Extending the Structure of an ABC Transporter to Atomic Resolution: Modeling and Simulation Studies of MsbA†

Jeff D. Campbell, Philip C. Biggin, Marc Baaden,‡ and Mark S. P. Sansom*

Laboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, United Kingdom

Received December 10, 2002; Revised Manuscript Received February 12, 2003

ABSTRACT: Molecular modeling and simulation approaches have been use to generate a complete model of the prokaryotic ABC transporter MsbA from *Escherichia coli*, starting from the low-resolution structure-based Cα trace (PDB code 1JSQ). MsbA is of some biomedical interest as it is homologous to mammalian transporters such as P-glycoprotein and TAP. The quality of the MsbA model is assessed using a combination of molecular dynamics simulations and static structural analysis. These results suggest that the approach adopted for MsbA may be of general utility for generating all atom models from low-resolution crystal structures of membrane proteins. Molecular dynamics simulations of the MsbA model inserted in a fully solvated octane slab (a membrane mimetic environment) reveal that while the monomer is relatively stable, the dimer is unstable and undergoes significant conformational drift on a nanosecond time scale. This suggests that the MsbA crystal dimer may not correspond to the MsbA dimer in vivo. An alternative model of the dimer is discussed in the context of available experimental data.

ATP binding cassette (ABC) transporters are a diverse membrane protein family of some biomedical importance. In *Escherichia coli*, ABC transporters are encoded by almost 5% of the entire genome (*1*). ABC transporters can mediate the translocation of a variety of solutes (also referred to as allocrites) across cell membranes. These solutes include polyanionic lipids, amino acids, peptides, nucleotides, vitamins, complex drug molecules, and inorganic ions. In particular, ABC transporters are one of the major families of microbial multidrug efflux systems. Other families include the RND superfamily, the structure of one member of which, the AcrB protein, has recently been determined by X-ray crystallography (*2*).

Despite the varied composition and structure of the allocrites, and the diverse physiological roles of their transporters, several fundamental features are conserved across the ABC superfamily. All ABC transporters draw their energy from binding and hydrolysis of ATP, enabling them to move the allocrite across the cell membrane. ABC transporters possess two nucleotide-binding domains (NBD) that are located in the cytoplasm. The NBDs from different members of the family share a high sequence and structure conservation (*3*−*9*). Both NBD domains are arranged in a close association with one another. This is supported by experiments showing that communication between the NBD monomers is critically important for ATP binding and hydrolysis (reviewed in ref *10*). In addition to the NBDs, every ABC transporter dimer has two transmembrane domains (TMDs). The two recently determined ABC transporter TMD structures (the lipid A exporter MsbA from *E. coli*; PDB code 1JSQ (*11*) and the vitamin B_{12} importer BtuCD from *E. coli*; PDB code 1L7V (*9*)) and multiple sequence alignments with other transporters demonstrate that greater structural and sequence diversity exists in the TMDs than in the NBDs. The transmembrane part consists of a helical bundle with the most common arrangement being six helices per monomer (or per domain when the entire transporter is coded as a single polypeptide), as found in MsbA, the eukaryotic drug exporter P-glycoprotein (P-gp), and the eukaryotic chloride transporter CFTR. However, transporters with 10−12 helices per monomer have been found as, for example, BtuCD and the eukaryotic antigen transporter TAP.

To fully understand the transport cycle of a given medically significant ABC transporter, at least one intact structure of the protein at atomic detail is required, but it might even be necessary to characterize several intermediate states. However, the difficulties associated with expressing (*12*) and crystallizing membrane proteins are well-established (*13*), and to date only a few (∼60) membrane protein structures have been solved (see http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html for a summary), of which only two are bacterial ABC transporters. We therefore still eagerly await the experimental determination of a mammalian ABC transporter structure. A possible alternative approach is to employ homology modeling of a mammalian membrane protein based on a bacterial template of known structure. The two crystal structures mentioned above—MsbA and BtuCD—might be used as templates in homology modeling of mammalian ABC transporters. Unfortunately, both the low resolution (0.45 nm) and high number of missing atoms (all side chains and ca. 23% of residues in their entirety) preclude the X-ray structure of MsbA (a Cα trace only) being used directly as template for homology modeling. The BtuCD

† This research was supported by grants from the Wellcome Trust. M.B. thanks the EC for a grant (MCFI-2000-01210).

* To whom correspondence should be addressed. E-mail: mark@biop.ox.ac.uk. Tel: +44-1865-275371. Fax: +44-1865-275182.

‡ Current address: Institut de Biologie Physico-Chimique, Paris, France.

10.1021/bi027337t CCC: $25.00
Published on Web 03/13/2003

structure is higher resolution (0.35 nm) and is only missing a small number of atoms, but unfortunately, it has low sequence identity with many mammalian transporters. An important aspect of a homology modeling study is the sequence alignment (*14*). For the BtuCD TMD, it is rather difficult to construct a convincing alignment with mammalian ABC transporters such as P-gp, CFTR, or TAP.

In contrast to BtuCD, the transmembrane domains of MsbA are more closely related (by sequence homology) to P-gp than those from any other bacterial transporter (*11*). MsbA also aligns with TAP, even though the latter has additional (predicted) TM helices. Therefore, given the biomedical importance of mammalian ABC transporters, a more complete model of the MsbA structure represents an important first step toward their understanding.

In this paper, the 0.45 nm resolution MsbA Cα trace has been extended to generate a complete model of the MsbA monomer using a combination of homology modeling and molecular dynamics (MD) simulation techniques. We employ MD simulations to investigate the conformational dynamics of MsbA in a membrane-like environment. We also propose an alternative model to the X-ray structure for the MsbA dimer.

MATERIALS AND METHODS

General Modeling and Simulation Methods. Homology models were generated using MODELLER 6v2 (*15*). Side chain conformations were optimized using SCAP (*16*). Model quality was assessed using WHAT-IF (*17*) and PROCHECK (*18*). Simulations were run using the GROMACS v3.0 (www.gromacs.org) molecular dynamics simulation package (*19*) with parameter set ffG43a2, a modified GROMOS96 force field.

