The First View of an ABC Transporter: The X-ray Crystal Structure of MsbA from *E. coli*

Lutz Schmitt*[a]

KEYWORDS:

ABC transporters · membrane proteins · multidrug resistance
structure – activity relationships · structure elucidation

Membrane biochemistry is a very exciting and fast-moving research area. Our understanding of membranes and the proteins embedded within the two-dimensional fluid of lipids is still limited. The processes occurring at this interface are even less understood. Chang and Roth have recently determined the first X-ray crystal structure of an ATP binding casette (ABC) transporter, MsbA from *Escherichia coli* (Figure 1; ATP = adenosine triphosphate).[1] This structure represents a major achievement in the field of membrane biochemistry. Biological membranes are a protective shield against a hostile environment. However, the two-dimensional, impermeable nature of biological membranes[2] creates a severe and sometimes life-threatening problem. Living organisms have to take up nutrients, extrude harmful substances, and of course exchange information. During evolution many transmembrane proteins have evolved, to ensure that living organisms are able to survive and multiply.

One of the most common families of transmembrane proteins found in all three kingdoms of life is the family of ABC transporters.[3] The substrates of this superfamily range from small inorganic ions (such as chloride ions) to amino acids, sugars, drugs, and even large proteins. Despite such diversity, all members of the family of ABC transporters share a common blueprint, which comprises of four modules: two transmembrane domains (TMDs) and two nucleotide- or ATP-binding domains (NBDs). In all cases, the energy released during ATP hydrolysis by the NBD is used to translocate the substrate. Every arrangement of these four domains is possible. In bacteria, separate polypeptide chains commonly make up each of the four modules. However, one NBD and one TMD might be fused on a single protein, to generate a so-called half-size transporter. In eukaria, a single polypeptide chain generally makes up all four domains. This architecture corresponds to the so-called full-size transporter; the half-sized transporter is the exception.[4]

On a functional and structural level, it is now commonly accepted that the NBDs provide only the energy for translocation while the TMDs confer substrate specificity. The NBDs, which define an ABC transporter, contain three conserved sequence motifs. The Walker A motif (consensus sequence: GXXGXGKST, where X denotes any amino acid), the Walker B motif (consensus sequence: hhhhD, where h denotes any hydrophobic amino acid),[5] and the signature motif or "C-loop" (consensus sequence: LSGQQR), which is specific for ABC transporters. As a consequence, NBDs are well-conserved among this protein family. On the other hand, TMDs share little sequence homology and even the number of α-helices seems to vary from one ABC transporter to the other. A six α-helix core is the minimal requirement for an ABC transporter, but variations are generally the rule rather than the exception.[3]

The best-characterized ABC transporters are bacterial ones, for example, the histidine permease[6] or maltose transporters.[7] In humans, the cystic fibrosis transmembrane conductance regulator (CFTR),[8] transporter associated with anti-

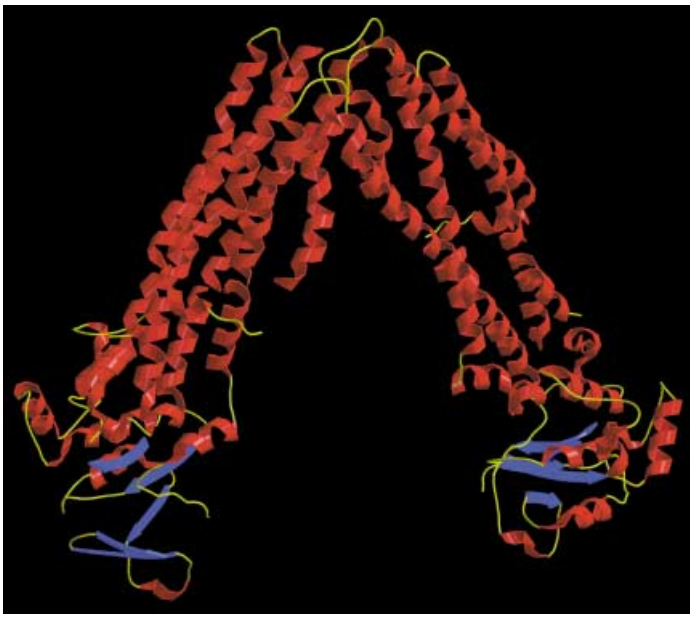

Figure 1. *Three-dimensional structure of the homodimeric MsbA. α-helices are represented in red, β-strands in blue, and coils in yellow. The figure was generated by using the MOLSCRIPT[31] and Raster3D[32] programs based on the Protein Databank entry 1JSQ.*

