ELSEVIER

Biochimica et Biophysica Acta 1564 (2002) 38–46

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

www.bba-direct.com

Insertion of carrier proteins into hydrophilic loops of the *Escherichia coli* lactose permease

Christian K. Engel [a,1], Lu Chen [a], Gilbert G. Privé [a,b,*]

[a] *Division of Molecular and Structural Biology, Ontario Cancer Institute, 610 University Avenue, Toronto, Ontario, Canada M5G 2M9*
[b] *Departments of Medical Biophysics and Biochemistry, University of Toronto, 610 University Avenue, Toronto, Ontario, Canada M5G 2M9*

Received 26 February 2002; accepted 7 March 2002

Abstract

We describe the design and characterization of a set of fusion proteins of the *Escherichia coli* lactose (lac) permease in which a set of five different soluble ''carrier'' proteins (cytochrome$_{b562}$, flavodoxin, T4 lysozyme, β-lactamase and 70 kDa heat shock ATPase domain) were systematically inserted into selected loop positions of the transporter. The design goal was to increase the exposed hydrophilic surface area of the permease, while minimizing the internal flexibility of the resulting fusion proteins in order to improve the crystallization properties of the membrane protein. Fusion proteins with insertions into the central hydrophilic loop of the lac permease were active in transport lactose, although only the fusion proteins with *E. coli* cytochrome$_{b562}$, *E. coli* flavodoxin or T4 lysozyme were expressed at near wild-type lac permease levels. Eight other loop positions were tested with these three carriers, leading to the identification of additional fusion proteins that were active and well-expressed. By combining the results from the single carrier insertions, we have expressed functional ''double fusion'' proteins containing cytochrome$_{b562}$ domains inserted in two different loop positions.

Keywords: Lactose permease; Fusion protein; Membrane protein; Crystallization

1. Introduction

The central difficulty in the structural analysis of membrane proteins by diffraction methods is the production of well-ordered crystals. In order to understand how detergent-solubilized membrane proteins form three-dimensional crystals, it is essential to consider the properties of the fundamental particle that assembles into the crystalline array, the protein–detergent complex (PDC). The PDC consists of two distinct types of surfaces: a central micelle-like detergent belt that protects the hydrophobic transmembranous segments of the protein from the bulk aqueous solvent, and two flanking extramembranous caps which have surface properties that are similar to those of soluble proteins. Most of the α-helical membrane proteins that have been crystallized from detergent solutions contain large hydrophilic domains, and in these cases, it is commonly observed that most of the intermolecular contacts that make up the crystalline lattice are due to polar interactions between these surfaces (reviewed in Refs. [1,2]), although exceptions are known [3,4]. The central detergent-coated belt region of the PDC is generally considered to have a flexible and dynamic surface since it is formed largely by the amphiphile headgroups. Although detergent–detergent contacts may be important in the crystal [3,5], these are not usually the ordered, regular determinants of the lattice. Thus, membrane proteins that have large extramembranous domains are considered to be promising candidates for crystallization because the lattice contacts between these extended polar domains can result in the formation of large cavities that can accommodate the detergent-coated domains.

We wish to broaden the scope membrane proteins that can be studied by X-ray crystallography to include proteins that do not inherently contain large polar domains. In particular, proteins that consist of a series of transmembrane α-helices linked only by relatively short polar loops, such as many of the proteins involved in secondary transport [6], do not have an architecture with high potential for crystal formation. We have used the *Escherichia coli* lactose (lac) permease as a

Abbreviations: ADA, adenosine deaminase; ECL, enhanced chemiluminescence; HSC, 70 kDa heat shock protein ATPase domain; IPTG, isopropyl β-D-thiogalactoside; PDB, Protein Data Bank; PDC, protein–detergent complex; TEM, β-lactamase

* Corresponding author. Tel.: +1-416-946-2971; fax: +1-416-946-6529.
E-mail address: prive@uhnres.utoronto.ca (G.G. Privé).
[1] Current address: Aventis Pharma Deutschland GmbH, D-65926 Frankfurt am Main, Germany.

0005-2736/02/$ - see front matter
PII: S0005-2736(02)00398-X

model system to develop methods to optimize the surfaces of membrane proteins for crystallization studies.

The lac permease is a monomeric protein consisting of 12 transmembrane α-helices connected by short hydrophilic loops, and catalyses the 1:1 coupled symport of protons and β-galactosides (reviewed in Ref. [7]). A battery of techniques that combine molecular biology with various biochemical and biophysical approaches have culminated in a packing model and a reaction mechanism [8], but the direct visualization of the protein at the atomic level has resisted all efforts to date. Because the permease lacks large extramembranous domains, we are pursuing a strategy to re-engineer the surface features of this transporter. The design goal is to create a PDC with a larger exposed polar surface area relative to the parent protein with the aim of increasing the potential for crystallization.

We have made a series of *internal* fusion proteins of the lac permease with various "carrier" proteins in which the carrier domain is inserted into the selected hydrophilic loops between successive transmembrane helices of the permease. A carrier is an unrelated soluble, monomeric protein whose 3D structure is available in the Protein Data Bank (PDB), and whose N- and C-termini are close together in space at the surface of the protein [9,10]. Our fusion strategy shares some features of using complexes of membrane proteins with antibody fragments for crystallization [11–16], although in our case, the surface engineering is done at the genetic level and thus allows for a wider diversity in the resulting PDCs.