Simulation systems were generated by (i) protein insertion into an octane slab (see ref *20* for details); (ii) solvation of the protein/octane system using a preequilibrated box of SPC water molecules (*21*, *22*); and (iii) addition of randomly positioned Na^+ and Cl^- ions equivalent to 0.1 mM NaCl. The system was subjected to 100 steps of steepest descents energy minimization.

The protein coordinates of the resulting systems were then restrained while allowing solvent molecules to relax their positions and optimize interactions with the proteins during a 500 ps simulation. During this non-hydrogen protein, atoms were restrained harmonically using a force constant of 1000 kJ mol^{-1} nm^{-2}.

All simulations were performed in the NPT ensemble at 300 K. Long-range electrostatics were treated via a cutoff of 1.7 nm and van der Waals interactions with a 1.0 nm cutoff. The temperatures of the protein, octane, and solvent (both water molecules and ions) were coupled separately, using the Berendsen thermostat (*23*) with a coupling constant of $\tau_T = 0.1$ ps. The pressure was controlled using the Berendsen algorithm at 1 bar with a coupling constant $\tau_P = 1$ ps, using a uniform compressibility of 4.5×10^{-5} bar^{-1}. The integration time step was 2 fs, and coordinates and velocities were saved every 5 ps. The LINCS algorithm was used to restrain all bond lengths (*24*).

The ionization states of MsbA residues located in the hydrophobic core of the bilayer (i.e., K49, D41, R78, R183, R296) were taken as neutral because they were not deemed (by visual inspection) to be forming any salt bridges and were not found at inter-dimer or helix interfaces. The positioning of the octane slab was guided by aligning the bands of Trp residues in MsbA with the membrane/water interfaces.

Analysis. Simulations were analyzed using GROMACS routines and/or locally written code. Secondary structure was analyzed using DSSP (*25*). Molecular graphics diagrams were generated using VMD (*26*) and PovRay (http://www.povray.org).

RESULTS

The MsbA model was generated in three stages (Figure 1): (i) building the transmembrane domain (TMD; residues 1−343); (ii) modeling the nucleotide binding domain (NBD; residues 344−582); and (iii) assembly of the intact MsbA monomer from the TMD and NBD.

Transmembrane Domain. The MsbA TMD is comprised of six long α-helices that protrude into the cytoplasm for a short distance before making contacts with the NBD. The X-ray structure of MsbA (PDB code 1JSQ) consists of an incomplete Cα trace, based on the 0.45 nm resolution structure (*11*). Thus, no side chain or backbone atoms (other than Cα) are present. Also, a total of 39 entire residues (i.e., Cα atoms) are missing from the TMD crystal structure: nine from the N-terminus (residues 1−9) and 30 from the putative intracellular domain (residues 208−237).

The nine N-terminal residues were modeled as an α-helix continuous with the helix made up by residues 10−17. The resulting single long amphipathic helix from residues 1−17 thus sits parallel to the bilayer plane. The secondary structure of the 30 remaining missing TMD residues (Ala 207 to Arg 238) was predicted using QUANTA (http://www.accelrys.com/quanta/) and PSIPRED (*27*), and the residues were modeled as a short helix−loop−short helix motif. The fact that these residues were not resolved in the X-ray structure suggests that this region might be quite mobile.

Having defined Cα coordinates for all of the TMD (i.e., residues 1−343), other missing backbone atoms were generated using MAXSPROUT (*28*) and missing side chain atoms were added using MAXSPROUT and MODELLER, using the latter to select side chain rotamers.

Modeling the Nucleotide Binding Domain. At the time of writing, six high atomic reolution ABC transporter nucleotide binding domain structures (*3*−*6*, *8*, *9*) have been released. Only one of these, BtuCD (*9*), has been solved with its constituent transmembrane domain. All of the elucidated NBDs share a highly conserved topology, whereas the transmembrane domains of MsbA (*11*) and BtuCD (*9*) show marked differences. Multiple sequence alignments for the NBDs show that there is ∼30% sequence identity across the ABC transporter family, which allowed us to employ homology-modeling techniques (*29*).

Only 60% of the MsbA NBD Cα trace is resolved in the crystal structure, namely residues 343−582. The coordinates for 97 out of 240 residues were not determined, and much of the resolved structure does not superimpose well with the corresponding regions in other NBDs. The first missing region is for residues 344−418 of MsbA. It corresponds to the highly conserved β-subdomain and part of the F1-type binding core, as denoted by ref *8*. This region contains three

