malate molecules. (E) Positions of malate ions MAL5 and MAL3 in the periplasmic half of the porin channel. MAL3 apparently induces release of the bound MAL5.

Monomer 1 – Malate ion MAL1, originally placed on the exoplasmic side (Fig. 3A), was bound to the constriction zone by the interaction between one of its carboxylate groups and the guanidinium group of Arg75. Initial binding to Arg75 was accompanied by the loss of ~4 water molecules from the first solvation shell around the malate ion, reducing it from ~18 to ~14 in the simulation (Fig. 3C). With the exception of a short period of dissociation from this base (t = 2.5 to 3 ns), one carboxylate of MAL1 remained immobilized at this side chain, while substitution for the other carboxylate was possible. The second carboxylate group of MAL1 was largely desolvated (between t = 4.8 and 6.0 ns) by temporary interaction with basic residues of the opposite channel wall, reducing the number of solvating water molecules from ~14 to 9-8 (Fig. 3C). MAL1 was subsequently stably bound, clamped between the side chains of Arg75 and Arg133. This transition is reflected in a step-like reduction of the z-coordinate of MAL1 by about 3 Å toward the channel center (t = 5.7 ns; Fig. 3A).

Monomer 2 – Malate ion MAL6, coming from the periplasmic side, was immobilized after 1.25 ns by binding to the cluster-forming Arg38 of monomer 2 (Figs. 3B, 4D). Desolvation of the anion proceeded in a more continuous way at this site (Fig 3D). Sudden transfer of the substrate from Arg38 to Arg75, however, led to loss of 5 water molecules (t = 2.75 ns). Subsequently, MAL6 was mainly coordinated to the side-chains of Arg75 and Arg133, thereby filling the narrow constriction zone of Omp32.

In analogy to the results of the simulation in monomers 1 and 2, the side-chains of Arg75 and Arg133 were found to be two of the main interaction sites for malate in our crystal structure. After 7 ns of simulated time, MAL1 is also bound to Arg75 and Arg133 (Fig. 4), but remains displaced from the position of the crystallographic binding site by 5-6 Å. Each malate carboxylate binds to one arginine, bridging their side chains. However, this appears to be a direct result of the insufficient sampling time, since the major residues involved in binding have been correctly identified. The difference can be ascribed to a 'memory effect' of the approach process, during which the side chains of Arg75 and Arg133 moved significantly toward the exoplasm to capture the substrate (Fig. 4), while their retraction carrying the substrate into the constriction zone is still incomplete after 7 ns.

The displacement of the binding site of MAL6 approaches only –2.5 Å in z-dimension. As in the crystal, one malate carboxylate binds both side chains of Arg75 and Arg133. In the closest position, the other carboxylate group of MAL6 interacts with the hydroxyl group of Ser 108 instead of that of Thr102 with respect to the X-ray structure. However, MAL6 has D-stereochemistry, as opposed to the L-malate bound in the crystal