A key design feature of these fusion proteins is that the carrier is introduced into a loop position of the target membrane protein, and we have inserted a series of candidate carrier proteins into each of nine different hydrophilic loops of the lac permease (Fig. 1). More traditional "head-to-tail" C- or N-terminal fusion proteins rarely crystallize, even when the two partners are soluble proteins, probably because of the dynamic and flexible nature of the fusion. In contrast, the internal fusion design generates particles that are internally rigid by virtue of the two peptidic linkages that connect the pair of domains. As a result, an internal fusion protein is expected to have fewer internal degrees of freedom then a head-to-tail fusion, and is thus more likely to crystallize. A preliminary "proof-of-principle" study of an internal fusion between two soluble proteins, the maltose binding protein and cytochrome$_{b562}$, has produced crystals that diffract to 2.2 Å (CKE, V. Ahn and GGP, in preparation, see also Ref. [17]). The use of head-to-tail fusions for the crystallization of membrane proteins has also been described [18,19]. A prototype of an internal carrier fusion protein containing a cytochrome$_{b562}$ insertion in the central hydrophilic loop of the lac permease [9,10] ("red permease") has produced well-ordered two-dimensional crystals resulting in a projection map calculated at a resolution of 2 nm [20]. In this paper, we describe the construction and characterization of an extended series of internal carrier fusion proteins of the lac permease. The biochemical properties of three of these fusion proteins in detergent solutions are investigated in the accompanying paper [21].

2. Materials and methods

2.1. Materials

[1-^{14}C]-Lactose and the enhanced chemiluminescence (ECL) detection kit were purchased from Amersham. Horseradish peroxidase conjugated anti-mouse IgG was obtained from Sigma, and Ni-NTA resin and Penta-His antibody (mouse anti-$(H)_5$) were from Quiagen. Deoxyoligonucleotides were purchased from ACGT Corp. (Toronto, Ontario). All site-directed mutagenesis was carried out with the Quikchange protocol (Stratagene).

2.2. Insertion of carrier proteins into the central loop of the permease

The template expression vector used in this study consisted of a cassette version of the lac permease gene under the control of a lac promoter (vector EMBL-X56095) with 10 additional histidine codons at the 5′ end of the gene. The histidine-tagged protein with no carrier insert is referred to as "wild-type." The construction of a lactose permease fusion protein with an *E. coli* cytochrome$_{b562}$ inserted into the central hydrophilic loop (red permease, "L6_cyt_N2C6") has been described [9,10]. In this plasmid, the cytochrome$_{b562}$ insert had flanking 5′ and 3′ *Xho*I restriction enzyme sites. To construct the other L6 expression vectors described in this paper, the 3′ *Xho*I site from the L6_cyt_N2C6 gene was replaced with a *Sac*I site, and the DNA for the inserts were amplified by PCR to introduce fragments with the appropriate 5′ *Xho*I and 3′ *Sac*I sites. Vector and PCR produces were digested with *Xho*I and *Sac*I and ligated. All constructs were confirmed by dideoxynucleotide sequencing over the full length of the fusion protein coding region. Five codons corresponding to the C-terminal linker in the L6_cyt_N2C6 vector were deleted to create the L6_cyt_N2C1 construct. For the other the carrier domains used in this study, the source of the template DNA for the PCR reactions were as follows: (i) *E. coli* flavodoxin, Met1 to His171 [22], amplified from *E. coli* K12 genomic DNA; (ii) bacteriophage T4 lysozyme, Met1 to Leu164 [23], amplified from a plasmid provided by B. Matthews (U Oregon); (iii) murine adenosine deaminase, Ala6 to Gln352 [24], amplified from I.M.A.G.E. consortium clone 577879; (iv) the ATPase domain from the 70 kDa human heat shock protein, Gly4 to Gly382 [25], amplified from I.M.A.G.E. consortium clone 565948; (v) *E. coli* TEM1 β-lactamase, His26 to Trp290 [26], amplified from plasmid pET32a DNA (Novagen). The exact sequence boundaries for the carriers were chosen from an inspection of the appropriate PDB entry in order to identify the closest approach between the ordered residues at each termini. The construction of the coding

regions for these fusion proteins introduced single residue linkers between the carrier and the lac permease, and are referred to as the "N1C1" series of constructs. Since the exact structure of the lactose permease central hydrophilic loop is unknown, and since linkers that are too short could destabilize the internal fusion proteins, a second set of fusion proteins was constructed that contained longer five-residue serine/glycine linkers at both the N- and C-terminal ends of the carrier domains. These constructs are referred to as the "N5C5" series.