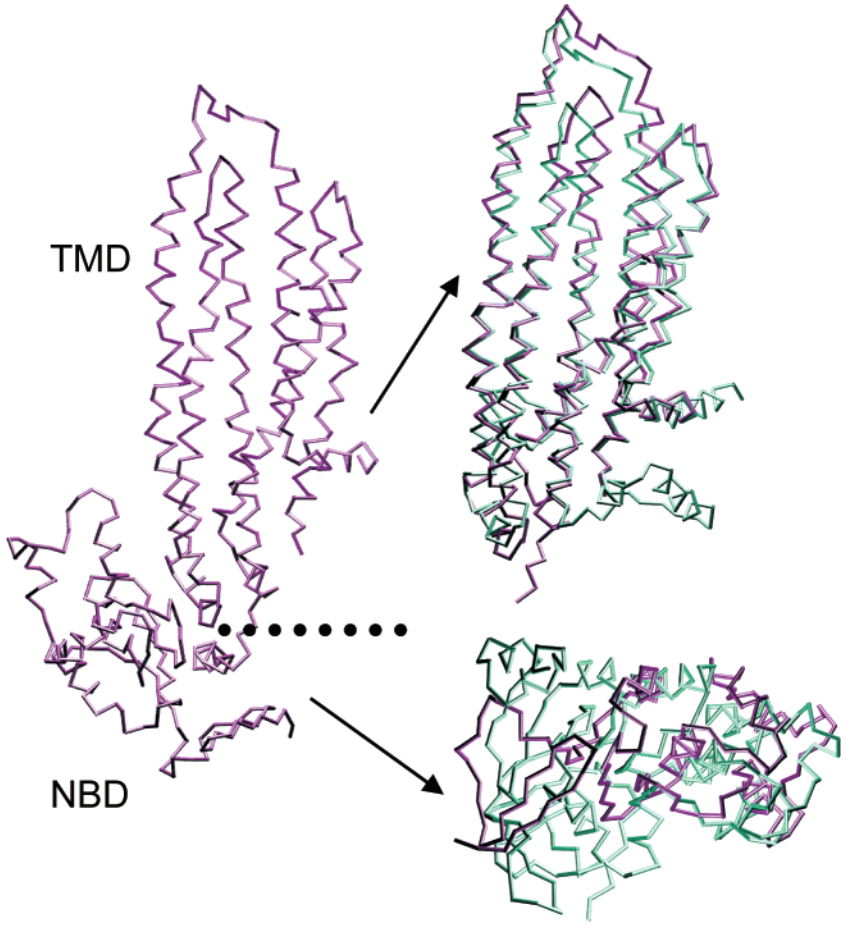

FIGURE 1: Modeling MsbA. The experimental Cα trace (PDB code 1JSQ) is shown on the left in purple. The modeled structure is shown on the right in cyan, superimposed on the corresponding fragment (purple) of the experimental structure. The dotted horizontal line indicates the division between the transmembrane domain (TMD, residues 1−343) and the nucleotide binding domain (NBD, residues 344−582).

antiparallel β-strands and the ATP binding helix. Also, within the first missing region is the functionally important Walker A motif (*30*−*33*), which binds the phosphate groups of ATP (and is N-terminal to the ATP binding helix). Since the root-mean-square deviation (RMSD) between any two NBD β-subdomains or F1-type binding cores is approximately 0.15 nm, we can assume that MsbA shares a similar topology in this region.

Following the β-subdomain is the ABC specific α-helical subdomain (residues 431−479). It is thought to be one of the important communication points between the transmembrane and the nucleotide binding domains (*6*, *8*). The α-helical subdomain is flanked at its N-terminus by an invariant Gln residue thought to bind the γ-phosphate of ATP and at its C-terminus by the ABC transporter signature sequence. The Cα coordinates for the MsbA α-helical subdomain are resolved in the X-ray structure, but they show little similarity with other NBDs. The closest fit for a superposition is 0.5 nm over 50 Cα atoms between the HisP and the MsbA α-helical subdomains. This is in contrast to an RMSD of ∼0.14 nm between any two other higher resolution NBD α-helical subdomains.

The best structural agreement between MsbA and the other existing NBDs is in the first part of the F1-type subdomain, which is only partially resolved in MsbA. These residues (419−424 and 483−527) have a least squares deviation of ∼0.3 nm from the corresponding residues in other NBDs. The C-terminal region of the MsbA F1-type core, which is missing 17 residues, bears no resemblance to the corresponding domain in any other NBD. The residues that are resolved in the X-ray structure (residues 548−564) appear to be the last two strands in the F1-type core. These strands exist in all of the other NBDs, but in MsbA their positions relative to the rest of the NBD are different. Although this region does show the greatest amount of variability, it seems unlikely that the marked difference between the positions of these strands in MsbA and other NBDs can be attributed to species/transporter specificity.

The method employed for modeling the full MsbA transporter involved first constructing a homology model of the entire NBD and then fitting this domain onto the transmembrane domain by least-squares superposition onto the existing MsbA NBD F1-type core Cα coordinates. The four highest resolution NBD structures (HisP (*5*), MalK (*3*), MJ1267 ADP-bound structure (*6*), and TAP1 (*4*)) were selected as templates for the MsbA homology models. The sequences were aligned using ClustalW (*34*), and the four templates were input into MODELLER.

For each template, 100 models were generated of which the one with the lowest MODELLER energy was selected. This lead to four MsbA NBD homology models from which to choose. A number of selection factors were considered to determine the most appropriate model: (i) template sequence homology to MsbA; (ii) template RMSD to available MsbA coordinates; (iii) model drift (stability) in MD simulations (see below); and (iv) template resolution. A summary of the

Table 1: MsbA NBD Homology Model Evaluation

template	X-ray resolution (nm)	Cα RMSD to MsbA[a] (nm)	% sequence identity	Cα RMSD over 2 ns MD (nm)
HisP	0.15	0.56	21	0.36
MJ1267	0.16	0.73	17	0.45
MalK	0.19	0.49	20	0.58
TAP1	0.24	0.77	40	0.36

[a] Fit and RMSD from Cα atoms of residues V419−V534 of MsbA.

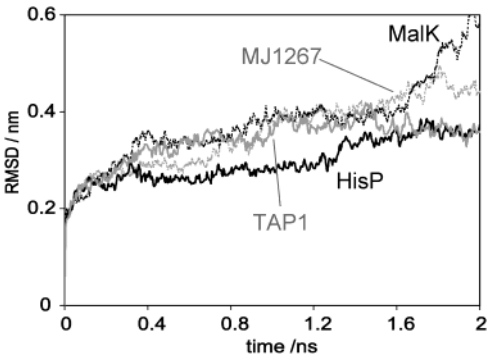

FIGURE 2: NBD structural drift**.** All residue Cα RMSDs vs time are shown for the four homology models of the MsbA NBD, derived from the structures of HisP (solid black line); TAP1 (solid gray line); MalK (dotted black line); and MJ1267 (dotted gray line).

four NBD models with respect to these factors is shown in Table 1. The RMS fit of homology models to the X-ray Cα coordinates of MsbA was done disregarding the C-terminus because this section in MsbA is radically different from all of the other NBDs. TAP1 bears the highest sequence identity to MsbA over the aligned NBDs. Surprisingly, the model based on TAP1 does not superimpose well with MsbA.

The final column in Table 1 represents the RMSD of each homology model from its starting coordinates (i.e., structural drift) after 2 ns of MD. As can be seen from Figure 2, at ∼1.0 ns into the simulation all of the models display a similar degree of drift. Between 1 and 2 ns the less stable models begin to unfold. It should be noted that the high RMSDs for MJ1267 and MalK are due to protein unfolding rather than a more restricted conformational change (determined by visual examination of the final structure). Because of its high stability in MD and its structural homology to the X-ray structure of MsbA, the HisP-based model of the MsbA NBD was used to generate the complete MsbA monomer and dimer models.

