Acta Cryst. (1994). D**50**, 375–379

Fusion Proteins as Tools for Crystallization: the Lactose Permease from *Escherichia coli*

BY GILBERT G. PRIVÉ*

Molecular Biology Institute, and Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA 90024-1570, *USA, and Howard Hughes Medical Institute, UCLA, Los Angeles, CA* 90024-1662, *USA*

GILLIAN E. VERNER, CINDY WEITZMAN AND KEVIN H. ZEN

Howard Hughes Medical Institute, UCLA, Los Angeles, CA 90024-1662, *USA*

DAVID EISENBERG

Molecular Biology Institute, and Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA 90024-1570, *USA*

AND H. RONALD KABACK

Howard Hughes Medical Institute, UCLA, Los Angeles, CA 90024-1662, *USA*

(*Received* 21 *November* 1993; *accepted* 20 *December* 1993)

Abstract

A novel strategy is presented for the crystallization of membrane proteins or other proteins with low solubility and/or stability. The method is illustrated with the lactose permease from *Escherichia coli*, in which a fusion is constructed between the permease and a 'carrier' protein. The carrier is a soluble, stable protein with its C and N termini close together in space at the surface of the protein, so that the carrier can be introduced into an internal position of the target protein. The carrier is chosen with convenient spectral or enzymatic properties, making the fusion protein easier to handle than the native molecule. Data are presented for the successful construction, expression and purification of a fusion product between lactose permease and cytochrome b_{562} from *E. coli*. The lactose transport activity of the fusion protein is similar to that of wild-type lactose permease, and the fusion product has an absorption spectrum in the visible range which is essentially identical to that of cytochrome b_{562}. The fusion protein has a higher proportional polar surface area than wild-type permease, and should have better possibilities of forming the strong directional intermolecular contacts required of a crystal lattice.

Introduction

Lactose (lac) permease of *Escherichia coli* is part of a large family of integral membrane proteins responsible for the transport of polar molecules across membranes. It is one of the most studied transporters, and represents an enormous class of proteins that are found in essentially all biological membranes. Almost all are thought to have the same basic structural arrangement: a series of 12 transmembrane α-helices crossing the membrane in a zigzag fashion (Griffith *et al.*, 1992). Although extensively studied, little is known about the tertiary structure of transporters, and the mechanism of action of these proteins remains largely a matter of speculation.

The most severe limitation in obtaining structural information on the lac permease is the lack of crystals. Like most other transporters, the lac permease is an extremely hydrophobic protein, having only short hydrophilic loops connecting the hydrophobic helices. Having only a small polar surface area, most of the mass of the lac permease is embedded within the membrane. Once solubilized, this protein is presumably surrounded by a ring of detergent that covers much of the surface area of the protein. Since the lattices of well diffracting protein crystals are generally made up from intermolecular contacts between polar groups on the protein surfaces, the lac permease seems to be a poor candidate for forming well ordered crystals. Indeed, many previous attempts at crystallizing lac permease have failed.

In this paper, we present a novel approach for the crystallization of extremely hydrophobic membrane proteins. The strategy consists of engineering a fusion product between the target hydrophobic pro-

* Corresponding author.

ISSN 0907-4449

tein and a stable, soluble 'carrier' protein. The fusion protein thus generated has a higher proportional polar surface area, and should have a better chance of forming strong directional intermolecular contacts. In a sense, these proteins are designed to mimic the bacterial photosynthetic reaction centers, which naturally have large extra-membraneous domains and can be induced to crystallize (Michel, 1982; Allen & Feher, 1984).

This basic strategy can be applied to any protein, soluble or membrane-bound, that resists crystallization because of stability and/or solubility problems. The use of fusion-protein technology has been used extensively to generate more stable, more soluble and more readily purified proteins (Nilsson, Forsberg, Moks, Hartmanis & Uhlén, 1992), and here we discuss the potential of fusion proteins as tools for crystallization. The method is illustrated with the successful construction of a fusion between the lac permease and cytochrome b_{652}. This particular carrier was chosen for reasons described below, and is but one of several possible choices. In constructing a set of fusions between the protein of interest and various carriers, the nature of the protein itself can be treated as a variable in the search for crystallization conditions.