2.3. Carrier insertions into other permease loops

In order to insert carrier domains into loops other than the central hydrophilic domain, the unique *Xho*I site in the original 10-His lac permease cassette gene was replaced with a *Sal*I site, resulting in a substitution of serine 194 to a threonine in the central hydrophilic loop. This vector was then used to generate a series of vectors in which tandem *Xho*I and *Sac*I sites were introduced into 8 of the remaining 10 hydrophilic loops that connect successive transmembrane helices. The mutagenesis resulted in the following amino acid changes in the permease sequence: L1: 36-DISSIRSSD-44 ; L2: 70-LGSSISSSLRK-73; L3: 101-YNSSVSSIV-109; L5: 159-GISSTSSNQ-167; L7: 251-FASSEQSSRV-260; L9: 309-ATSSLESSIL-318; L10: 337-ITSSFESSFS-346; L11: 370-GNSSESSGF-378 (residue numbering based for the wild-type sequence, with amino acid substitutions indicated by singly underlined residues and insertions indicated by double underlines). These vectors are referred to as the parent L*x* vectors, where *x* refers to the designated hydrophilic domain. Inserts for the cytochrome$_{b562}$ (N2C1 linkers), T4 lysozyme (N1C1 and N5C5 linkers) and flavodoxin (N1C1 and N5C5 linkers) carriers were prepared by digestion of the L6 constructs described in the section above with both *Xho*I and *Sac*I restriction enzymes, and ligated into the L*x* parent vectors prepared by reaction with the same two enzymes. All constructs were verified by DNA sequencing.

2.4. Shortened central loop

Two constructs were made to test the effect of deleting residues from the central hydrophilic domain of the permease. These were L6_cyt_Δ11, corresponding to a deletion of permease residues Ala195-His205, and L6_cyt_Δ19, corresponding to a deletion of residues Ala195-Ala213 (wild-type lac permease numbering).

2.5. Double fusion constructs

Lac permease "double fusion" proteins containing cytochrome$_{b562}$ inserts at positions L1 and L6, and L2 and L6 were constructed by introducing a fragment digested with *Xho*I and *Hind*III from the L6_cyt_N2C1 vector into either the L1_cyt_N2C1 or L2_cyt_N2C1 plasmid prepared by digestion with *Sal*I and *Hind*III. An L6/L10 double cytochrome$_{b562}$ fusion protein was made with a similar strategy by combining *EcoR*I/*Afl*II fragments of the L6_cyt_N2C1 and L10_cyt_N2C1 plasmids.

2.6. Lactose transport

E. coli T184 cells [27] were transformed with plasmids containing lac permease constructs, grown to an A_{600} of 1.0 and induced with 0.2 mM IPTG for 2 h. Cells transformed with a similar plasmid without the lac permease gene served as a negative control, and did not display any measurable transport. Cells were washed in 100 mM potassium phosphate pH 7.5, 10 mM $MgSO_4$ and resuspended to an A_{420} of 10.0 in the same buffer. [1-^{14}C]-Lactose was added to a final concentration of 0.4 mM, and cells were collected by rapid filtration at specified time points. Initial rates of lactose uptake were measured at 1 min, and steady state accumulation was measured at 60 min. The levels of accumulated lactose was measured by liquid scintillation counting of the filters [28]. All measurements were done in triplicate.

