Volume 280, number 2, 379–382 FEBS 09534 March 1991

ADONIS 001457939100344U

The structure of porin from *Rhodobacter capsulatus* at 1.8 Å resolution

M.S. Weiss[1], A. Kreusch[1], E. Schiltz[1], U. Nestel[2], W. Welte[2], J. Weckesser[3] and G.E. Schulz[1]

[1]*Institut für Organische Chemie und Biochemie der Universität, Albertstr. 21, D-7800 Freiburg i. Br., Germany,* [2]*Institut für Biophysik und Strahlenbiologie der Universität, Albertstr. 23, D-7800 Freiburg i. Br., Germany* and [3]*Institut für Biologie II, Mikrobiologie der Universität, Schänzlestr. 1, D-7800 Freiburg i. Br., Germany*

Received 4 January 1991

The structure of the porin from *Rhodobacter capsulatus* was determined at a resolution of 1.8 Å. The analysis started from a closely related crystal structure that had been solved at a medium resolution of 3 Å using multiple isomorphous replacement and solvent flattening. The new structure contains the complete sequence of 301 amino acid residues. Refinement of the model is under way; the present R-factor is 22% with good geometry. Except for the lengths of several loops, the resulting chain fold corresponds to the medium resolution model. The membrane channel is lined by a large number of ionogenic side chains with characteristic segregation of differently charged groups.

Porin; Membrane channel structure; β-Barrel topology; X-Ray structure; *Rhodobacter capsulatus*

1. INTRODUCTION

Porins are found in the outer membranes of Gram-negative bacteria, mitochondria and chloroplasts. They form channels for small hydrophilic molecules that are in most cases weakly ion-selective. Typical exclusion limits are around 600 Da [1–3]. Porins are quite resistant to denaturation by heat or detergents. They are trimeric and consist predominantly of β-pleated sheet structure as demonstrated spectroscopically [4]. In contrast, all the other structurally known integral membrane proteins are essentially α-helical [5–7]. Two-dimensional crystals of porin have been studied by electron microscopic methods yielding the general shape of the molecule [8–13]. Also, a number of three-dimensional porin crystals suitable for X-ray diffraction analysis have been reported [14–17]. The crystal structure of the porin from *Rhodobacter capsulatus* is known at medium resolution [18]. Using a related, new crystal form [19] and the amino acid sequence (E. Schultz, A. Kreusch, U. Nestel and G.E. Schulz, unpublished results) the resolution of the structure analysis has now been extended to 1.8 Å. Here, we report the chain fold of the 301 amino acids and describe the channel.

2. MATERIALS AND METHODS

An established procedure for the isolation of porin from *Rhodobacter capsulatus* strain 37b4 had given rise to crystals diffracting to a medium resolution of about 2.8 Å [3,16], the structure of which had been solved at low [20] and then at medium resolution [18]. This crystal form is now called 'form-A'. By modifying the purification procedure [21], in particular by dispensing with EDTA and replacing SDS by *N,N*-dimethyl-dodecylaminoxide (LDAO), we could grow the new crystal 'form-B' that diffracts to 1.8 Å resolution [19]. As demonstrated by SDS-polyacrylamide gel electrophoresis of dissolved crystals, the high resolution crystal form-B contains porin trimers, whereas crystal form-A contains monomers.

Both crystal forms belong to space group R3 with one monomer in the asymmetric unit and grow with nearly cubic habits to sizes of about (500 μm³). The unit cell dimensions are a_{hex} = 95.3 Å, c_{hex} = 146.8 Å in form-A [18] and a_{hex} = 92.3 Å, c_{hex} = 146.2 Å for the high resolution crystal form-B. The X-ray diffraction patterns of both crystal forms resemble each other at low resolution. The structure factor amplitude difference R_F in the range ∞ to 5.9 Å is 34%, with $R_F = 2 \cdot \Sigma|F_1 - F_2|/\Sigma(F_1 + F_2)$. Unit cell axes and R_F indicate clearly that the molecular packings are closely related; crystal form-B is a slightly contracted crystal form-A. For data collection of crystal form-B, the crystals were stored and handled at room temperature in a buffer containing 30% (w/v) polyethyleneglycol 600, 300 mM LiCl, 0.6% (w/v) *n*-octyltetraoxyethylene, 3 mM NaN_3 and 20 mM Tris-HCl at pH 7.2. In the resolution range ∞–3.6 Å), data were collected on a 4-circle diffractometer (Siemens-Nicolet). High resolution data up to 1.8 Å were measured at the synchrotron facility of the EMBL outstation at DESY, Hamburg. All data were merged to yield a final data set of 42914 reflections that is 99% complete in the range ∞–1.8 Å.