[a] *L. Schmitt*
Institute of Biochemistry, Biocenter N210
Johann Wolfgang Goethe University Frankfurt
Marie-Curie Strasse 9
60439 Frankfurt (Germany)
Fax: (+69) 79829-495
E-mail: lschmitt@em.uni-frankfurt.de

gen processing (TAP),[9]and multidrug resistance protein 1 (MDR1)[10]are the most prominent members of the ABC transporter family. Point mutations in CFTR, mainly F508Δ, are the genetic cause of cystic fibrosis, the most common inherited disease among Caucasians. MDR1 causes severe problems in chemotherapy. MDR1 is overexpressed in most tumors upon treatment with chemotherapeutics and results in resistance against the applied drugs. As a consequence, the drug dose has to be increased so that even healthy cells are affected and the chemotheraphy becomes ineffective. These two examples explain why a molecular understanding of ABC transporters that are involved in diseases, especially transporters in human, is more than worthwhile. But multidrug resistance is not only a human phenomenon. More or less every organism contains a transmembrane protein conferring resistance to many natural and synthetic drugs,[11] and in many cases these proteins belong to the family of ABC transporters.[12]

Despite two decades of intensive research and an overwhelming body of experimental data, little is know about the mechanisms of transport or the coupling of ATP hydrolysis and substrate translocation; there is also a lack of structural information. For example, MDR1 is able to transport more or less every hydrophobic drug known today.[10] Examples are vinca alkaloids (e.g., vinblastine), antibiotics (e.g., actinomycin D), Taxol, protein-synthesis inhibitors (e.g., puromycin), DNA intercalators (e.g., ethidium bromide), toxic peptides (e.g., valinomycin), or fluorescent dyes (e.g., rhodamine). How can a single protein achieve such diversity while maintaining specificity? This knowledge is, of course, a prerequisite for the development of drugs or specific inhibitiors that would open up new avenues not only in cancer therapy but also in the treatment of infectious diseases spread through bacteria, which have achieved multidrug resistance to many commonly used drugs. In the last three years, several three-dimensional structures of NBDs from ABC transporters have been solved by X-ray crystallography.[13–17] However, no structural information for the TMDs of an ABC transporter was available. The only exceptions were the low-resolution structures of human MDR1[18, 19] and multidrug resistance related protein 1 (MRP1)[20] obtained from single-particle analysis and two-dimensional crystals.

In Science,[1] Geoffrey Chen and Christopher B. Roth from the Department of Molecular Biology at The Scripps Research Institute in La Jolla, USA, reported the first three-dimensional structure of a complete ABC transporter, MsbA from *E. coli* (Figure 1). MsbA transports lipid A. It shares around 30% sequence homology with human MDR1. However, in contrast to MDR1, which is a full-size transporter, functional MsbA is assembled from two half-size transporters. It might very well serve as a structural model for many of the ABC transporters which give rise to the phenomena of drug resistance.

The bottleneck of modern X-ray crystallography is the availability of well-ordered three-dimensional crystals. Due to the techniques of modern molecular biology, more or less every water-soluble protein can be produced in quantities sufficient for structural analysis. However, structural investigations of membrane proteins are still hampered by the limited amounts of protein. In addition, another important parameter has to considered: the choice of the right detergent.[21, 22] Taking these points together, even the nonexpert can understand why the structure determination of a membrane protein is more than an ordinary challenge.

The milestone of the MsbA structure was achieved by a tour de force. In order to obtain X-ray suitable crystals, a tremendous effort was undertaken. A total of 20 different MDR-mediating ABC transporters from 14 different organisms were cloned, overexpressed, purified, and investigated for their ability to crystallize. The incredible number of 96 000 crystallization trials was performed with around 20 different detergents. At the end of this long torture, 35 crystal forms were obtained. Out of these crystals, MsbA had the best diffraction quality. However, even after this impressive struggle, native crystals of MsbA diffracted only to 6.2 Å and showed strong anisotropic diffraction. Nevertheless, the authors proceeded and applied a so-called "refinement strategy", which was intended to strengthen lattice contacts in order to improve diffraction quality. After another tour de force, which included an intensive screening of detergents, detergent concentrations, temperatures, and inorganic and organic compounds, the diffraction limit was raised to 4.5 Å through $OsCl_3$. As this was still not enough, the procedure of refinement also had to be modified to the needs of this particular structure. It is beyond the scope of this article to describe the efforts undertaken in the work, but it finally resulted in an electron density of MsbA, which was of sufficient quality to trace the protein backbone chain. The quality of the structure determined, even in light of the moderate resolution, is indicated by the *R* factor (27%, $R_{free} = 38\%$).