Experimental

A cassette version of the *lacY* gene (EMBL-X56095) under control of the *lac* promoter/operator was used as the plasmid vector. This *lacY* gene has an *Xho*I restriction site in a position encoding the middle cytoplasmic loop (*L*6), and contains a 3′ addition encoding a biotinylation domain used for purification by monomeric avidin affinity chromatography (Consler *et al.*, 1993). The cytochrome b_{562} gene was obtained from *E. coli* genomic DNA by the polymerase chain reaction (PCR). The primers were based on the Genebank sequence entry ECCYBC (Trower, 1993), and were designed so that the amplified fragment would have *Xho*I restriction sites at both ends. Both the vector and the PCR fragment were digested with *Xho*I, and the two were then joined with T4 DNA ligase. The first and second junctions were constructed with the amino-acid sequences DAPSSA*ADLE* and *QKYR*SFLISSSATV, where the non-italicized letters refer to residues contributed by the permease, underlined letters are additional 'linker' residues included as a result of the cloning strategy, and the italicized letters are amino acids from the cytochrome. Various clones were tested for the proper orientation of the insert by restriction digestion analysis, and the correctness of the final selected construct was confirmed by DNA sequencing across both *Xho*I junction sites. Transport activity measurements were performed as described in Consler *et al.* (1993).

Expression of the fusion product in *E. coli* and the preparation of urea/cholate washed membranes were performed as described for the wild-type protein by Viitanen, Newman, Foster, Wilson & Kaback (1986). Protein was solubilized from the membranes with 3–5% dodecyl maltoside, and the biotinylated permease was purified from the crude membrane extract by affinity chromatography on monomeric avidin columns (Consler *et al.*, 1993). Purification buffers generally contained 1 m*M* dithiothreitol (DTT) to prevent the oxidation of cysteine side chains on the permease; this also maintained the cytochrome heme in the reduced form. For the spectra shown in Fig. 4, reducing agents were not used in the chromatography buffers, so that the oxidized spectrum of the pure protein could be recorded. DTT was then added up to 1 m*M*, and the reduced spectrum was measured.

Design of the fusion

The 30-amino-acid central loop of the permease separates the first six transmembrane helices from the last six, and was chosen as the site for inserting a hydrophilic 'carrier' domain. An internal loop was chosen over a simple N- or C-terminal fusion, since the latter would generate 'dumbbell' shaped molecules having two domains linked by a (presumably) flexible hinge. Since the target is to make a solid compact structure with as few internal degrees of freedom as possible, a fusion product with two links between the partners (*i.e.* internal fusion within a loop) best satisfies this requirement (Fig. 1).

Several factors were considered in the selection of the carrier protein. A checklist of desirable features in a carrier included the following: (1) Soluble monomer, single domain. (2) Stable compact structure; known to crystallize, structure in the Protein Data Bank (PDB; Bernstein *et al.*, 1977). (3) C and N termini close together and at the surface of the protein (for an 'internal' fusion within a loop). (4) No disulfide bonds (the permease is stored in DTT). (5) *E. coli* protein – easily cloned and expressed. (6) Size – larger proteins would presumably be better

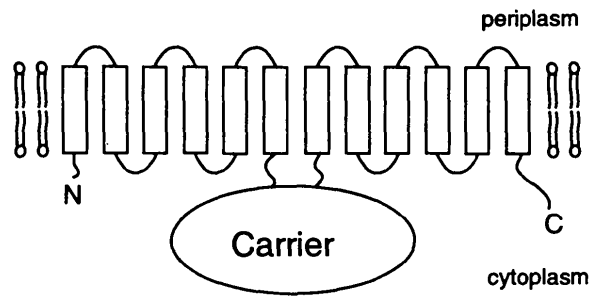

Fig. 1. Schematic of the fusion between the carrier protein and the lac permease.

since they would tend to dominate the behavior of the system. (7) Enzymatic activity and/or color – as an aid to monitoring the protein stability and integrity. A colored protein has the added advantage that its behavior is easily monitored in crystallization trials. (8) Affinity purification system – purification of the fusion is simplified if the carrier can be used for affinity chromatography.