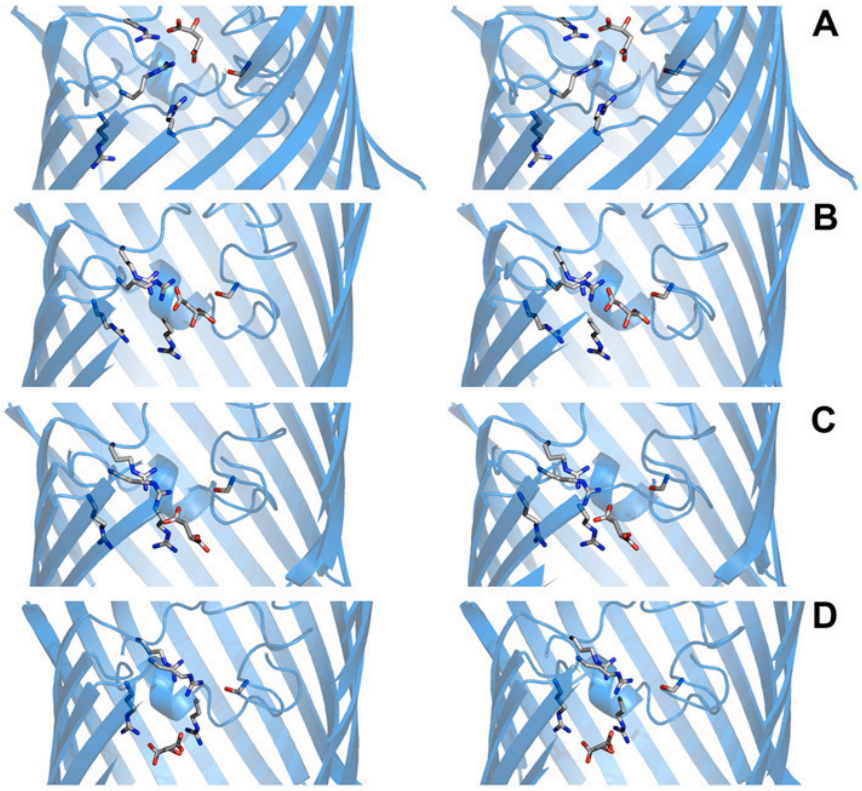


FIGURE 4. Snapshots (stereo views) from the 7 ns MD trajectory showing MAL1, originally in the exoplasmic solvent space, clamped between Arg133 and Arg75 after t = 6 ns (A). MAL6, coming from the periplasmic side, is bound to Arg133, Arg75 and Ser108 after t = 6 ns (B), MAL6 was also found jointly coordinated by Arg75 and Arg329 (C, t = 7 ns) and clamped between Arg38 and Arg76 before it reached the constriction zone (t = 2 ns).

structure. It is interesting to note that *D. acidovorans* preferably consumes L-malate but is able to grow on D-malate as well (Engelhardt, unpublished results). Since binding of D-malate has not been thoroughly tested by crystallography, we can only speculate that this binding mode would be assumed by D-malate in Omp32. However, it seems plausible that the electrostatic attraction of the arginine cluster together with a favorable positioning of serines and threonines on the opposite channel wall allows for D-malate binding.

Overall, the binding partners of MAL1 and MAL6 in the simulation are in favorable agreement with the binding situation found in the crystal structure (Figs. 2C, D). In spite of the remaining differences, this means that the natural substrate of Omp32 was successfully docked near the crystallographically determined binding site by a completely unbiased MD simulation. In explaining the differences, the discrepancy in temperatures between crystallography and simulation, the limited timescale of the simulation and the presence of inorganic ions in the crystal structure have to be taken into account. For instance, the charges on Arg 38 and Arg329 bind sulfate ions next to malate at Arg75 in the crystal structure (Fig. 1F). While one carboxylate group is continuously bound to the guanidinium group of Arg75, the second carboxylate is temporarily attracted by Arg38 and Arg329 in the simulation (Fig. 4C).

Monomer 3 – As observed in monomer 2, a malate ion (MAL5) was bound in monomer 3 by coordination of one of its carboxylate groups to Arg38 (t = 2.2 ns). From that site, it was analogously transferred to the side-chain of the binding-site forming Arg75. A second malate ion, originally placed into the exoplasmic solution, was rapidly attracted toward the periplasmic vestibule of monomer 2, in fact crossing the periodic boundary (t = 4.2 to 5.4 ns; Fig. 3E). It bound to the side chain of Arg329 for ~ 300 ps (from t = 5.5 to 5.8 ns). However, the interaction between the substrate anions did not push the first malate ion further toward the center, but instead liberated both molecules from their protein-bound states, thereby preventing them to reach the binding site. Their motion appeared correlated after t ~ 5.5 ns, and was essentially directed toward the periplasm.