The structure presented by Chang and Roth answers many questions, but it also leaves many questions open and even raises extra questions. The overall structure of MsbA at 4.5 Å resolution is shown in Figure 1. The crystal structure is consistent with the fact that MsbA forms a homodimer as the functional unit. Perhaps most important is the fact that the TMDs are solely composed of α-helices. This proves, beyond any doubt, that the substrate pathway (TMD) is in agreement with the secondary structure prediction proposed by many other laboratories and supported by biochemical evidence. In the case of MsbA, the TMDs are each composed of six α-helices, which have a tilt angle of 30 – 40° with respect to the bilayer normal. This and many other details of the structure agree with two decades of experimental work and give us confidence. The NBDs, as far as they are visible within the experimental electron density, agree with the recently published structures of isolated NBDs.[13–17] However, the N-terminal region (residues 341 – 418), which includes the Walker A motif, is not visible. The exact position and role of the NBDs are therefore under speculation. The first surprise of the structure is the presence of a third type of domain, which has not been suggested before. This intracellular domain (ICD), which corresponds to residues 97 – 139 (ICD1), 193 – 252 (ICD2), and 302 – 327 (ICD3), is helical in nature (three helices in the case of ICD 1 and two helices each for ICD2 and ICD3) and connects the NBD and the TMD in each half-sized transporter. IDC1 is positioned directly "above" the ABC signature motif (LSGGQQ) of the NBD and is

thought to transmit information between the NBD and TMD (see Figure 2). This transmission very likely occurs through rearrangement of the helical IDC.

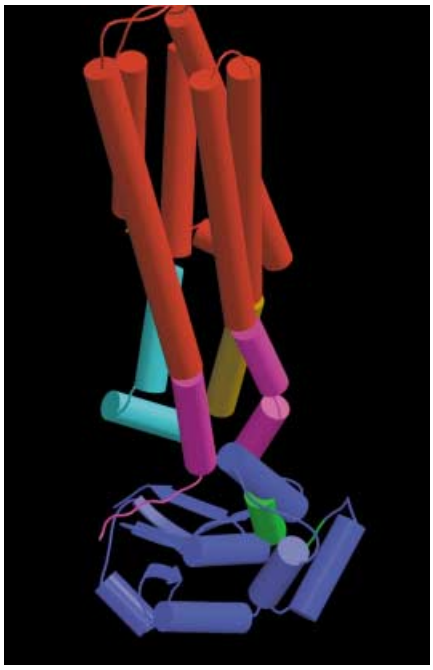

Figure 2. *Organization of the ICD. For simplicity only a monomeric unit of MsbA is shown. The TMD is shown in red, the NBD in dark blue and the ICD in purple (ICD1), light blue (ICD3), and orange (ICD2). The C-loop within the NBD is highlighted in green. Helix 2 of ICD 1 is positioned above the C-loop, which implys a possible pathway for signal transduction. Please note that the second helix of ICD2 is hidden in this orientation. For further details, please see text. The figure was generated as in Figure 1 by rotation of 90° in plane.*

Obviously, the three-dimensional architecture of the TMDs is the most exciting and surprising part of the structure. The helices form a cone-like structure with two large openings within the bilayer section of the protein (Figures 1 and 3). The part of MsbA located in the outer leaflet is closed, while the part positioned in the inner leaflet of the membrane is widely open. The two openings facing the bilayer are around 25 Å wide and located solely within the inner (cytosolic) leaflet of the bilayer. The base of the formed chamber, which is located at the cytosolic side of the bilayer is roughly 45 Å wide. This chamber can easily accommodate the substrate, lipid A. The opening of the chamber within the inner leaflet of the bilayer guarantees free entry of the substrate, while the closed structure within the outer leaflet prohibits entry or exit. Therefore, the unidirectional translocation of lipid A from one side of the bilayer to the other is understandable. The openings of the chamber are defined by transmembrane helix 2 (TM2) from one monomer within the dimer and TM5 from the other. However, the whole chamber is formed from side chains of all the TMs. Another interesting point is the charge distribution within the chamber. While the part of the substrate binding side located at the inner leaflet contains a cluster of positively charged amino acids, the part located in the outer leaflet is hydrophobic in nature. As suggested in Figures 1 and 3, the chamber creates a large perturbation within the inner leaflet. Intuitively, one would expect that this arrangement generates a lot of stress on the membrane and that the impermeable nature of the bilayer might be endangered.