Several potential candidates were found from a survey of the structures in the PDB (Bernstein *et al.*, 1977). Further consideration of the specific cloning requirements in the lac permease vector system led to the choice of cytochrome b_{562} from *E. coli*. This 106-amino-acid protein folds as a four-α-helix bundle in which a heme group is tightly bound in the core of the molecule (Fig. 2). The protein is readily crystallized (Itagaki & Hager, 1966), its structure has been solved (Mathews, Bethge & Czerwinski, 1979; Lederer, Glatigny, Bellamy & Mathews, 1981) and refined against 1.4 Å diffraction data (PDB entry 256B). The Cα—Cα distance between residues 3 and 106 is only 11.1 Å, so that insertion of the cytochrome into the central loop of the lac permease should cause only minor strain on either structure. (The exact structural nature of the permease/cytochrome linkage is unknown, but the residues introduced at the junction sites are expected to add some flexibility in linking the two molecules together.) This protein has additional features which lead to its selection over other candidates. Cytochrome b_{562} readily binds ferriprotoporphyrin IX *via* a non-covalent interaction in which the heme iron is coordinated by a methionine and histidine, giving the protein a red color and a distinctive absorption spectrum in the visible range. This cytochrome can be overexpressed in *E. coli* (Nikkila, Gennis & Sligar, 1991; Trower, 1993) and is relatively heat stable [the midpoint transition temperatures for the reduced, oxidized and apo forms of the protein are 354, 340 and 325 K, respectively (Feng & Sligar 1991; Fisher 1991)]. The protein still retains a well defined structure even in the apo form. Although cytochrome b_{562} is normally targeted to the periplasm, removal of the cleavable signal sequence prevents the export of the protein, yielding protein which is indistinguishable from the mature wild-type material (Nikkila *et al.*, 1991). Cytochrome b_{562} does not contain any cysteine residues, which is important for the purposes described here, since lac permease contains eight cysteines, all of which are thought to be in the reduced state. Thus, the presence of disulfide bonds in the carrier protein would be incompatible with the presence of reducing agents required to maintain lac permease stability.

Cytochrome b_{562} does not satisfy item (8) from the above checklist. Instead, we chose to express the *lacY*/L6cyt fusion protein with a 100-amino-acid biotinylation domain at the C terminus for purification purposes. A factor Xa protease site is engineered into the region linking the permease to the biotinylation domain, so that this domain can be removed following affinity purification on a monomeric avidin column (Consler *et al.*, 1993). This protein is referred to here as *lacY*/L6cyt/CXB. We have also expressed a version of the fusion protein with a tag of six consecutive histidines at the permease C-terminus instead of the biotinylation domain for column purification on a NiNTA resin. Details of the purification of the 6His protein will be published elsewhere (Privé & Kaback, in preparation).

A potential weakness in the use of cytochrome b_{562} as a carrier is its small size [see item (6) above]. It is difficult to predict *a priori* which choice may be best suited as an aid in crystallizing the permease, and although it seems reasonable to assume that a larger

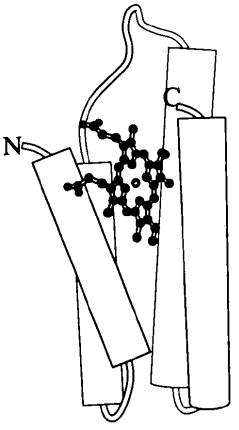

Fig. 2. Structure of cytochrome b_{562}. This schematic (Kraulis, 1991) was created using the coordinates of entry 256B of the PDB (Bernstein *et al.*, 1977).