Final Model Assembly. First, an intact MsbA monomer model was assembled. The homology modeled HisP-based NBD was attached to the TMD part by least squares superposition of the available F1-type subdomain coordinates of the X-ray structure of MsbA (419−424 and 483−527) with the NBD. This fit placed the α-helical subdomain within reasonable agreement to the MsbA X-ray α-helical subdomain. Consequently, the linker region of MsbA was too distant from the NBD. The corresponding linker residues (330−342) join the TMD to the NBD and are built as helix residues in the X-ray structure. After fitting, the linker helix crossed through the ATP-binding helix immediately following the Walker A loops. The linker helix was therefore displaced manually such that it did not cause steric conflicts and would not interfere with ATP binding.

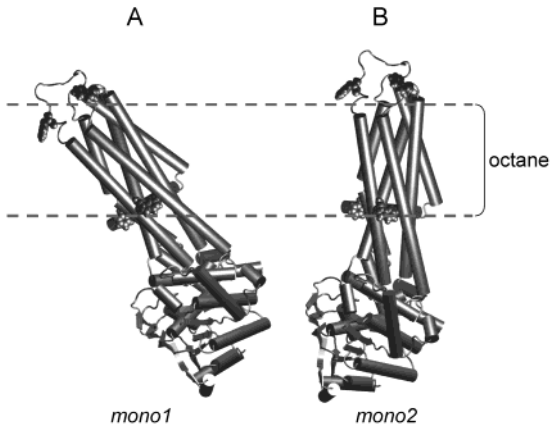

FIGURE 3: Simulation setups. The starting orientations of the MsbA monomer in simulations *mono1* (tilt angle derived from the crystallographic dimer) and *mono2* (tilt angle set to ca. 0°) are shown in panels A and B, respectively. The two bands of Trp residues (shown is space−fill format) were used to position the monomers relative to the octane slab (indicated by the broken horizontal line).

A MsbA dimer was assembled by least-squares fitting the monomer model onto the TMDs in the X-ray dimer. Local steric conflicts resulting from the dimer-fitting procedure were reduced by energy minimization using GROMACS v3.0 with the method of steepest-descents. Both PROCHECK and WHATIF indicated that 92% of the residues of the MsbA model had backbone torsion angles in the favored region of a Ramachandran plot, as might be expected for such a modeling strategy (*35*). All other protein-health characteristics were within acceptable ranges.

MD Simulations. In an attempt to explore the conformational stability of the resultant monomeric MsbA model, multinansecond MD simulations were performed with the monomer embedded in a membrane-mimetic octane slab, fully solvated on either side. Two possible initial orientations of the monomeric MsbA model with respect to the slab were explored (Figure 3). The first simulation (*mono1*—see Table 2) started from the monomer oriented with respect to the octane slab at an angle similar to that derived from inserting the crystallographic MsbA dimer in a membrane with the dimer axis parallel to the membrane normal. An alternative orientation was explored (*mono2*) in which the monomer was rotated 25° such that the TMD helices ran approximately parallel to the membrane normal (i.e., the *z* axis). A simulation for the full MsbA dimer embedded in an octane slab with the dimer axis parallel to *z* was also run (*dim1*).

A simple measure of overall protein stability can be obtained by plotting the RMSD of the protein structure (fitted onto its initial structure) as a function of time. A continuous increase in RMSDs to high values indicates a conformationally labile model. From Figure 4A, it is evident that the RMSD for the *mono1* simulation stabilizes after 1 ns, without any further drift. When the contributions from the NBD and TMD are plotted separately, it appears that the larger component of the overall protein drift is due to the TMD rather than the NBD. In the *mono2* simulation, where the protein was started in a rotated conformation, no stabilization is observed (yielding a final RMSD of 0.65 nm). The total Cα RMSD for *mono2* (not shown) increases monotonically throughout the 2 ns simulation. In both *mono1* and *mono2*, the TMD contributes the majority of the structural drift.

Table 2: MsbA Simulations

protein	simulation name	setup	total number of atoms	duration (ns)	final Cα RMSD, all residues (nm)[a]	final Cα RMSD, TMD (nm)[a]	final Cα RMSD, NBD (nm)[a]
MsbA monomer	Mono1	initial tilt angle from X-ray dimer	79 585	2	0.55	0.50	0.35
MsbA monomer	Mono2	initial tilt angle set to ca. 0°	60 436	2	0.65	0.68	0.40
MsbA dimer	dim1	dimer from X-ray structure	132 946	1	0.70	0.65	0.30

[a] For all Cα atoms of each model, averaged over the last 0.2 ns of simulation.

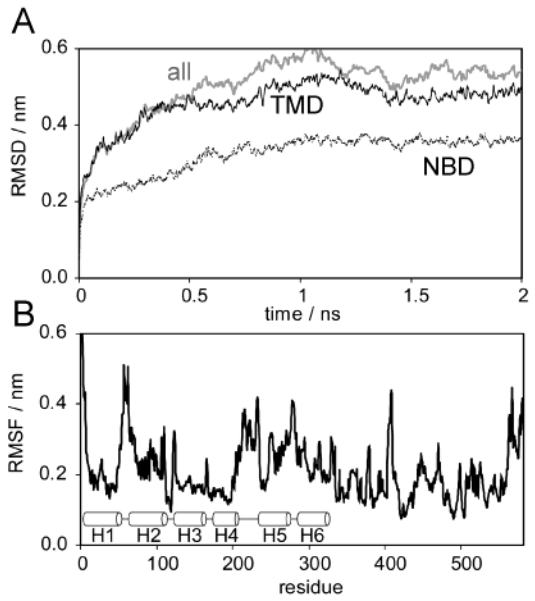

FIGURE 4: Simulation *mono1*. (A) Cα RMSDs vs time are shown for all residues (solid gray line); TMD helices fitted on the TMD and measured relative to the TMD (solid black line); and NBD fitted on the NBD and measured relative to the NBD (dotted black line). (B) Cα RMSFs vs residue number. The locations of TM helices H1−H6 are indicated.

Visual inspection of protein secondary structure in the *mono2* simulation revealed that the protein was unfolding. We will therefore only present further analysis of the *mono1* simulation, where the overall secondary structure remains intact and the protein drift appears to have stabilized after ca. 1 ns.