Based on the structure of MsbA, which was obtained in the absence of substrate and any nucleotide, Chang and Roth have derived a possible mechanism of substrate transport for MsbA and for ABC transporters in general. The model is based on the presented structure and a body of available biochemical data. In each monomer, the ICD seems to be a "sensing unit", which transmits signals from the TMD to the NBD and vice versa by conformational changes. Lipid A binds to the open chamber from the inner leaflet of the bilayer. Information of this event is conducted through the ICDs to the NBDs; this triggers ATP hydrolysis. Such substrate-induced stimulation of ATPase activity has also been demonstrated for MDR1 (see, for example, ref. [23]) or TAP.[24] A conformational change of the NBDs upon ATP hydrolysis is proposed to induce an interaction between both NBDs. However, the N-terminal regions of the NBDs are not visible in the electron density and their exact position and interaction is speculative. However, such a scenario would result in a rearrangement of the whole molecule. The reorganization of the NBDs influences the TMDs. The chamber is lined with a cluster of charges, which creates an energetically unfavorable situation for lipid A. Such charges also imply the presence of bound solvent. However, as pointed out above, the TMDs contain an asymmetrical charge and polarity distribution: charged and highly polar in the lower part (chamber), while hydrophobic in the upper part. Together with the interaction of the NBDs upon ATP hydrolysis and a subsequent structural reorganization, lipid A flips over

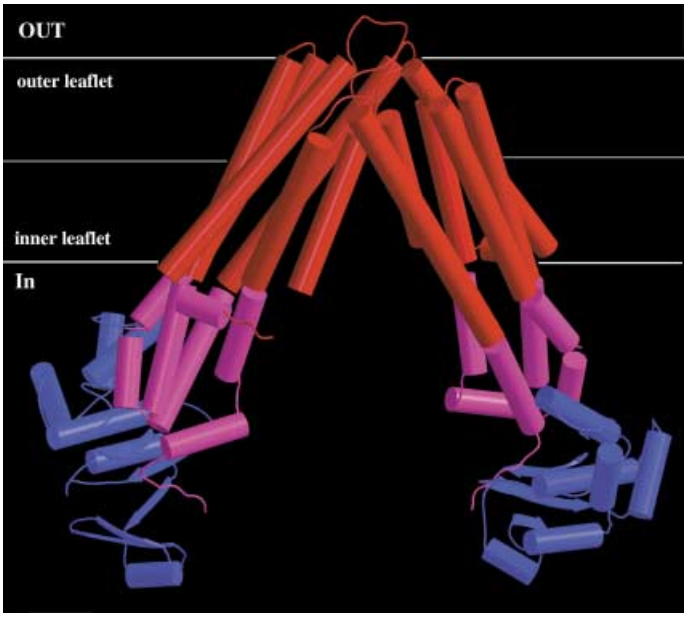

Figure 3. *Domain organization of the homodimeric MsbA. The TMDs are given in red, the ICDs in purple, and the NBDs in blue. The orientation of MsbA is identical to the one shown in Figure 1. The orientation and location of the putative bilayer is indicated with solid lines. The middle line indicates the border between the inner and outer leaflets of the membrane bilayer. The figure was generated as in Figure 1.*