While Arg38, Arg76, and Arg329 formed salt bridges to inorganic anions such as sulfate or chloride in the crystal structure and were thereby fully or partially saturated, the same sites provided temporary binding sites for malate on the

periplasmic side of the constriction zone in the MD simulation. It is conceivable that these residues constitute a pathway of substrate release from the constriction zone into the periplasm. Fig 3E illustrates that detachment of malate from the constriction zone can be observed, effected by interaction with a second malate ion. This indicates that in spite of the higher concentration on that side, malate is capable of leaving its binding site toward the periplasm. This step was facilitated by the higher density of positively charged residues in the periplasmic vestibule. Similar to desolvation, resolvation of malate is more continuous on the periplasmic side (Fig 3D), leading to a smoother energy gradient. Recently, Kosztin and Schulten (50) suggested that asymmetrically shaped energy landscapes in specific channel proteins may lead to an accumulation of substrate molecules by nonequilibrium membrane fluctuations via a ratchetlike mechanism. Omp32 may be representative of such a system, since the shape of the free energy profile (calculated for Cl^-) is reminiscent of a ratchet, exhibiting a much gentler slope from the channel constriction zone toward the periplasm than in the direction of the exoplasm (26), an asymmetry that can also be observed in conductance measurements (24).

Electrophysiological Measurements – Efficient and directed attraction of malate by the electrostatic field and specific association of the substrate in the constriction zone draw the picture of a channel that is adapted for the translocation of such molecules. This is also reflected in conductance measurements with Omp32 inserted in planar lipid membranes. Electrophysiological measurements in KCl showed that Omp32 is a good conductor at low salt concentrations compared to other porins; a feature that we could attribute to Cl^- immobilization inside the channel (26), resulting in a non-linear current-concentration curve. Corresponding experiments with sodium L-malate as an electrolyte revealed kinetics with a strong substrate-binding characteristic (Fig. 5). Taking into account that cations may still contribute to the recorded current in an almost linear and concentration-dependent manner, the saturation effect for malate ions is expected to be even stronger. Estimates of half-saturation by L-malate are clearly below 10 mM, which is in the physiological concentration range and could also be expected regarding the affinity of maltose binding in maltoporin (51, 52). It is critical to determine the binding constant accurately by means of conductance measurements in a straightforward manner. Since malate is ionic and contributes to the current and Cl^- (particularly at higher concentrations) interferes with the amino acid residues being part of the malate binding site (see above) blocking experiments, according to those performed with uncharged oligosaccharides and maltoporin in KCl solutions (e.g. 51), will be biased.

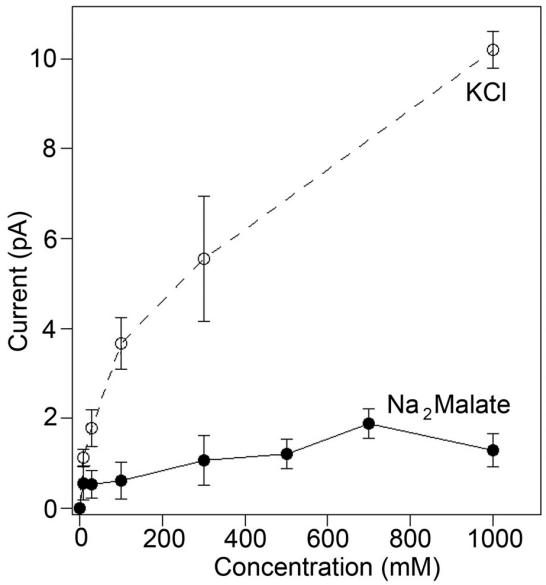


FIGURE 5. Ion currents through Omp32 trimers at 20 mV. The data were recorded in solutions of KCl (open symbols) or Na_2-L-malate (filled symbols) and determined from current steps occurring upon insertion of porin molecules into the lipid bilayer. The plotted data are means and standard deviations from Gaussians fitted to the histograms of current steps. The experimental curves illustrate moderate interaction with Cl^- and strong malate binding in the Omp32 channels. The KCl curve is derived from (24).