To further characterize the conformational fluctuations of the MsbA monomer over the 2 ns simulation, the root-mean squared fluctuation (RMSF) of the Cα coordinates from their starting positions were calculated as a function of residue number for *mono1* (Figure 4B). The RMSF plot reveals that the largest fluctuation occurs in the large extracellular loop (Pro50−Val70) between helices 1 and 2. Helices 1, 3, and 4 show low fluctuations throughout the simulation, while helices 2, 5, and 6 display a larger degree of mobility (see below). The most mobile segment of the NBD is the loop formed by residues 402−420.

The conformational fluctuations over the course of the simulation can be further unpacked by examining their time-averaged secondary structure content (Figure 5). Marked loss of initial secondary structure occurs at the N-terminus and in the intracellular loop (Ala 207−Arg 238), both of which were entirely model-built because of the absence of experimental data. It should be noted that some loss of helicity was observed in helix H3 (at residues 142−143). This same loss of helicity was also observed in three other MsbA screening simulations that were run to test different models

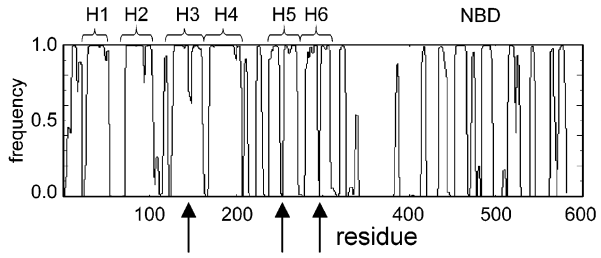

FIGURE 5: Secondary structure analysis. The time-averaged frequency of α-helical secondary structure vs residue number is shown for simulation *mono1*. The locations of the TM helices and of the NBD are indicated. The arrows indicate regions where significant loss of α-helicity occurs in TM helices H3, H5, and H6.

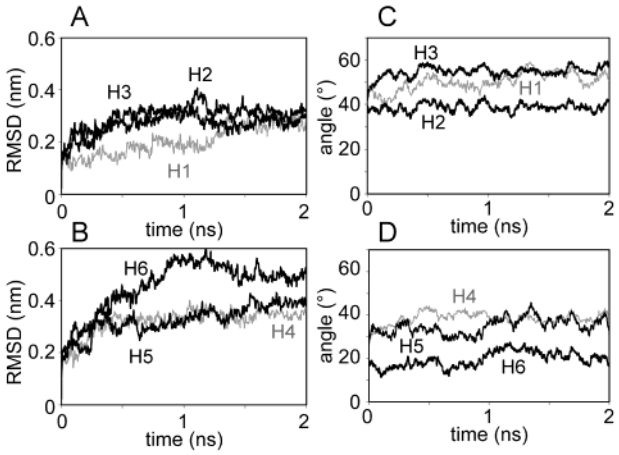

FIGURE 6: Intra-helix motions. The intra-helix RMSDs (A and B) and tilt angles with respect to the octane slab normal (C and D) are shown as a function of time. Intra-helix RMSDs were calculated by fitting each helix on itself and measuring the RMSD relative to the fit over time.

(data not shown). Significant helix distortions also occurred in H5 and H6, which are the most dynamic helices as indicated by RMSF values.

A large number of fluctuations in secondary structure occur in the NBD during the first nanosecond of simulation, but after that time the NBD conformation remains constant. Loss of helicity occurs in the ATP binding helix (residues 382−390), which is immediately N-terminal to the Walker A motif. This loss of helicity is likely due to the placement of the linker helix, which was modeled to pass through the NBD, close to the active site. The presence of this linker helix destabilizes the secondary structure of neighboring segments.

To investigate further the substantial motions in the TMD, the RMSDs of each TM helix relative to itself and tilt angles with respect to the membrane normal were plotted over simulation time (Figure 6). During the *mono1* simulation, helices H1−H3 display low (ca. 0.3 nm) and stable RMSD values. H6 shows considerable conformational drift, and H4

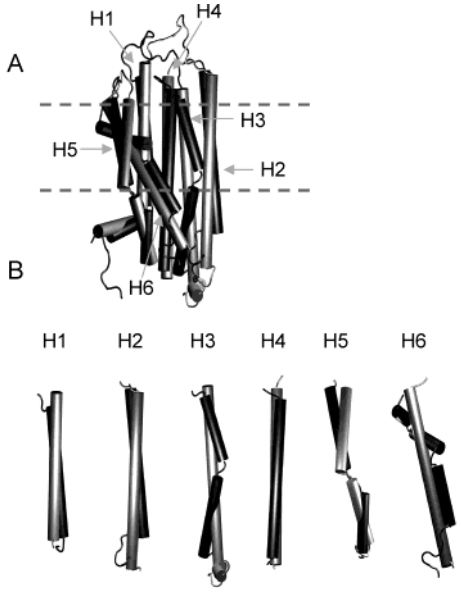

FIGURE 7: Post-simulation structure fitted onto the starting conformation for *mono1*. (A) The post 2 ns TMD structure (dark gray) fitted onto the starting TMD structure (pale gray). (B) The post 2 ns structure (dark gray) fitted onto the starting structure (pale gray) for each individual helix.