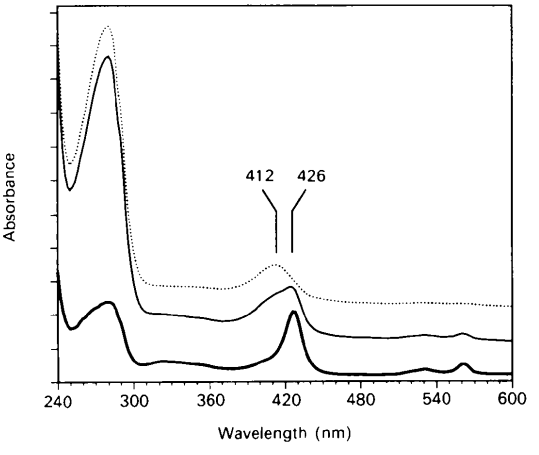

Fig. 3. Overexpression of the permease fusion. Dotted line: spectrum of a dodecyl maltoside extract from *E. coli* membranes without induction of the *lacY*/L6cyt/CXB gene. Thin line: same as above only after induction with 0.2 m*M* IPTG. Thick line: purified *lacY*/L6cyt/CXB after monomeric avidin column chromatography.

carrier would be better, a small protein like cytochrome b_{562} may in fact prove to be a good choice. The many other positive features of this protein lead to its selection, but it is by no means the only possibility. Just as it is impossible to predict the correct pH or salt concentration required for the crystallization of a given protein, the 'crystallization potential' of a particular fusion construct must simply be tested experimentally.

Results

Since cytochrome b_{562} is a soluble protein normally found in the *E. coli* periplasm, membrane preparation from cells expressing the *lacY*/L6cyt/CXB fusion are free from any endogenous cytochrome b_{562}. The distinctive spectrum of this particular cytochrome makes for an easy spectroscopic assay of the abundance of the fusion protein in crude membrane extracts.* This cytochrome is readily reduced in 1 m*M* DTT, leading to a absorbance maximum at 426 nm from the Soret band. Preparations of similar membrane extracts in 1 m*M* DTT from cells not induced for *lacY*/L6cyt/CXB expression have a single, symmetrical peak at 412 nm (Fig. 3). Based on both the 426/412 nm and 280/426 nm ratios one can estimate the level of expression of the protein from crude extracts and readily quantitate the purity of the final preparation.

Authentic cytochrome b_{562} has absorption bands at 562, 531, 427 and 324 nm in the reduced state, and at 418 and 363 nm in the oxidized form (Itagaki & Hager, 1966). This agrees with the maxima seen in the visible spectrum of purified *lacY*/L6cyt/CXB (Fig. 4). The shape, positions and relative heights of the peaks in the spectrum of *lacY*/L6cyt/CXB are in excellent agreement with published spectra. The only significant difference between the two is the peak at 280 nm in *lacY*/L6cyt/CXB due to the abundance of aromatic residues found in the lac permease. Cytochrome b_{562} alone absorbs only very weakly at 280 nm, since it contains very few aromatic residues.

The measured midpoint reduction potential for *lacY*/L6cyt/CXB in a buffer containing 0.02% dodecyl maltoside, 200 m*M* NaCl and 20 m*M* Hepes pH 7.4 is approximately +150 mV (D. Knaff, personal communication), which is similar to the value of +113 mV for wild-type cytochrome b_{562} in 0.25 m*M* Tris buffer pH 7.0 (Itagaki & Hager, 1966).

The lactose transport activity of *lacY*/L6cyt/CXB in intact cells is very similar to that of wild-type lac permease, with respect to both initial rate and steady-state level of lactose accumulation (Fig. 5). Thus, the transport activity of the fusion protein is not significantly impaired by the presence of the fusion domains in the central loop and at the C terminus.