The analysis of crystal form-B was started using a partially refined medium resolution model [18] which at this point consisted of 310 residues with an 'X-ray sequence' that had been read from the electron density map. Because of the close packing relationship between crystal forms-A and -B, the starting model as defined in crystal form-A was positioned into crystal form-B solely by shortening the unit cell axes from A- to B-lengths. The model position was then improved by using the rigid body refinement option of the program package XPLOR [22] in the resolution range 10–6 Å on our VaxStation-3100. Within 40 cycles the *R*-factor dropped from 49% to 37% showing that the model was placed correctly. In the subsequent refinement we used standard XPLOR protocols on a CRAY-YMP (HLRZ, Jülich), extended the resolution range gradually to 2.1 Å and modified the model several times on the basis of weighted ($F_{obs} - F_{calc}$)- and ($2F_{obs} - F_{calc}$)-maps [23] as well as OMIT-maps [24].

Correspondence address: G.E. Schulz, Institut für Organische Chemie und Biochemie, Albertstr. 21, D-7800 Freiburg i.Br., Germany

During the refinement procedure we repeatedly tried to incorporate peptide sequences. To avoid prejudice, comparison between X-ray and real sequence was carried out with a sequence alignment program using special similarity matrices [25]. When the *R*-factor reached the 30% level in the resolution range 6–2.1 Å, we obtained first reliable signals at the N- and C-termini. From then on the whole 'X-ray-sequence' could be gradually converted to the real sequence of 301 residues. The resolution was then extended to 1.8 Å and bound water molecules and Ca^{2+} ions were added. At present, the *R*-factor is 22% in the range 10–1.8 Å for 2219 protein atoms, 161 water molecules and 2 assigned Ca^{2+} ions. The refinement is being continued.

3. RESULTS AND DISCUSSION

At an *R*-factor of 22% in the range 10–1.8 Å with good geometry and known sequence, the model can be considered final with respect to chain fold and secondary structure. The observed β-sheet topology is identical to the one of the medium resolution model [18]. As sketched in Fig. 1, the 16 β-strands are all antiparallel and form a barrel twisted in the usual right-handed manner with the highest possible 'neighborhood correlation' [26], i.e. the most simple topology. This corroborates the rising suspicion that the chain folds of integral membrane proteins have much simpler topologies than those of water-soluble proteins. The observed 'porin-barrel' seems to be of a more general nature, as the medium resolution model [18] has given clear rotation function signals in the diffraction patterns of crystals of OmpF-[15], PhoE- and LamB-porin [17] of *Escherichia coli* (J. Rosenbusch, H. Jansonius, R. Pauptit, T. Schirmer, D. Tsernoglou, M. Parker, personal communication).

Apart from the 171 β-sheet residues (= 57%) there are 18 residues in 3 short α-helices (= 6%) and 26 residues in hydrogen-bonded reverse turns (= 9%). These numbers may change slightly on further refinement. The lower rim of the barrel as given in Fig. 1 is constructed in a rather regular way. It contains the N- and C-termini, two β-strand connections (= loops) with 3 or more residues, and five adjacent short 2-residue loops. Noticeably, the termini and the two larger loops are part of the trimer interface, whereas the five 2-residue loops are at the membrane-exposed side of the barrel.