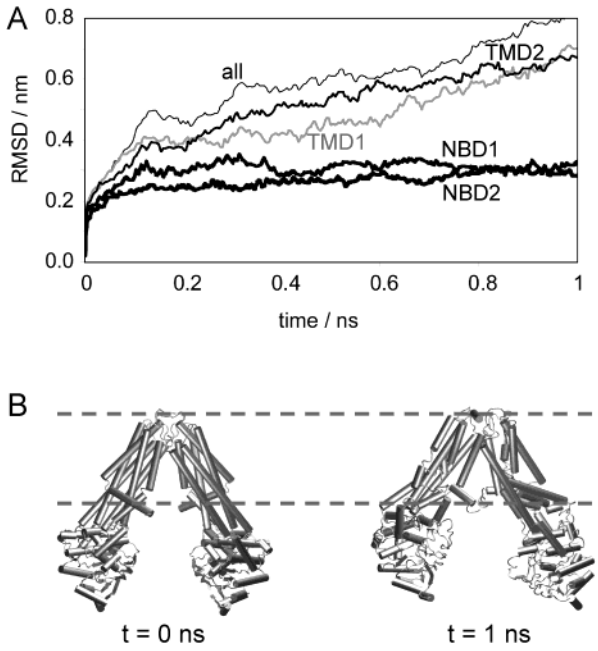

FIGURE 8: Dimer simulation. (A) Cα RMSDs vs time for all residues (thin black line); TMD1 (medium black line); TMD2 (medium gray line); and the two NBDs (thick black lines) are shown. (B) Structure of the dimer at the start ($t = 0$ ns) and end ($t = 1$ ns) of the simulation.

and H5 are intermediate. The helix angles with respect to the membrane normal show fluctuations throughout the 2 ns. These fluctuations are most pronounced for H3 and H6. The helix motions can be further assessed by overlaying the 2 ns structure of the protein onto its starting configuration (Figure 7). This shows that the high RMSD for H6 in the *mono1* simulation is due to a distortion of the helix (at Leu 294 and Met 295) within the hydrophobic core. Some helix distortion also occurs in H3 and H5. In H3, which displayed the highest amount of tilt in the *mono1* simulation, the lower half of the helix rotates such that it becomes more parallel with H2. The helix distortions in H5 and H6 are probably due to the presence of proline residues (Pro 253 and 297, respectively) that have been shown to distort transmembrane helices both in X-ray structures (*36*) and during MD simulations (*37*−*39*).

Dimer Simulation. Despite the fact that the monomeric MsbA model is relatively stable when inserted into the octane slab at the crystallographic angle, the dimer model shows considerable instability in its TMDs (Figure 8). In particular, the Cα RMSD monotonically increases over just 1 ns of simulation. The high RMSD is coupled to a significant loss of secondary structure and helix packing (not shown). Both NBDs, in contrast, were quite stable. The agreement between the RMSDs of NBD1 and NBD2 and between the RMSDs of TMD1 and TMD2 is good, suggesting that the observed drift is reproducible on the time scale of this simulation. Visual inspection (Figure 8B and see below) suggests the TMD instability was coupled to changes at the monomer/monomer interface.

DISCUSSION

Critical Evaluation of the Model. The protocols employed for modeling the intact MsbA structure made use of several existing algorithms and programs and were derived from the low resolution X-ray structure. The quality of the MsbA model is acceptable based on the assessment of the protein-checker programs employed. The TM helix packing at several points in the MsbA model is such that ca. 0.4 nm inter-helix gaps exist. These gaps were a result of modeling the TMD of MsbA without significant deviation of the Cα trace from the original X-ray coordinates. It was felt that since the only experimental data for the MsbA structure is the crystal structure, reasonable adherence to the Cα trace should be maintained during the modeling process.

A significant amount of structural rearrangement occurs in the TMD for both monomer simulations. Although the large degree of motion in the TMDs is not typical of membrane protein simulations, it may be related to the fact that significant repacking of the TMD occurs in P-glycoprotein (P-gp) (*40*) and other transporters during the transport cycle; therefore, the TMDs must inherently be very flexible. Indeed, the helix that displayed the largest amount of motion has been demonstrated to be of functional importance. During the *mono1* simulation, H6 displays the largest intra-helix RMSD. In P-gp, H6 has been implicated in substrate binding (*41*) and may be exposed to both aqueous and lipid environments during the transport cycle (*40*), confirming that dramatic conformational changes must occur in this helix.

The amount of motion observed during the simulations might be attributed to a lack of tight helix packing, which could cause protein instability during simulation. This lack of tight packing arises from the close adherence to the crystallographic MsbA Cα coordinates. The nature of the membrane mimetic might also influence protein stability. However, early screening simulations demonstrated that the use of a lipid bilayer rather than an octane slab did not improve stability (not shown). Furthermore, simulations of a number of membrane proteins (*20*) and models (*42*) in octane have shown the latter to be a suitable membrane mimetic for such studies. Use of a full phospholipid bilayer would require significantly longer simulation times to reach

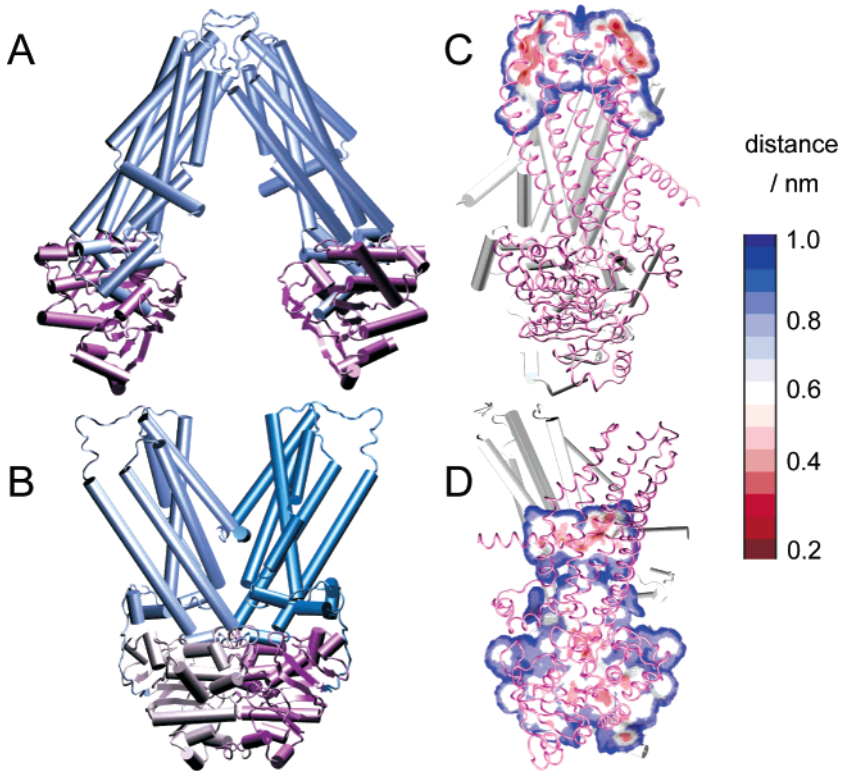

FIGURE 9: Alternative models of the dimer. The crystallographic dimer is shown in panels A and C; the alternative model of the dimer in panels B and D. The TMDs and NBDs are colored blue and purple, respectively. In panels C and D, the dimer contacts are shown using a contact distance map (colored on a 0.2−1.0 nm scale) sandwiched between two monomers, the foremost of which is shown in tube format and the rearmost in cylinder format. (These diagrams were generated using adsi module from MolSurfer (*51*)).

equilibrium than an octane slab. Furthermore, much longer simulations would have been required to reveal conformational changes.