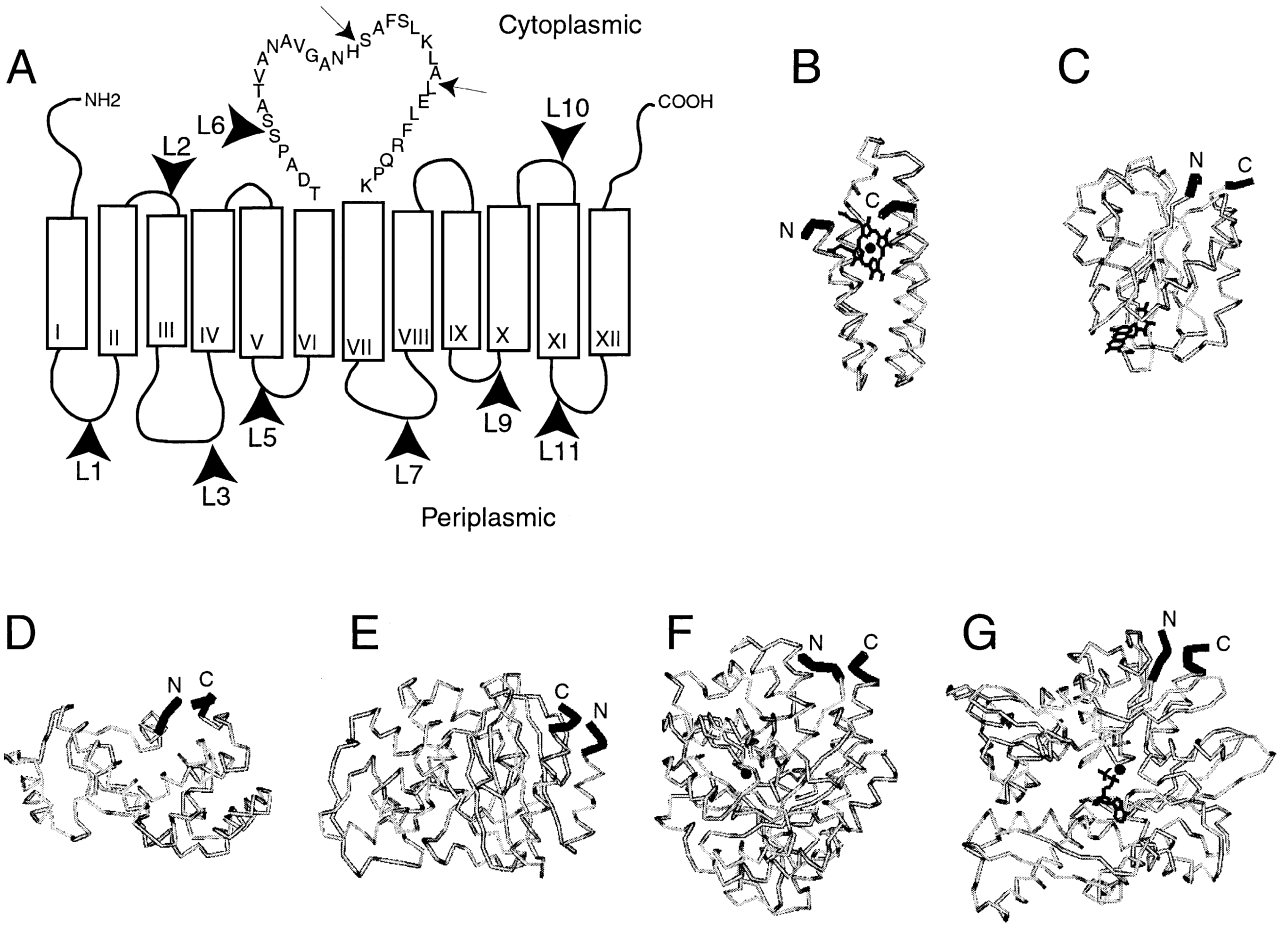
2.7. Western blots

Protein expression levels were measured by Western blot analysis of total membranes. Transformed *E. coli* T184 cells (1.5 ml) at an A_{420} of 10.0 were collected by centrifugation and resuspended in 0.5 ml lactose transport wash buffer. The cells were lysed by sonication and cell debris was removed by centrifugation. Membranes were pelleted by ultracentrifugation and directly taken up in 20 μl SDS-PAGE sample buffer. After SDS gel electrophoresis, the proteins were blotted on a nitrocellulose membrane. The blots were developed using a mouse Anti-Penta-His primary antibody, a peroxidase conjugated antimouse IgG secondary antibody, and visualized by ECL.

3. Results

3.1. Characterization of internal L6 fusion proteins

The internal fusion proteins carrying cytochrome$_{b562}$ (cyt), T4 lysozyme (lyso), flavodoxin (fla), β-lactamase (TEM), adenosine deaminase (ADA) or the 70 kDa heat shock protein ATPase domain (HSC) in the central hydrophilic loop of lactose permease are described in Fig. 1. In the cases of lysozyme, flavodoxin, ADA and β-lactamase, we investigated two constructs for each insertion domain, one with minimal length linkers at the N- and C-termini of the inserted carrier domain, referred to as N1C1, and one with extra amino acids at the two protein junctions (N5C5). These were characterized for expression and lactose transport activity (Fig. 2A,B). Western blot analysis showed strong expression signals for wild-type lac permease, the cytochrome$_{b562}$ fusion and the two lysozyme constructs, while the



H

Construct	Carrier	Permease	N-linker	Insert	C-linker	Permease
wild type	-	...TDAPS				SATVANAVGANHSAFSLKLALELFRQPK...
L6_cyt_N2C6	cytochrome b562	...TDAPS	SA	ADLED-(96aa)-HQKYR	SFLISS	SATVANAVGANHSAFSLKLALELFRQPK...
L6_cyt_N2C1	cytochrome b562	...TDAPS	SA	ADLED-(96aa)-HQKYR	S	SATVANAVGANHSAFSLKLALELFRQPK...
L6_cyt_Δ11	cytochrome b562	...TDAPS	SA	ADLED-(96aa)-HQKYR	S	S SAFSLKLALELFRQPK...
L6_cyt_Δ19	cytochrome b562	...TDAPS	SA	ADLED-(96aa)-HQKYR	S	S LELFRQPK...
L6_lyso_N1C1	T4-lysozyme	...TDAPS	S	MNIFE-(154aa)-AYKNL	S	SATVANAVGANHSAFSLKLALELFRQPK...
L6_lyso_N5C5	T4-lysozyme	...TDAPS	SGGSG	MNIFE-(154aa)-AYKNL	GSGGS	SATVANAVGANHSAFSLKLALELFRQPK...
L6_fla_N1C1	flavodoxin	...TDAPS	S	MNMGL-(161aa)-EMAEH	S	SATVANAVGANHSAFSLKLALELFRQPK...
L6_fla_N5C5	flavodoxin	...TDAPS	SGGSG	MNMGL-(161aa)-EMAEH	GSGGS	SATVANAVGANHSAFSLKLALELFRQPK...
L6_TEM_N1C1	β-lactamase	...TDAPS	S	HPETL-(254aa)-LIKHW	S	SATVANAVGANHSAFSLKLALELFRQPK...
L6_TEM_N5C5	β-lactamase	...TDAPS	SGGSG	HPETL-(254aa)-LIKHW	GSGGS	SATVANAVGANHSAFSLKLALELFRQPK...
L6_ADA_N1C1	adenosine deaminase	...TDAPS	S	AFNKP-(336aa)-YREYQ	S	SATVANAVGANHSAFSLKLALELFRQPK...
L6_ADA_N5C5	adenosine deaminase	...TDAPS	SGGSG	AFNKP-(336aa)-YREYQ	GSGGS	SATVANAVGANHSAFSLKLALELFRQPK...
L6_HSC_N1C1	heat shock ATPase	...TDAPS	S	GPAVG-(368aa)-AILSG	S	SATVANAVGANHSAFSLKLALELFRQPK...