However, why is malate bound at all inside the channel? Berezhkovskii and Bezrukov (19) showed in a recent study that a substrate binding site in porins is advantageous for translocation because the increased rate of attraction due to an energy minimum may outweigh the reduced efflux rate. They found that substrate-specific outer membrane channels exhibit an optimized balance between substrate binding and escape from the channel, resulting in overall increased flux rates of the solutes. Porin specificity for malate and an optimized flux are indeed of importance with respect to the physiological needs of *D. acidovorans*; this bacterium cleaves naturally occurring poly(β-L-malic acid) by an outer-membrane bound enzyme and consumes malate as sole carbon source for growth (16).

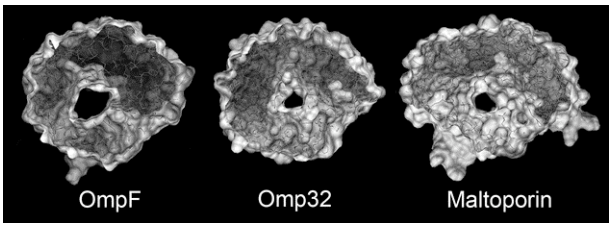


FIGURE. 6. Comparison of the smallest cross sections in the constriction zone of the 16-stranded porins OmpF from *E. coli* (PDB entry 2OMF) and Omp32 from *D. acidovorans* (1E54), and of the 18-stranded sugar-specific maltoporin from *E. coli* (1MAL). The images show the "inside" of the porin surface from the position of the smallest pore cross section toward the perplasmic side of the monomers. The figures were generated with DINO.

CONCLUSIONS

Substrate specificity has been attributed to the trimeric 18-stranded porins while the 16-stranded ones are currently regarded as being unspecific in general. Our study challenges this classification since the 16-stranded porin Omp32 possesses binding capacity for malate that is utilized by the bacterium. Substrate binding is mediated by amino acid residues located on individual loops inside the channel in both porin types. We therefore suggest uncoupling the structural classification of porins from their functional potential. Instead, the minimal cross section of the channel, mainly formed by the porin-specific loops and protrusions, offer a much better indicator for efficient substrate-protein interactions. Fig. 6 illustrates that the minimal cross section of the Omp32 pore constriction is the smallest one among the trimeric porins including the sugar-specific maltoporin (and the sucrose porin ScrY), thus, reflecting the adaptation to a smaller substrate. This criterion also applies for other outer membrane channels, e.g., the siderophore-binding proteins, possessing 22 β-strands and a large plug inside the barrel (53), and to the 12- and 14-stranded transporters Tsx and FadL, being specific for nucleosides and long-chain fatty acids, respectively (17, 18). Substrate-binding is obviously independent of the barrel size and was aquired by very different outer membrane proteins in the course of optimizing substrate selection and translocation (19). Our findings suggest to look for other (16-stranded) porins exhibiting substrate specificity in bacteria with particular nutrient requirements or adaptation to substances at low concentration in the organisms´ natural environment.

Acknowledgments – We thank our colleague Erik Roth for a sample of recombinant Omp32, Johann Lubieniecki for skillful technical assistance, Marc Fuhrmans for preparation of crystals, the staff of the EMBL Grenoble outstation and ESRF for support and operation of the JSBG beamlines, Bert de Groot for critical reading of the manuscript, Wolfgang Baumeister and Dieter Oesterhelt for generous support. This work was supported by a grant from the DFG (En 144/3).

Footnote: The atomic coordinates and structure factors (codes 2FGQ and 2FGR) have been deposited in the Protein Data Bank.

The abbreviations used are: MLT, crystal structure of Omp32 containing malate; NAT, crystal structure of Omp32; PAP, porin-associated protein.