MsbA Dimer. The MsbA molecule was relatively unstable in its dimeric form and as a monomer when inserted into the octane after rotating ca. 25° toward the bilayer normal. Thus, the MsbA monomer seems to prefer a tilted orientation in the bilayer but is destabilized in its crystallographic dimer. This result suggests that the MsbA dimer may be a result of crystal packing rather than a model of one of the in vivo dimeric states of the transporter molecule.

In an attempt to explore an alternative dimer, the X-ray structure of the NBD dimer of MJ0796 was used to construct a new configuration for MsbA. The MsbA NBDs were first homology modeled based on the MJ0796 dimer (ref *7*; PDB code 1L2T). The MsbA monomers were then least-squares fitted onto MJ0796 yielding a structure similar to an alternative dimer presented previously (*43*). When the monomers of MsbA are joined in this alternative arrangement, helix packing improves forming as a result of the new dimer contacts (Figure 9). This would be anticipated to improve the simulation stability of the dimeric model. Since the NBDs were positioned using the MJ0796-based MsbA NBD dimer, they are in agreement with the consensus nucleotide sandwich dimer arrangement that has been demonstrated in Rad50, BtuCD, MalK, and P-gp (*44*−*47*), respectively. In addition, the alternative dimer is compatible with low resolution EM data from P-gp, TAP, MRP, and YvcC (*40*, *48*−*50*). Thus, the alternative dimer may provide a model of MsbA that is more consistent with existing ABC transporter structural data.

ACKNOWLEDGMENT

Our thanks to the Oxford Supercomputing Centre for access to resources. J.D.C. thanks Christ Church and the Canadian Scholarship Trust for Hugh Pilkington and Kenneth-Carter scholarships, respectively. Our thanks to Jonathan Cuthbertson for analysis of transmembrane predictions and Ian Kerr for useful discussions.

REFERENCES

1. Linton, K. J., and Higgins, C. F. (1998) The *Escherichia coli* ATP-binding cassette (ABC) proteins, *Mol. Microbiol. 28*, 5−13.
2. Murakami, S., Nakashima, R., Yamashita, E., and Yamaguchi, A. (2002) Crystal structure of bacterial multidrug efflux transporter AcrB, *Nature 419*, 587−593.
3. Diederichs, K., Diez, J., Greller, G., Muller, C., Breed, J., Schnell, C., Vonrhein, C., Boos, W., and Welte, W. (2000) Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis*, *EMBO J. 19*, 5951−5961.
4. Gaudet, R., and Wiley, D. C. (2001) Structure of the ABC ATPase domain of human TAP1, the transporter associated with antigen processing, *EMBO J. 20*, 4964−4972.
5. Hung, L. W., Wang, I. X. Y., Nikaido, K., Liu, P. Q., Ames, G. F. L., and Kim, S. H. (1998) Crystal structure of the ATP-binding subunit of an ABC transporter, *Nature 396*, 703−707.
6. Karpowich, N., Martsinkevich, O., Millen, L., Yuan, Y. R., Dai, P. L., MacVey, K., Thomas, P. J., and Hunt, J. F. (2001) Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter, *Structure 9*, 571−586.
7. Smith, P. C., Karpowich, N., Millen, L., Moody, J. E., Rosen, J., Thomas, P. J., and Hunt, J. F. (2002) ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer, *Mol. Cell 10*, 139−149.