into the upper part, which is energetically favored. Such a model is in agreement with the observed vectorial transport of lipid A. After flipping, the substrate is properly oriented to enter the outer leaflet of the membrane and complete the transport cycle. This, of course, requires the rearrangement of TM2 and TM5 to create an opening into the outer leaflet. Finally, the extrusion of lipid A sends a signal to the NBDs, probably through the ICDs again, which induces ADP – ATP exchange or spontaneous ADP release. This brings the system back to the ground state. One ATP molecule is consumed per transported substrate during the proposed cycle. This agrees with data obtained for MDR1.[25, 26] However, the "tilting" movement of the TMDs is energized by ATP hydrolysis, while the flipping of lipid A is driven by charge and polarity gradients along the TMDs. Such a tilting would impose a large amount of physical stress on both leaflets of the membrane and require lipid reorganization to counterbalance the different space requirements of MsbA during the transport cycle. The proposed mechanism raises another question: What is the driving force of the transport process? A recognition step has to take place, because MsbA transports lipid A with high specificity. After ATP hydrolysis, hydrophobic interactions drive lipid A to flip into the upper part of MsbA. At this stage and based on the proposed model, it is not obvious how the substrate is released into the outer leaflet of the bilayer. Are lateral, two-dimensional density gradients involved or is it simply a diffusion-controlled process? Further structural and biochemical investigations are necessary to clarify this point and prove the proposed transport cycle of MsbA.

Of course, Chang and Roth do not propose that the presented model holds for all ABC transporters, especially not for those transporting hydrophilic substrates. Nevertheless, the "tilting model" derived from the structure might serve as a general scheme for MDR-mediating ABC transporters. However, it has been shown that LmrA, the MDR1 homologue of *Lactococcus lactis*, extrudes the substrate into the extracellular medium[27] and not, like MsbA, into the outer leaflet. Even from a structural point of view, contradicting results exist. The low-resolution structure of MDR1 derived from two-dimensional crystals clearly shows a large extracellular opening (around 25 Å) in MDR1 in the absence of substrate and nucleotide.[18] A similar observation was made for MRP1.[20] Additionally, two-dimensional crystals of MDR1 in different functional states of the NBDs displayed large conformational changes of the TMDs.[19] Apart from the extracellular opening, cross-linking studies performed with MDR have shown that helices of the TMDs are in close proximity,[28, 29] although they are far apart in the structure of MsbA. From these data, Rosenberg et al. derived a model, in which the binding of ATP is used for substrate translocation from the inner to the outer leaflet.[19] This goes in hand with a reduced affinity of substrate. In contrast, Chang and Roth propose that hydrolysis is employed for the tilting of the TMDs while substrate flipping is a more or less spontaneous process. Additional biochemical data indicate that the NBDs strongly interact and act in an alternating fashion.[30] No evidence can be derived from the presented structure to clarify this point.

Despite the open questions and differences between the available structural data, it has been now demonstrated that ABC transporters can be crystallized and their structure solved by X-ray crystallography. The MsbA structure is only a first, very important and exciting step towards further understanding the structure and function of ABC transporters. Of course, the dimer interface of the TMDs seen in the structure might not be the biologically relevant one. Further biochemical analysis is necessary to prove this, but in favor of the observed interface is the fact that Chang and Roth used only protein that corresponded to the dimeric state of MsbA for the crystallization set-ups. The low sequence similarity of ABC transporters within the transmembrane region explains the different substrate specificity but might also imply different transport pathways. Such a situation is not in favor of the conservation of structure and function in biological systems. Nevertheless, evolutionarily related ABC transporters,[4] such as CFTR and MRP5 or TAP and hemolysin B, display very different substrate specificity. Many puzzling questions in the field of ABC transporters still wait to be answered. A lot of structural investigations will have to be undertaken until a clear picture of the structure – function relationship of ABC transporters will arise. However, the structure of Chang and Roth will guide future biochemical and biophysical studies that will help us to understand the molecular mechanisms of the extremely large family of diverse membrane transport proteins.[3]

I would like to apologize to all colleagues whose work has not been referenced appropriately. I thank Carsten Horn, Nils Hanekop, Stanislav Gorbulev, and Dr. Robert Tampé (Institute of Biochemistry, Johann Wolfgang Goethe University Frankfurt) for many stimulating and exciting discussions about ABC transporters. Our work is supported by the Deutsche Forschungsgemeinschaft through the Emmy Noether program and the Graduiertenkolleg "Protein function on an atomic level" (Philipps University Marburg).