Fig. 1. (A) Schematic representation of the lac permease (adapted from Ref. [33]), with the insertion site of the carriers indicated by the thick arrowhead, and the C-terminal positions of the two truncations in the central loop indicated by the small arrows. (B–G) Cα representations of the soluble proteins used as carrier domains for insertion into the central loop of the lactose permease. The N- and C-termini of the proteins used in the fusion inserts are indicated. (B) Cytochrome$_{b562}$, (PDB ID: 256B; Cα3 to Cα106 distance: 11.1 Å; cofactor: protoporphyrin IX with bound Fe^{2+}), (C) flavodoxin (PDB ID: 1AG9; Cα1 to Cα171 distance: 14.9 Å; cofactor: FMN), (D) T4 lysozyme (PDB ID: 2LZM; Cα1 to Cα164 distance: 11.4 Å), (E) β-lactamase (PDB ID: 1FQG; Cα26 to Cα290 distance: 7.6 Å), (F) adenosine deaminase (PDB ID: 2ADA; Cα6 to Cα352 distance: 17.3 Å; cofactor: Zn^{2+}), (G) heat shock ATP hydrolase (PDB ID: 3HSC; Cα4 to Cα382 distance: 10.4 Å; cofactor: Mg^{2+}ADP). (H) Amino acid sequences in the region of the junctions between the lac permease and the carrier moieties in the L6 series of constructs. The lac permease residues that flank the inserts are serine 193 and serine 194 at the N- and C-terminus, respectively (wild-type lac permease numbering).