In contrast, the upper barrel rim of Fig. 1 contains quite irregularly folded chain segments; no loop is shorter than 3 residues. The largest loop contains 44 residues and runs between the fifth and the sixth β-strands (β5 and β6) into the inside of the barrel where it defines the channel size and shape (see also Fig. 2B of [18]). As can be visualized in Fig. 2A and B, it contains the short helix α2. The second largest loop consists of 20 residues between β11 and β12, it forms an α-helix of 7 residues (α3) and protrudes significantly from the non-polar layer of the membrane (Fig. 2A). The third largest loop contains 13 residues between β3 and β4 and forms another short helix α1; it also participates in interface contacts. The other loops of the upper barrel rim of Fig. 1 contain 4, 12, 9, 3 and 7 residues. Altogether, the upper rim is much less regular than the lower one.

It should be noted that the most substantial correction of the medium resolution structure [18] was applied to loops β7→β8 and β9→β10. All α-helices turned out as right-handed, confirming the chirality derived earlier from the β-strand twists [18]. The weak density region of previous maps [18,20], which had puzzled us

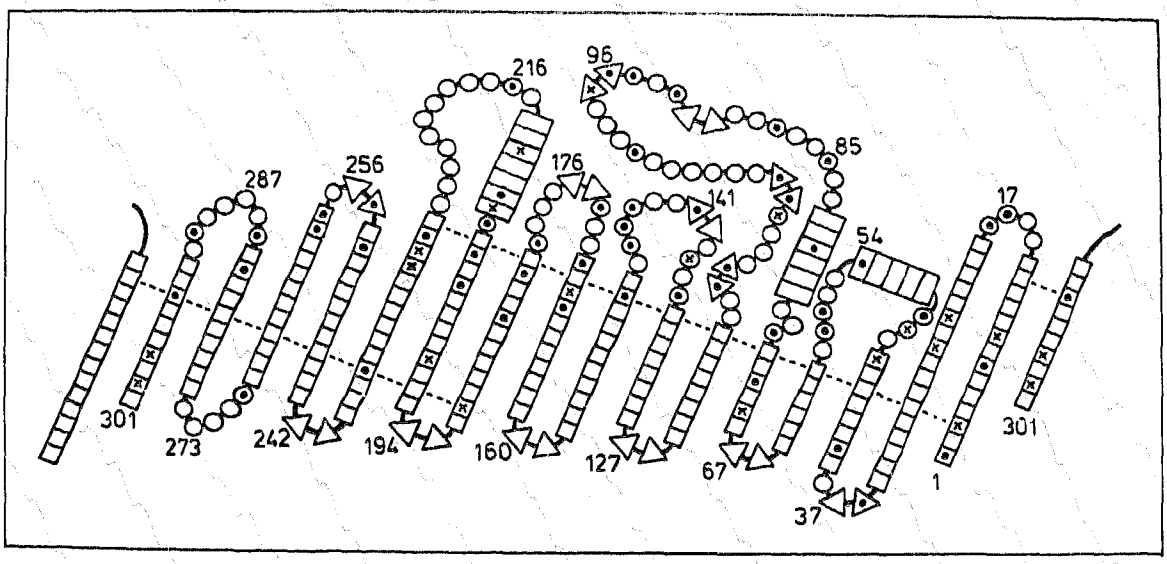

Fig. 1. Topology of the 16-stranded β-sheet barrel of porin as viewed from the barrel outside. The barrel axis is vertical. Care has been taken to correlate the apex residue of each loop with those observed in the spatial structure. The secondary structure assignments are given by appropriate marks (quadrangle = β, rectangle = α, triangle = hydrogen bonded reverse turn). Several residue positions are labelled. Hydrogen bonding between β-strands is parallel to the dashed lines. Marking all Asp and Glu by dots (•) and all His, Lys and Arg by crosses (×) clarifies the charge distribution. The loop in the barrel that defines the channel size runs from residues 74–117. It can also be visualized in Fig. 2B of [18].

in connection with a very high apparent M_r observed in SDS-gel electrophoreses, was clearly an artefact caused by using the wrong M_r in the solvent-flattening procedure.