8. Yuan, Y. R., Blecker, S., Martsinkevich, O., Millen, L., Thomas, P. J., and Hunt, J. F. (2001) The crystal structure of the MJ0796 ATP-binding cassette—Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter, *J. Biol. Chem. 276*, 32313−32321.
9. Locher, K. P., Lee, A. T., and Rees, D. C. (2002) The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism, *Science 296*, 1091−1098.
10. Higgins, C. F. (2001) ABC Transporters: physiology, structure and mechanism—an overview, *Res. Microbiol. 152*, 205−210.
11. Chang, G., and Roth, C. B. (2001) Structure of MsbA from *E. coli*: a homologue of the multidrug resistance ATP binding cassette (ABC) transporters, *Science 293*, 1793−1800.
12. Grisshammer, R., and Tate, C. G. (1995) Overexpression of integral membrane-proteins for structural studies, *Q. Rev. Biophys. 28*, 315−422.
13. Rosenbusch, J. P. (2001) Stability of membrane proteins: relevance for the selection of appropriate methods for high-resolution structure determination, *J. Struct. Biol. 136*, 144−157.
14. Martin, A. C., MacArthur, M. W., and Thornton, J. M. (1997) Assessment of comparative modeling in CASP2, *Proteins: Struct., Funct., Genet. Suppl. 1*, 14−28.
15. Sali, A., and Blundell, T. L. (1993) Comparative protein modeling by satisfaction of spatial restraints, *J. Mol. Biol. 234*, 779−815.
16. Xiang, Z. X., and Honig, B. (2001) Extending the accuracy limits of prediction for side-chain conformations, *J. Mol. Biol. 311*, 421−430.
17. Vriend, G. (1990) WhatIf—a molecular modeling and drug design program, *J. Mol. Graph. 8*, 52−.
18. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) Procheck—a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
19. Lindahl, E., Hess, B., and van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis, *J. Mol. Model. 7*, 306−317.
20. Bond, P., Faraldo-Goméz, J., and Sansom, M. S. P. (2002) OmpA—A pore or not a pore? Simulation and modelling studies, *Biophys. J. 83*, 763−775.
21. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. (1981) *Intermolecular Forces*, Reidel, Dordrecht, The Netherlands.
22. Berweger, C. D., van Gunsteren, W. F., and Mullerplathe, F. (1995) Force-field parametrization by weak-coupling—reengineering SPC water, *Chem. Phys. Lett. 232*, 429−436.
23. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath, *J. Chem. Phys. 81*, 3684−3690.
24. Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations, *J. Comput. Chem. 18*, 1463−1472.
25. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577−2637.
26. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD—Visual Molecular Dynamics, *J. Mol. Graph. 14*, 33−38.
27. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server, *Bioinformatics 16*, 404−405.
28. Holm, L., and Sander, C. (1991) Database algorithm for generating protein backbone and side-chain coordinates from a C-α trace: application to model-building and detection of coordinate errors, *J. Mol. Biol. 218*, 183−194.
29. Rost, B. (1999) Twilight zone of protein sequence alignments, *Protein Eng. 12*, 85−94.
30. Lapinski, P. E., Neubig, R. R., and Raghavan, M. (2001) Walker A lysine mutations of TAP1 and TAP2 interfere with peptide translocation but not peptide binding, *J. Biol. Chem. 276*, 7526−7533.
31. Saveneau, L., Daniel, S., and van Endert, P. M. (2001) Distinct functions of the ATP binding cassettes of transporters associated with antigen-processing: a mutational analysis of Walker A and B sequences, *J. Biol. Chem. 276*, 22107−22113.
32. Powe, A. C., Al-Nakkash, L., Li, M., and Hwang, T. C. (2002) Mutation of Walker-A lysine 464 in cystic fibrosis transmembrane conductance regulator reveals functional interaction between its nucleotide-binding domains, *J. Physiol. 539*, 333−346.
33. Vergani, P., Basso, C., Nairn, A. C., and Gadsby, D. C. (2002) Effects on CFTR Cl^- channel gating of Walker A lysine mutation K464A imply allosteric interaction between NBDS, *Biophys. J. 82*, 1162.
34. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
35. Krieger, E., Koraimann, G., and Vriend, G. (2002) Increasing the precision of comparative models with YASARA NOVA—a self-parametrizing force field, *Proteins: Struct., Funct., Genet. 47*, 393−402.
36. Cordes, F. S., Bright, J. N., and Sansom, M. S. P. (2002) Proline-induced distortions of transmembrane helices, *J. Mol. Biol. 323*, 951−960.
37. Tieleman, D. P., Shrivastava, I. H., Ulmschneider, M. B., and Sansom, M. S. P. (2001) Proline-induced hinges in transmembrane helices: possible roles in ion channel gating, *Proteins: Struct., Funct., Genet. 44*, 63−72.
38. Bright, J. N., Shrivastava, I. H., Cordes, F. S., and Sansom, M. S. P. (2002) Conformational dynamics of helix S6 from Shaker potassium channel: simulation studies, *Biopolymers 64*, 303−313.
39. Bright, J. N., and Sansom, M. S. P. (2003) The flexing/twirling helix: exploring the flexibility about molecular hinges formed by proline and glycine motifs in transmembrane helices, *J. Phys. Chem. B 107*, 627−636.
40. Rosenberg, M. F., Velarde, G., Ford, R. C., Martin, C., Berridge, G., Kerr, I. D., Callaghan, R., Schmidlin, A., Wooding, C., Linton, K. J., and Higgins, C. F. (2001) Repacking of the transmembrane domains of P-glycoprotein during the transport ATPase cycle, *EMBO J. 20*, 5615−5625.
41. Ambudkar, S. V., Dey, S., Hrycyna, C. A., Ramachandra, M., Pastan, I., and Gottesman, M. M. (1999) Biochemical, cellular, and pharmacological aspects of the multidrug transporter, *Annu. Rev. Pharmacol. Toxicol. 39*, 361−398.
42. Capener, C. E., and Sansom, M. S. P. (2002) MD Simulations of a K channel model—sensitivity to changes in ions, waters, and membrane environment, *J. Phys. Chem. B 106*, 4543−4551.
43. Thomas, P. J., and Hunt, J. F. (2001) A snapshot of Nature's favorite pump, *Nature Struct. Biol. 8*, 920−923.
44. Hopfner, K. P., Karcher, A., Shin, D. S., Craig, L., Arthur, L. M., Carney, J. P., and Tainer, J. A. (2000) Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily, *Cell 101*, 789−800.
45. Locher, K. P., Lee, A. T., and Rees, D. C. (2002) The *E. coli* BtuCD structure: A framework for ABC transporter architecture and mechanism, *Science 296*, 1091−1098.
46. Fetsch, E. E., and Davidson, A. L. (2002) Vanadate-catalyzed photocleavage of the signature motif of an ATP-binding cassette (ABC) transporter, *Proc. Natl. Acad. Sci. U.S.A. 99*, 9685−9690.
47. Loo, T. W., Bartlett, M. C., and Clarke, D. M. (2002) The "LSGGQ" motif in each nucleotide-binding domain of human p-glycoprotein is adjacent to the opposing Walker A sequence, *J. Biol. Chem. 277*, 41303−41306.
48. Rosenberg, M. F., Callaghan, R., Ford, R. C., and Higgins, C. F. (1997) Structure of the multidrug resistance P-glycoprotein to 2.5 nm resolution determined by electron microscopy and image analysis, *J. Biol. Chem. 272*, 10685−10694.
49. Velarde, G., Ford, R. C., Rosenberg, M. F., and Powis, S. J. (2001) Three-dimensional structure of transporter associated with antigen processing (TAP) obtained by single particle image analysis, *J. Biol. Chem. 276*, 46054−46063.
50. Chami, M., Steinfels, E., Orelle, C., Jault, J. M., Di Pietro, A., Rigaud, J. L., and Marco, S. (2002) Three-dimensional structure by cryo-electron microscopy of YvcC, an homodimeric ATP-binding cassette transporter from *Bacillus subtilis*, *J. Mol. Biol. 315*, 1075−1085.
51. Gabdoulline, R. R., Wade, R. C., and Walther, D. (1999) MolSurfer: two-dimensional maps for navigating three-dimensional structures of proteins, *Trends Biochem. Sci. 24*, 285−287.

BI027337T