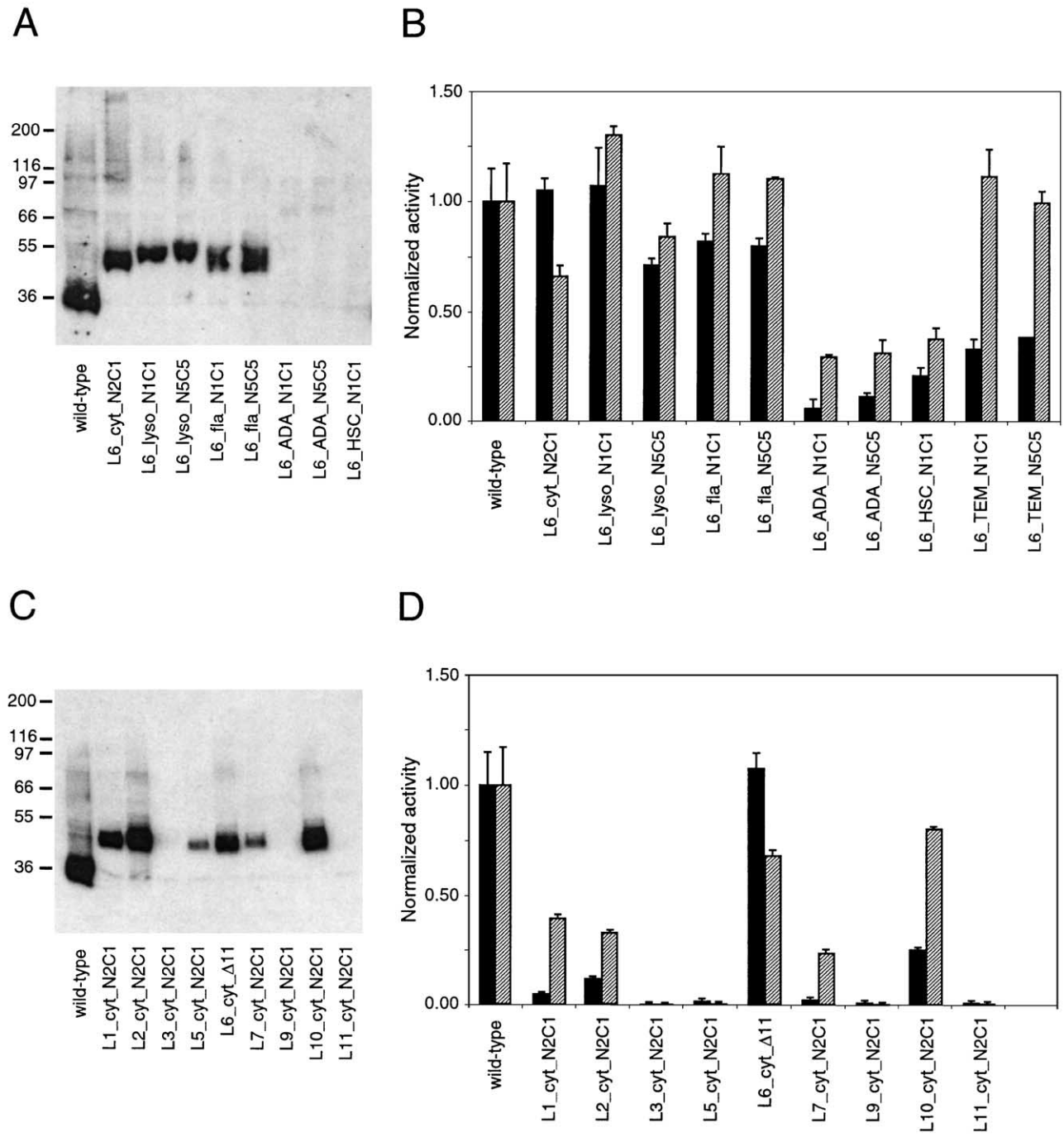


Fig. 2. (A) Expression levels of the L6 fusion proteins as analyzed by Western blotting of total membranes. (B) Lactose transport activity of the L6 fusions. The initial rate and steady state level of accumulation are indicated by black bars and hatched bars, respectively. The rates were normalized to the transport activity of wild-type lac permease. (C) Expression levels of the cytochrome$_{b562}$ insertions into each of nine hydrophilic loops of the lac permease as analyzed by Western blotting of total membranes. (D) Transport activity of the lac permease-cytochrome$_{b562}$ fusion proteins. The initial rate and steady state level of accumulation are indicated by black bars and hatched bars, respectively. The rates were normalized to the transport activity of wild-type lac permease.

flavodoxin fusion proteins have a slightly reduced expression level, and only weak or no expression could be detected for the ADA and HSC fusion proteins. Expression levels were also very low for the β-lactamase fusions (data not shown). As is commonly observed with very hydrophobic membrane proteins, the SDS complexes migrated at anomalously high rates in the gel, giving smaller than expected apparent molecular weights. Surprisingly, the rates of migration of the fusion proteins correlated only roughly with the size of the inserted carrier domains. The initial rate of lactose transport in the L6_cyt_N2C1 and L6_lyso_N1C1 proteins was comparable to that of wild-type permease, while

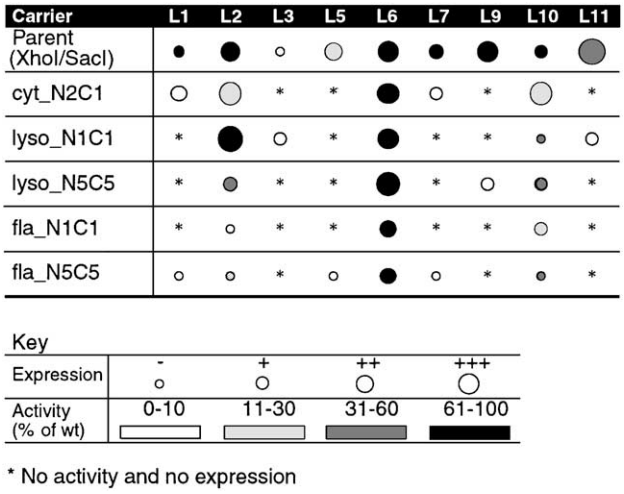


Fig. 3. Summary of the expression and activity of the fusions with five different carriers inserted into each of nine different loops of the permease. Larger circles represent higher levels of expression, and darker shading indicates higher levels of transport activity.

the N5C5 lysozyme and both flavodoxin fusion proteins exhibited slightly reduced initial rates of transport. Most of these constructs reached a steady state level of lactose accumulation that was comparable to wild-type lac permease. The ADA, HSC and β-lactamase fusion proteins had a lower, but significant level of activity, suggesting that functional protein was produced, but at levels below the threshold of detection in the Western blots used here.

3.2. Insertions into other loops

Insertions of carrier proteins into loops other than the central hydrophilic domain were less well tolerated, as shown by the series containing the cyt_N2C1 insertions (Fig. 2C,D). In these cases, the L1, L2 and L10 fusion proteins were expressed at roughly wild-type levels, but these had significantly reduced initial rates of transport. The expression and activity data for the entire series of the nine L*x* parent clones (containing no inserted domains), and 45 permease/carrier fusion proteins are summarized in Fig. 3. The L6 position was clearly the most tolerant of the fusion sites, followed by the L2, L10 and L1 sites in decreasing order of suitability. Insertion into any of the five remaining sites (L3, L5, L7, L9, L11) resulted in the loss of essentially all useable expression and activity. The L3 series was the most poorly behaved of this set, as neither the parent construct, nor any of the insertion fusions, produced viable protein. The length of the linkers was not an important factor, as can be seen by a comparison of the N1C1 and N5C5 series in the lysozyme or flavodoxin fusion proteins. Consistent with the L6 results, the cytochrome$_{b562}$ and lysozyme fusion proteins were better behaved than the flavodoxin fusion proteins across most of the insertion sites. The cell paste from bacteria producing the well-expressed L1, L2, L6 and L10 cytochrome$_{b562}$ fusion proteins was visibly red. In addition, the purified cytochrome$_{b562}$ fusion proteins had similar UV/visible spectra, with an A280/A426 ratio of approximately 1, indicating full occupancy of heme in the cytochrome$_{b562}$ domain of the fusion proteins [9].