The β-sheet geometry is given in Fig. 2A. As reported earlier [18], the angles between β-strands and the trimer axis vary from about 60° at the interface to about 30° at the membrane-facing side at the barrel. The trimer contains 3 distinct channels separated from each other by walls with an axial length of about 20 Å. The

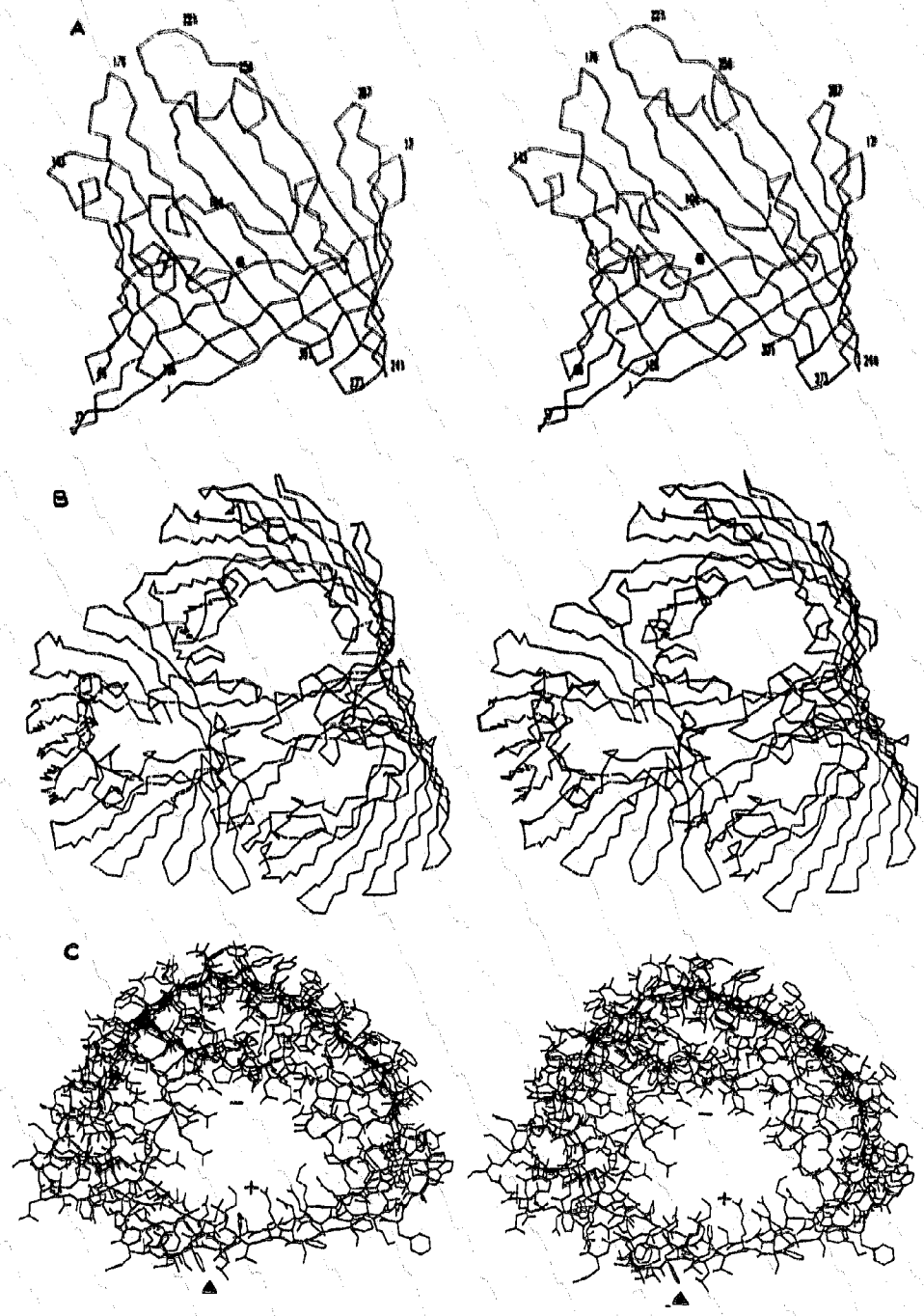

Fig. 2. Stereoviews of the three-dimensional porin structure. (A) C_α-backbone model of a porin monomer of 301 residues as viewed approximately from the trimer axis. Some residue numbers are given. The distance between highest (residue 221) and lowest (residue 37) C_α-atom in this drawing is 58 Å. (B) C_α-backbone model of a porin trimer of 3 × 301 residues as viewed at an angle of about 40° from the trimer axis. The drawing shows the height differences of the walls: between the 3 channels the wall height is only about 20 Å while it is about 40 Å at the outside, i.e. at the interface to the membrane. (C) Complete model (all non-hydrogen atoms) of a porin monomer polypeptide chain viewed along the trimer axis (marked by ▲) perpendicular to the membrane. Six positively charged Lys and Arg side chains below the mark (+) protrude into the channel. A large number of Asp and Glu residues line the other side of the channel forming an arc surrounding the mark (−).

membrane-facing trimer circumference has an axial length of about 40 Å, which is enough to cover completely the approximately 30 Å thick non-polar part of the membrane [27].

The side chains of most of the amino acid residues are positioned unambiguously in the electron density. Their properties follow the expected pattern. Around the equatorial circumference of the trimer there are non-polar side chains contacting the non-polar inner part of the membrane, whereas the side-chains lining the channel are polar. The trimer interface is polar as well. In this interface, we observe the following main chain hydrogen bonds per monomer: 1-N...301'-O (salt bridge), 25-N...53''-O, 36-N...271'-O, 53-O...25'-N, 271-O...36''-N and 301-O...1''-N (salt bridge).

The sequence contains 70 residues with ionogenic side chains, 51 of which are Asp and Glu while there are only 19 His, Lys and Arg. The distribution of the ionogenic side chains is remarkable. If one divides the monomer shown in Fig. 1 at the barrel equator one observes only 7 negative and 7 positive charges in the lower half whereas there are as many as 44 negative and 12 positive charges in the upper half. The large excess of negative charges in the upper half of the barrel is obvious. The channel is electrostatically asymmetric and should therefore have different ionic diffusion rates in both directions.