After purification on monomeric avidin columns, *lacY*/L6cyt/CXB runs as a single broad band on silver-stained SDS–PAGE gels, with an apparent molecular weight (M_r) of approximately 63 kDa (data not shown). The calculated molecular weight (MW) is 69 kDa, and this anomalous migration is consistent with the differences seen in wild-type lac permease (M_r = 33 kDa; MW = 46 kDa) and permease carrying only the C-terminal biotinylation domain (M_r = 48 kDa; MW = 57 kDa). In preliminary experiments, *lacY*/L6cyt/CXB has resisted

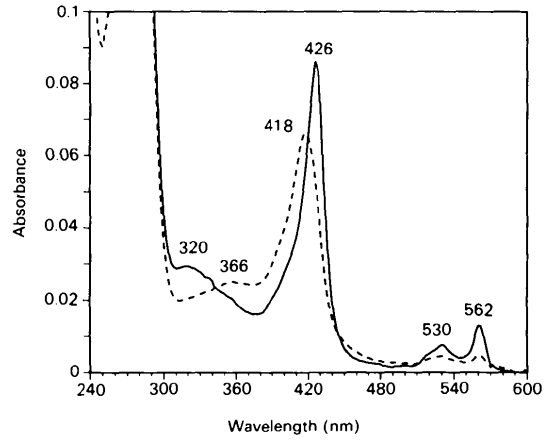

Fig. 4. Spectra of purified *lacY*/L6cyt/CXB in the oxidized (dashed line) and reduced (solid line) states.

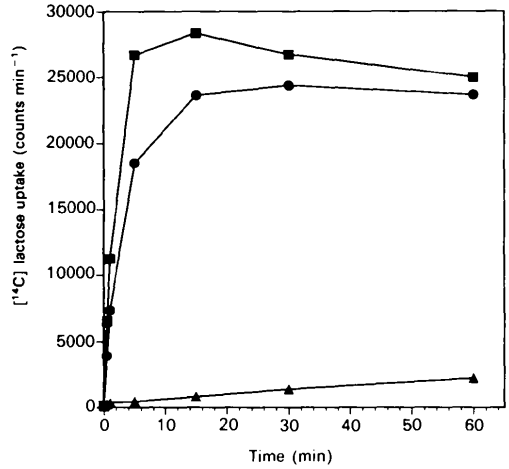

Fig. 5. Transport activity of *lacY*/L6cyt/CXB. Uptake of [1-^{14}C]lactose was measured in *E. coli* transformed with a vector without a *lacY* gene (pT7-5, ▲), a vector for the wild-type *lacY* gene (■), and a vector carrying the *lacY*/L6cyt/CXB gene (●). The fusion protein has essentially the same activity as wild-type lac permease.

* The spectrophotometric measurements reported here assume that all of *lacY*/L6cyt/CXB protein contains bound heme. This was verified by trial additions of excess protohemin, which should readily reconstitute any apo protein (Itagaki & Hager, 1966). The addition of protohemin had no effect on the intensity of the cytochrome absorbance peaks.

cleavage of the C-terminal biotinylation domain by factor Xa, and so the entire intact protein has been used in the early crystallization trials.

Lac permease tends to form aggregates once it is solubilized, and a great advantage of working with solutions of the colored fusion protein is that it is immediately obvious if the protein has precipitated. The fusion appears to be more stable than the wild-type protein, and solutions of *lacY*/L6cyt/CXB are stable at 1–4 mg ml^{-1} for over a week at 277 K in 0.02% dodecyl maltoside as judged by the red color of the solution, and the lack of aggregates seen by SDS–PAGE. Long-term stability of a protein is usually considered as a prerequisite for crystallization, and this fact alone makes the fusion protein a better candidate for crystallization than the wild-type permease.

In preliminary crystallization trials, it can easily be seen if the *lacY*/L6cyt/CXB is in solution, has precipitated, or has denatured. The requisite that protein crystals be red greatly simplifies the interpretation of the experiments, since salt or detergent crystals can be immediately identified by their lack of color.

Discussion

The *lacY*/L6cyt/CXB fusion product can be expressed in large quantities and is readily purified. Each part of the fusion product behaves completely normally, based on the cytochrome absorption spectrum and lactose transport activity. All parts of the fusions are thus correctly folded, and in particular, the cytochrome forms stable complexes with ferroprotoporphyrin IX *in vivo*.