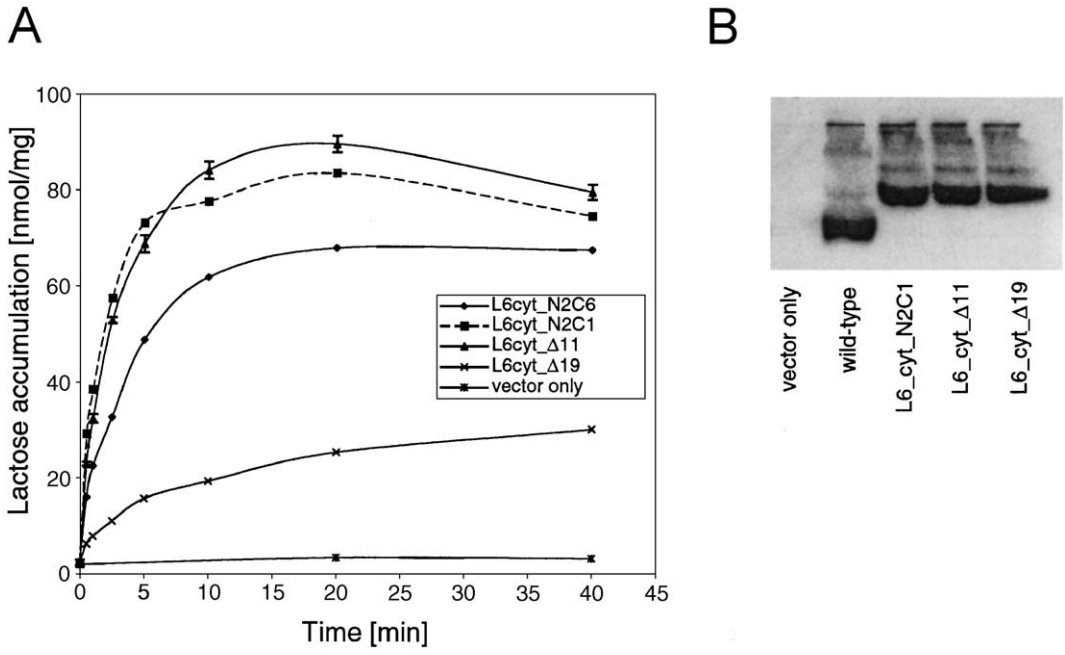


Fig. 4. Expression and transport activity for cytochrome$_{b562}$ fusion proteins with deletions in the central hydrophilic loop. (A) Transport activity as measured by lactose uptake in intact cells. (B) Expression levels of the L6 deletion mutants as analyzed by Western blotting of total membranes.

Thus, the cytochrome$_{b562}$ moiety of these fusion proteins is able to bind heme when it is localized to either the cytoplasmic or periplasmic side of the membrane.

3.3. Shortened central permease loop

For a given carrier, similar levels of expression and activity were observed in the N1C1 and N5C5 series of fusion proteins (Figs. 2 and 3), suggesting that in these cases, the linkers were not interfering with the folding of either of the protein domain. Because our goal was to reduce the flexibility between the protein pairs by minimizing the length of the linker peptides, we investigated the effect of shortening the lac permease central hydrophilic loop in the L6 cytochrome$_{b562}$ fusion proteins (Fig. 4). The initial rate of transport was higher for the N2C1 and Δ11 proteins than for the wild-type and N2C6 proteins. However, the L6_cyt_Δ19 fusion protein had markedly reduced transport activity even though it was well expressed, suggesting that this degree of shortening was not fully compatible with the structural requirements of the permease across the transport cycle, or resulted in a folding defect in the permease.

3.4. Double cytochrome$_{b562}$ fusion proteins

The L1/L6, L2/L6 and L6/L10 double cytochrome$_{b562}$ fusions were well-expressed, and had activities similar to those of the L1, L2 and L10 single cytochrome$_{b562}$ proteins, respectively (data not shown). Thus, the L1, L2 and L10 insertions are compatible with the presence of the well-tolerated L6 insertion, and in these cases, the effects on the fusion protein are essentially additive. In these double fusions, the ratio of the absorbances at 426 and 280 nm in purified protein preparations is roughly twice that of the single cytochrome$_{b562}$ fusion proteins, indicating that both of the cytochrome$_{b562}$ carrier domains contain bound heme.