The orientation of porin in the membrane is not yet clear. Assuming that the porin trimer has a single channel at one side of the membrane and three channels at the other [8], Hoenger et al. [13] found that the three-channel-side of OmpF porin of *E. coli* binds a lipopolysaccharide (LPS) molecule and should therefore face the extracellular space. In our model, the three-channel-side corresponds to the smooth barrel rim at the bottom of Fig. 1 and Fig. 2A. At this side of the three barrels, our model has no suitable binding place for an LPS molecule; the channels are polar, and there is no non-polar pocket at the threefold axis between the barrels that could accommodate even a single aliphatic chain.

Even more conspicuous than the axial charge difference of the channel is a pronounced azimuthal charge separation. There are 6 adjacent Lys and Arg side chains pointing from the channel wall at the interface side into the channel interior. These residues are marked by '+' in Fig. 2C. The remaining arc-like inner surface of the channel consists of a much larger number of negatively charged side chains, the center of which is marked by '−'. The excess of negative charges should give rise to a strong electric field that dominates the fewer positive charges. This is in good agreement with the observation that the 6 Lys and Arg side chains are all adjacent in space and oriented in parallel (Fig. 2C). Since this is an energetically very unfavorable arrangement, one should expect that they are oriented by an electric field. The two identified Ca^{2+} ions are localized within the negatively charged arc of the channel (Fig. 2C), where they are embedded in a large excess of negative charges. We expect to find more Ca^{2+} ions on further refinement.

Acknowledgements: We thank the team of the EMBL outstation at Hamburg for help in data collection and processing. Furthermore, we thank the HLRZ Jülich for granting computing time on a CRAY-YMP. The work was supported by the Graduiertenkolleg Polymerwissenschaften and by the Sonderforschungsbereiche 60 and 176. Coordinates will be deposited in the Protein Data Bank after the refinement has been completed.