REFERENCES

1. Nikaido, H. (2003) *Microbiol. Mol. Biol. Rev.* **67,** 593-656
2. Koebnik, R., Locher, K. P., and van Gelder, P. (2000) *Mol. Microbiol.* **37,** 239-253
3. Schulz, G. E. (2003) *Adv. Protein Chem.* **63,** 47-70
4. Schirmer, T. (1998) *J. Struct. Biol.* **121,** 101-109
5. Schirmer, T., Keller, T. A., Wang, Y. F., and Rosenbusch, J. P. (1995) *Science* **267,** 512-514
6. Forst, D., Welte, W., Wacker, T., and Diederichs, K. (1998) *Nat. Struct. Biol.* **5,** 37-46
7. Weiss, M. S., Abele, U., Weckesser, J., Welte, W., Schlitz, E., and Schulz, G.E. (1991) *Science* **254,** 1627-1630
8. Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius, J. N., and Rosenbusch, J. P. (1992) *Nature* **358,** 727-733
9. Kreusch, A., and Schulz, G. E. (1994) *J. Mol. Biol.* **243,** 891-905
10. Hirsch, A., Breed, J., Saxena, K., Richter, O. M., Ludwig, B., Diederichs, K., and Welte, W. (1997) *FEBS Lett.* **404,** 208-210
11. Dutzler, R., Rummel, G., Alberti, S., Hernandez-Alles, S., Phale, P., Rosenbusch, J., Benedi, V., and Schirmer, T. (1999) *Structure Fold. Des.* **7,** 425-434
12. Engelhardt, H., Gerbl-Rieger, S., Krezmar, D., Schneider-Voss, S., Engel, A., and Baumeister, W. (1990) *J. Struct. Biol.* **105,** 92-102
13. Baldermann, C., Lupas, A., Lubieniecki, J., and Engelhardt H. (1998) *J. Bacteriol.* **180,** 3741-3749
14. Horowitz, H., Gilroy, Sh., Feinstein, S., and Gilardi, G. (1990) *J. Clin. Microbiol.* 28, 143-145
15. Willems, A., de Vos, P., and de Ley, J. (1992) in *The Prokaryotes Vol. III*, eds. Balows, A., Trüper, H. G., Dworkin, M., Harder, W., & Schleifer, K. H. (Springer, New York), pp. 2583-2590
16. Gödde, C., Liebergesell, M., and Steinbüchel, A. (1999*) FEMS Microbiol. Lett.* **173**, 365-372
17. Ye, J., and van den Berg, B. (2004) *EMBO J.* **23,** 3187-3195
18. van den Berg, B., Black, P. N., Clemons, W. M., Jr., and Rapoport, T. A. (2004) *Science* **304,** 1506-1509