There are many examples where the removal of a flexible domain or the truncation of a chain terminus has been used as a successful strategy in the crystallization of proteins that would otherwise not crystallize. As an alternative to the removal of an unfavorable portion of a target molecule, the addition of favorable components has proven to be an excellent approach. Notably, muscle actin could only be crystallized as a 1:1 complex with DNase I from *E. coli* (Suck, Kabsch & Mannherz, 1981), and crystals of human immunodeficiency virus reverse transcriptase were grown as a complex with an Fab fragment and a DNA oligomer (Jacobo-Molina *et al.*, 1991). The fusion-protein strategy has many features in common with these approaches, except that stoichiometry and binding affinity are not a concern, since in a fusion the two proteins are locked together by covalent bonds. In addition, the repertoire of possible fusion products is much greater than of the antibody complexes, and features such as color or enzymatic activity can be engineered into the fusion protein.

Both the nature of the carrier protein and the position of the fusion construct are being tested as variables in the search for crystallization conditions. With the lac permease, the hydrophilic loops between the transmembrane helices are obvious choices as carrier insertion points. The particular loop, the exact position within the loop, and the addition of 'linker' polypeptide between the permease and the carrier may be just as important as the choice of the carrier itself. Long linkers would tend to make the fusion protein more flexible and probably less likely to crystallize, while a direct junction between target and carrier may distort one or the other protein, and not allow the two proteins to explore enough of their partner's surface in order to find favorable non-bonded contacts. Short linkers seem to be the best choice in allowing the two halves of the fusion product to relax enough to fold properly, and yet keep the entire particle as compact as possible.

GGP was supported by a postdoctoral fellowship from the American Cancer Society, California Division and by Grant GM-31599. We thank David Knaff for the midpoint reduction potentials on *lacY*/L6cyt/CXB.

References

ALLEN, J. P. & FEHER, G. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 4795–4799.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MAYER, E. F., BRYCE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **122**, 535–542.

CONSLER, T. G., PERSSON, B. L., JUNG, H., ZEN, K. H., JUNG, K., PRIVÉ, G. G., VERNER, G. E. & KABACK, H. R. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 6934–6938.

FENG, Y. & SLIGAR, S. G. (1991). *Biochemistry*, **30**, 10150–10155.

FISHER, M. T. (1991). *Biochemistry*, **30**, 10012–10018.

GRIFFITH, J. K., BAKER, M. E., ROUCH, D. A., PAGE, M. G., SKURRAY, R. A., PAULSEN, I. T., CHATER, K. F., BALDWIN, S. A. & HENDERSON, P. J. (1992). *Curr. Opin. Cell Biol.* **4**, 684–695.

ITAGAKI, E. & HAGER, L. P. (1966). *J. Biol. Chem.* **241**, 3687–3695.

JACOBO-MOLINA, A., CLARK, A. D. JR, WILLIAMS, R. L., NANNI, R. G., CLARK, P., FERRIS, A. L., HUGHES, S. H. & ARNOLD, E. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 10895–10899.

KRAULIS, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

LEDERER, F. A., GLATIGNY, P. H. B., BELLAMY, H. D. & MATHEWS, F. S. (1981). *J. Mol. Biol.* **148**, 427.

MATHEWS, F. S., BETHGE, P. H. & CZERWINSKI, E. W. (1979). *J. Biol. Chem.* **254**, 1699–1706.

MICHEL, H. (1982). *J. Mol. Biol.* **158**, 567–572.

NIKKILA, H., GENNIS, R. B. & SLIGAR, S. G. (1991). *Eur. J. Biochem.* **202**, 309–313.

NILSSON, B., FORSBERG, G., MOKS, T., HARTMANIS, M. & UHLÉN, M. (1992). *Curr. Opin. Struct. Biol.* **2**, 569–575.

SUCK, D., KABSCH, W. & MANNHERZ, H. G. (1981). *Proc. Natl Acad. Sci USA*, **78**, 4319–4323.

TROWER, M. K. (1993). *Biochem. Biophys. Acta*, **1143**, 109–111.

VIITANEN, P., NEWMAN, M. J., FOSTER, D. L., WILSON, T. H. & KABACK, H. R. (1986). *Methods Enzymol.* **125**, 429–452.