REFERENCES

[1] Nikaido, H. and Vaara, M. (1985) Microbiol. Rev. 49, 1–32.
[2] Benz, R. and Bauer, K. (1988) Eur. J. Biochem. 176, 1–19.
[3] Flamman, H.T. and Weckesser, J. (1984) J. Bacteriol. 159, 410–412.
[4] Nabedryk, E., Garavito, R.M. and Breton, J. (1988) Biophys. J. 53, 671–676.
[5] Deisenhofer, J., Epp, O., Miki, K., Huber, R. and Michel, H. (1985) Nature 318, 618–624.
[6] Feher, G., Allen, J.P., Okamura, M.Y. and Rees, D.C. (1989) Nature 339, 111–116.
[7] Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E. and Downing, K.H. (1990) J. Mol. Biol. 213, 899–929.
[8] Engel, A., Massalski, A., Schindler, H., Dorset, D.L. and Rosenbusch, J.P. (1985) Nature 317, 643–645.
[9] Kessel, M., Brennan, M.J., Trus, B.L., Bisher, M.E. and Steven, A.C. (1988) J. Mol. Biol. 203, 275–278.
[10] Sass, H.J., Büldt, G., Beckmann, E., Zemlin, F., van Heel, M., Zeitler, E., Rosenbusch, J.P., Dorset, D.L. and Massalski, A. (1989) J. Mol. Biol. 209, 171–175.
[11] Jap, B.K. (1989) J. Mol. Biol. 205, 407–419.
[12] Rachel, R., Engel, A.M., Huber, R., Stetter, K.-O. and Baumeister, W. (1990) FEBS Lett. 262, 64–68.
[13] Hoenger, A., Gross, H., Aebi, U. and Engel, A. (1990) J. Struct. Biol. 103, 185–195.
[14] Garavito, R.M. and Rosenbusch, J.P. (1980) J. Cell Biol. 86, 327–329.
[15] Garavito, R.M., Jenkins, J., Jansonius, J.N., Karlsson, R. and Rosenbusch, J.P. (1983) J. Mol. Biol. 164, 313–327.
[16] Nestel, U., Wacker, T., Woitzik, D., Weckesser, J., Kreutz, W. and Welte, W. (1989) FEBS Lett. 242, 405–408.
[17] Stauffer, K.A., Page, M.G.P., Hardmeyer, A., Keller, T.A. and Pauptit, R.A. (1990) J. Mol. Biol. 211, 297–299.
[18] Weiss, M.S., Wacker, T., Weckesser, J., Welte, W. and Schulz, G.E. (1990) FEBS Lett. 267, 268–272.
[19] Kreusch, A., Weiss, M.S., Welte, W., Weckesser, J. and Schulz, G.E. (1991) J. Mol. Biol. 217, 9–10.
[20] Weiss, M.S., Wacker, T., Nestel, U., Woitzik, D., Weckesser, J., Kreutz, W., Welte, W. and Schulz, G.E. (1989) FEBS Lett. 256, 143–146.
[21] Kreusch, A. (1990) Diploma thesis, Universität Freiburg i.Br.
[22] Brünger, A.T., Kuriyan, D. and Karplus, M. (1987) Science 35, 458–460.
[23] Read, R.J. (1986) Acta Crystallogr. sect. A 42, 140–149.
[24] Bhat, T.N. and Cohen, G.H. (1984) J. Appl. Crystallogr. 12, 794–805.
[25] Thieme, R., Pai, E.F., Schirmer, R.H. and Schulz, G.E. (1981) J. Mol. Biol. 152, 763–782.
[26] Schulz, G.E. and Schirmer, R.H. (1978) Principles of Protein Structure, pp. 84–87, Springer-Verlag, New York.
[27] Lewis, B.A. and Engelman, D.M. (1983) J. Mol. Biol. 166, 211–217.