19. Berezkhovskii, A. M., and Bezrukov, S. M. (2005) *Biophys. J.* **88,** L17-L19
20. Dutzler, R., Schirmer, T., Karplus, M., and Fischer, S. (2002) *Structure (Camb.)* **10,** 1273-1284
21. Robertson, K. M., and Tieleman, D. P. (2002) *FEBS Lett.* **528,** 53-57
22. Nestorovich, E. M., Danelon, C., Winterhalter, M., and Bezrukov, S. M. (2002) *Proc. Natl. Acad. Sci.* USA **99,** 9789-9794
23. Ceccarelli, M., Danelon, C., Laio, A., and Parrinello, M. (2004) *Biophys. J.* **87,** 58-64
24. Mathes, A., and Engelhardt, H. (1998) *Biophys. J.* **75,** 1255-1262
25. Zeth, K., Diederichs, K., Welte, W., and Engelhardt, H. (2000) *Structure Fold. Des.* **8,** 981-992
26. Zachariae, U., Helms, V., and Engelhardt, H. (2003) *Biophys. J.* **85,** 954-962
27. Suenaga, A., Komeiji, Y., Uebayasi, M., Meguro, T., Saito, M., and Yamato, I. (1998) *Biosci. Rep.* **18,** 39-48
28. Danelon, C., Suenaga, A., Winterhalter, M., and Yamato, I. (2003) *Biophys. Chem.* **104,** 591-603
29. Kabsch, W. (1988) *J. Appl. Cryst.* **21,** 67-71
30. Vagin, A., and Teplyakov, A. (2000) *Acta Crystallogr.* D **56,** 1622-1624
31. Collaborative Computational Project, No. 4 (1994) *Acta Crystallogr*. D **50,** 760-763
32. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr*. A **47,** 110-119
33. Morris, R. J., Perrakis, A., and Lamzin, V. S. (2003) *Methods Enzymol.* **374,** 229-244
34. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Cryst.* **26,** 283-291
35. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491-497
36. Woolf, T. B., and Roux, B. (1994) *Proc. Natl. Acad. Sci.* USA **91,** 11631-11635
37. Woolf, T. B., and Roux, B. (1996) *Proteins* **24,** 92-114
38. Im, W., and Roux, B. (2002) *J. Mol Biol.* **319,** 1177-1197
39. Im, W., and Roux, B. (2002) *J. Mol. Biol.* **322,** 851-869
40. Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R., and Klein, M. (1983) *J. Chem. Phys.* **79,** 926-935
41. Brooks, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem.* **4**, 187-217
42. MacKerell, A. D., Bashford, D., Bellot, M., Dunbrack, R. L., Evanseck, J. D., Field., M. J., Fischer, S., Gao, J., Guo, H., Ha, S., McCarthy, D. J., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E. III, Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) *J. Phys. Chem.* B **102,** 3586-3616
43. Ryckaert, J., Ciccotti, G., and Berendsen, H. (1977) *J. Comput. Chem.* **23,** 327-341
44. Essmann, U., Perera, L., Berkowitz, M., Darden, T., Lee, H., and Pedersen, L. (1995) *J. Chem. Phys.* **103,** 8577-8593
45. Grizot, S., and Buchanan, S. K. (2004) *Mol. Microbiol.* **51**, 1027-1037
46. Prinz, T., and Tommassen, J. (2000) *FEMS Microbiol. Lett.* **183**, 49-53
47. Meyer, J. E. W., Hofnung, M., and Schulz, G. E. (1997) *J. Mol. Biol.* **266**, 761-775
48. Zachariae, U., Koumanov, A., Engelhardt, H., and Karshikoff, A. (2002) *Protein Sci.* **11,** 1309-1319
49. Domene, C., Bond, P. J., and Sansom, M .S. P. (2003) *Adv. Prot. Chem.* **66,** 159-193
50. Kosztin, I., and Schulten, K. (2004) *Phys. Rev. Lett.* **93,** 238102
51. Winterhalter, M. (1999) *Coll. Surfaces* **A 149**, 547-551
52. Jordy, M., Andersen, C., Schülein, K., Ferenci, T., and Benz, R. (1996) **259**, 666-678
53. Buchanan, S. K., Smith, B. S., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., van der Helm, D., and Deisenhofer, J. (1999) *Nat. Struct. Biol.* **6,** 56-63

Supplemental Information

Supplemental Methods – Purification of Omp32

Cells of *D. acidovorans* strain JL0 were grown under conditions described elsewhere (1). Cells were harvested by centrifugation, resuspended in buffer A (100 mM NaCl, 100 mM KCl, 3 mM $MgCl_2$ and 10 mM Hepes, pH 7) and broken by *french press*. The suspension was centrifuged at 75000 *g* for 1h to isolate the bacterial membranes. To remove inner membranes the pellet was homogenized and resuspended using buffer A containing 3% octyl-polyoxyethylene (o-POE). The sample was stirred for 1 h at room temperature and then spinned down for 30 min at 75 000 *g*. This step was repeated twice. A third resuspension step was carried out in the presence of lysozyme at a concentration of 1 mg/ml and the extract was incubated for 3h at 30°C. The material was centrifuged at 75000 *g* and the clear supernatant added to a 10 ml MonoQ column (MonoQ 16/10, Amersham), which was equilibrated with buffer A containing 0.5% o-POE. The column was washed with ten column volumes of the same buffer before the protein was eluted with a linear gradient of the buffer A solution containing 0.5% o-POE and 500 mM NaCl for 20 column volumes. Protein fractions predominantly containing Omp32 were collected, concentrated to 20 mg/ml, and applied to size exclusion chromatography (Superdex 200, 16/60; Amersham) in buffer A and 0.5% o-POE. The purity of the protein was determined by SDS-PAGE.