[1] G. Chang, C. B. Roth, *Science* **2001**, *293*, 1793 – 1800.
[2] S. J. Singer, G. L. Nicolson, *Science* **1972**, *175*, 720 – 731.
[3] C. F. Higgins, *Annu. Rev. Cell Biol.* **1992**, *8*, 67 – 113.
[4] W. Saurin, M. Hofnung, E. Dassa, *J. Mol. Evol.* **1999**, *48*, 22 – 41.
[5] J. E. Walker, M. Saraste, M. J. Runswick, N. J. Gray, *EMBO J.* **1982**, *1*, 945 – 951.
[6] G. F.-L. Ames, *Int. Rev. Cytol.* **1992**, *137A*, 1 – 35.
[7] M. Ehrmann, R. Ehrle, E. Hofmann, W. Boos, A. Schlosser, *Mol. Microbiol.* **1998**, *29*, 685 – 694.
[8] D. C. Gadsby, A. C. Nairn, *Physiol. Rev.* **1999**, *79*, 77 – 107.
[9] L. Schmitt, R. Tampé, *ChemBioChem* **2000**, *1*, 16 – 35.
[10] M. M. Gottesman, I. Pastan, *Annu. Rev. Biochem.* **1993**, *62*, 385 – 427.
[11] H. W. van Veen, W. N. Konings, *Semin. Cancer Biol.* **1997**, *8*, 183 – 191.
[12] H. W. van Veen, W. N. Konings, *Biochim. Biophys. Acta* **1998**, *1365*, 31 – 36.
[13] L. W. Hung, I. X. Y. Wang, K. Nikaido, P. Q. Liu, G. F. L. Ames, S. H. Kim, *Nature* **1998**, *396*, 703 – 707.
[14] K. Diederichs, J. Diez, G. Greller, C. Muller, J. Breed, C. Schnell, C. Vonrhein, W. Boos, W. Welte, *EMBO J.* **2000**, *19*, 5951 – 5961.
[15] R. Gaudet, D. C. Wiley, *EMBO J.* **2001**, *20*, 4964 – 4972.
[16] N. Karpowich, O. Martsinkevich, L. Millen, Y. R. Yuan, P. L. Dai, K. MacVey, P. J. Thomas, J. F. Hunt, *Structure* **2001**, *9*, 571 – 586.

[17] Y. R. Yuan, S. Blecker, O. Martsinkevich, L. Millen, P. J. Thomas, J. F. Hunt, *J. Biol. Chem.* **2001**, *276*, 32313 – 32321.
[18] M. F. Rosenberg, R. Callaghan, R. C. Ford, C. F. Higgins, *J. Biol. Chem.* **1997**, *272*, 10685 – 10694.
[19] M. F. Rosenberg, G. Velarde, R. C. Ford, C. Martin, G. Berridge, I. D. Kerr, R. Callaghan, A. Schmidlin, C. Wooding, K. J. Linton, C. F. Higgins, *EMBO J.* **2001**, *20*, 5615 – 5625.
[20] M. F. Rosenberg, Q. Mao, A. Holzenburg, R. C. Ford, R. G. Deeley, S. P. Cole, *J. Biol. Chem.* **2001**, *276*, 16076 – 16082.
[21] R. M. Garavito, D. Picot, *Methods* **1990**, *1*, 57 – 69.
[22] R. M. Garavito, D. Picot, P. J. Loll, *J Bioenerg. Biomembr.* **1996**, *28*, 13 – 27.
[23] K. Szabo, E. Welker, Bakos, M. Muller, I. Roninson, A. Varadi, B. Sarkadi, *J. Biol. Chem.* **1998**, *273*, 10132 – 10138.
[24] S. Gorbulev, R. Abele, R. Tampé, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3732 – 3737.
[25] G. D. Eytan, R. Regev, Y. G. Assaraf, *J. Biol. Chem.* **1996**, *271*, 3172 – 3178.
[26] A. B. Shapiro, V. Ling, *Eur. J. Biochem.* **1998**, *254*, 189 – 193.
[27] H. W. van Veen, A. Margolles, M. Muller, C. F. Higgins, W. N. Konings, *EMBO J.* **2000**, *19*, 2503 – 2514.
[28] T. W. Loo, D. M. Clarke, *J. Biol. Chem.* **1996**, *271*, 15414 – 15419.
[29] T. W. Loo, D. M. Clarke, *J. Biol. Chem.* **1996**, *271*, 27482 – 27487.
[30] A. E. Senior, *Acta Physiol. Scand.* **1998**, *163*, 213 – 218.
[31] P. J. Kralaukis, *J. Appl. Crystallogr.* **1991**, *24*, 946 – 950.
[32] E. A. Merritt, D. J. Bacon, *Methods Enzymol.* **1997**, *277*, 505 – 524.