1. Baldermann, C., Lupas, A., Lubieniecki, J., and Engelhardt H. (1998) *J. Bacteriol.* **180,** 3741-3749

Supplemental Data: Table I Data collection and refinement statistics

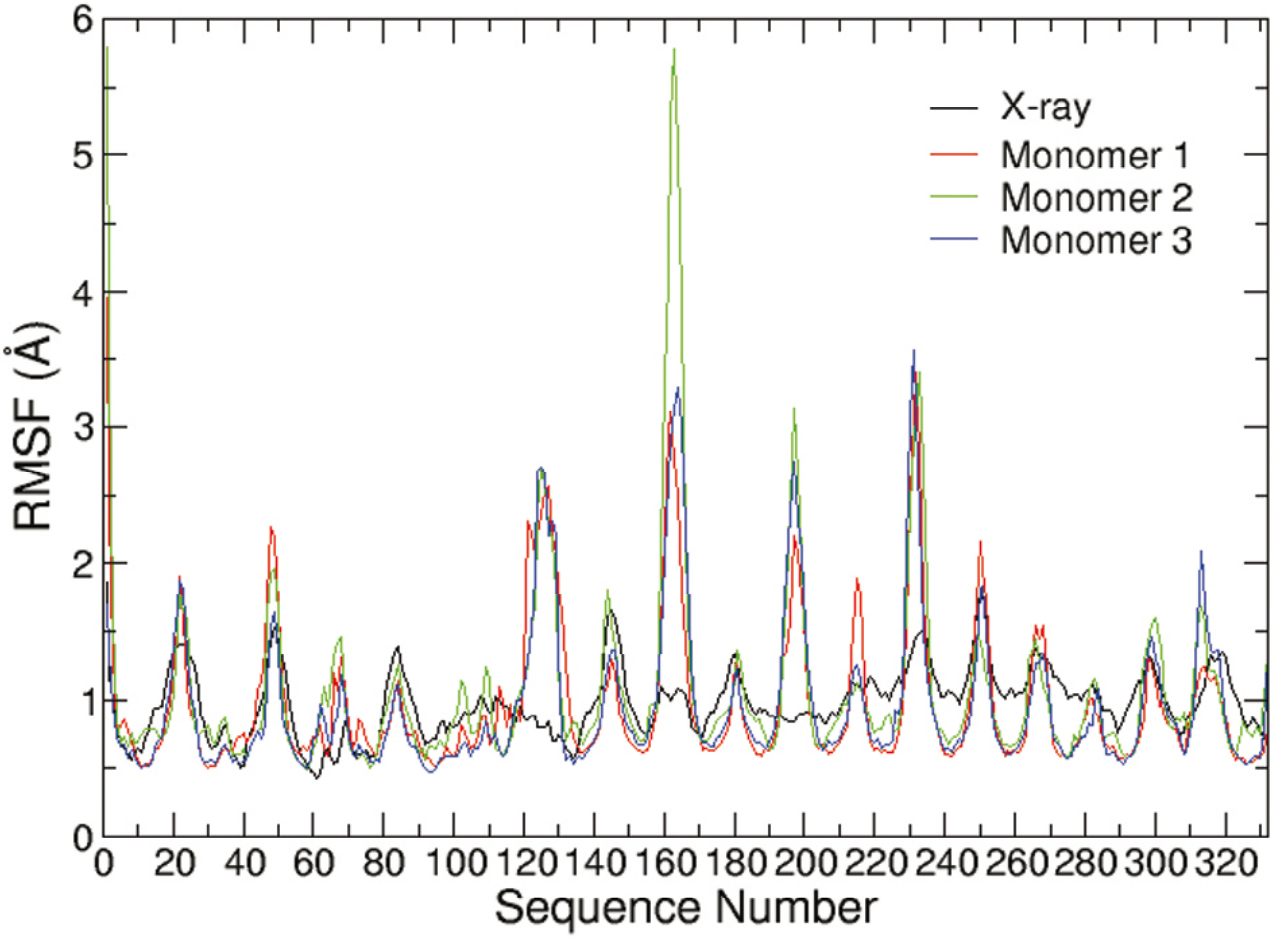
Dataset	MLT	NAT
Data collection		
X-ray Source/detector system	ID14-EH4/ADSC Quantum 4R	ID14-EH4/ADSC Quantum 4R
Wavelength [Å]	1.0053	1.0053
Space group	P63	P63
Cell constants [Å]	a=b=106.21, c=93.21	a=b=106.41, c=93.79
Resolution range [Å]	20-1.45 (1.54-1.45)	20-1.5 (1.59-1.5)
No. unique reflections	105491 (16363)	96292 (15412)
Redundancy	2.7	2.7
Completeness [%]	98.1 (94.2)	96.1 (87.6)
$^{a}R_{merge}$ [%]	6.9 (44.9)	8.3 (61.3)
$I/\sigma(I)$	8.66 (2.38)	7.59 (1.78)
Refinement		
Resolution range [Å]	15-1.45 (1.45-1.48)	20-1.5 (1.54-1.5)
No. of unique reflections	98333 (5173)	87899 (4618)
$R^{b}_{cryst}/R^{c}_{free}$ [%]	15.0/17.0 (27.2/28.1)	15.0/17.9 (29.2/33.0)
No. of protein atoms	2448	2527
No. of waters	377	409
No. of ligand atoms (MLT/BOG)	9/20	-/-
No. of ions (SO_4^{2-}/Ca^{2+})	2/3	2/3
r.m.s.d. of bond length [Å^2]	0.017	0.021
r.m.s.d. of bond angles [degrees]	1.7	1.7
Ave. *B* factor of protein atoms [Å^2]	14.8	16.4
Ramachandran plot statistics		
residues in most favoured region [%]	89.6	90.9
residues in add. allowed regions [%]	10.0	8.3
residues in gen. allowed regions [%]	0	0.7
residues in disallowed regions [%]	0.4	0

Values in parentheses refer to the highest resolution shell

[a] $R_{merge} = \sum_{\text{unique reflections}}(\sum^{N}_{i=1} |I_i - \_ |)/\sum_{\text{unique reflections}}(\sum^{N}_{i=1} I_i)$, where *N* represents the number of equivalent reflections and *I* the measured intensity.

[b] $R_{cryst} = \sum |F_{obs} - F_{calc}|/\sum|F_{obs}|$.

[c] R_{free} was calculated using 5% randomly selceted reflections.



Supplemental Figure RMS fluctuations of the backbone atoms, averaged over each residue, calculated from the experimental B-factors (PDB 1E54) and from the MD simulation. Similar to the study on OmpF by (1), a uniform offset constant of 0.28 Å was subtracted from the crystallographic data to account for the presence of static disorder. The high RMSF value of the N-terminus is reflected by its missing experimental electron density in the MLT structure.

The discrepancies in the extracellular loops L4 (residues 158-169), L5 (195-200), and L6 are especially pronounced and can be ascribed to relief from crystal constraints in this region and the higher temperature used in the simulation. The difference in protrusion 2 (residues 118-130) is probably directly related to substrate-protein interactions. The protrusion is part of the channel-lining loop L3, which is involved in binding malate ions.

1. Im, W., and Roux, B. (2002) *J. Mol Biol.* **319,** 1177-1197