4. Discussion

It was not expected that all of the designed fusion proteins would produce material suitable for crystallization trials, and the modular system was designed as a system to readily generate a large number of fusion proteins that could be screened to identify a subset of proteins suitable for further study. The majority of the L6 fusion proteins were active in lactose transport, and half of these could be expressed at levels comparable to the wild-type protein. In addition to the increase in polar surface area, features such as purification tags or color (in the case of the cytochrome$_{b562}$ and flavodoxin fusion proteins) can be introduced by this strategy. A 100-residue biotinylation domain from the oxaloacetate decarboxylase of *Klebsiella pneumoniae* has previously been inserted into this position of the permease, and the resulting protein was active and well-expressed [29]. Fusion proteins containing larger inserts (β-lactamase, ADA, HSC) showed significant lactose transport activity, but were expressed only at lower levels and thus were not suitable for purification and further characterization. The products of these larger constructs were roughly over 700 amino acids, and recombinant expression of proteins of this size in *E. coli* is often problematic, even with soluble proteins. For insertions into other loop positions, the rate of success was significantly lower, and only the cytochrome$_{b562}$ and lysozyme inserts into the L1, L2, L6 and L10 positions are being pursued in crystallization trials.

The lac permease and many other proteins in its class can readily tolerate mutations and breaks in the loops structures [7,30]. In a pioneering study, McKenna et al. [31] considered the effect of inserting histidine residues into the hydrophilic domains of the lac permease, and found that either two or six histidine insertions into the L2, L8 and L9 positions had serious negative effects on the protein, while all other sites were well tolerated. A "split lac permease" with a break in the backbone chain at the L2 site was functional [32], and we find here that the fusion inserts at this position also result in functional protein, indicating that the effect on the transport activity depends on the exact nature of the loop modifications. In general, the insertion of the carrier proteins presented here have more deleterious effect on the lac permease than the large number of insertions, deletions or interruptions in the backbone structure that have been described for this protein [7]. The low levels of expression in these fusion proteins could be the result of problems at the level of expression, insertion and/or short lifetime in the membrane.

In the internal fusion design, the optimal linker length that allows for the proper folding of the protein partners depends on the distance between the ordered terminal residues of the carrier. The selected carriers introduce a gap of approximately 10 Å between the Cα positions and the permease loops (Fig. 1), and molecular modeling indicates that these distances can be easily accommodated by N and C linkers of only a few amino acids each. Because of the additional flexibility of interface residues formally designated as belonging to "carrier" or "permease," carrier inserts whose termini are within this range are easily accommodated with very short linkers. The properties of proteins with long (N2C6 or N5C5) or short (N2C1 or N1C1) linkers were similar in all cases tested, suggesting that the poor expression of the L6 ADA, HSC or TEM fusion proteins, or of the L1, L3, L5, L7, L9 or L11 lysozyme or flavodoxin proteins were not due to internal structural strain resulting from the insertion of the carrier. Further support that the linkages between the permease and the carriers are suitably relaxed in the L6 series is provided by the observation that eleven residues from the permease central polar loop could be removed from the cytochrome$_{b562}$ fusion without affecting the transport activity of the protein (Fig. 4). Weinglass and Kaback [33] describe similar dele-

tions made in this loop of the wild-type lac permease. They removed five amino acids without any effect on expression or activity, while a deletion of 12 loop residues showed increased rates of protein turnover and decreased transport activity, and a 20-residue deletion resulted in trace levels of insertion and activity. In our case, larger loop deletions were well tolerated by the protein, presumably because of the additional contribution of the inserted cytochrome$_{b562}$ to the overall distance between the two transmembrane helices.

The optimization of the length of the two linkers that connect the protein domains is clearly an important issue for the strategy of using internal fusion proteins for crystallization. In principle, short, rigid linkers are favored for crystallization, but a limited amount of flexibility may be nonetheless tolerated, since it would allow the two protein domains to sample a wider range of conformations. T4 lysozyme, for example, is well-known as a protein that is readily crystallized, and yet exhibits a large hinge-bending between the amino and carboxy-terminal domains [34] and Ay et al. [17] have demonstrated variability of a hinge angle in the crystal structure of an internal fusion protein between two soluble domains.

There have been several examples of stabilizing ''difficult'' proteins via the formation of complexes with antibody fragments for the purposes of crystallization with both soluble and integral membrane proteins [11–16,35–37]. In the cases of the bacterial cytochrome *c* oxidase [11–13] and the yeast cytochrome bc_1 complex [15], the complex with a cognate Fv fragment had several advantages: it allowed the rapid purification of the PDC–Fv complex, the Fv stabilized the structure by recognizing a conformational epitope, and most significantly, the Fv was an important contributor to the protein–protein contacts in the crystal lattice. Note that antibody fragments that bind selectively to a particular conformation state of the target protein may reduce the aggregation problem and improve the crystallizability of the protein by restricting the internal dynamics of the system [1]. Although the antibody fragment approach is a powerful method for the study of membrane proteins, a major limitation is the availability and production of suitable monoclonal antibodies and their fragments specific for each target protein. The carrier fusion approach circumvents this problem, and can be used in conjunction with antibody fragment methods.

Acknowledgements

Support for this work was provided in part by a grant from The Canadian Institutes for Health Research (MT-13335). C.K.E. was supported by Fellowships from the Deutsche Forschungsgemeinschaft and from the Canadian Institutes of Health Research. We thank Inge Klitsie, Clare McGregor, Jonathan Burns, Nora Eifler, Jimmy Woo and Yan Gao for technical assistance, Ron Kaback for helpful discussions, and Brian Matthews for the T4 lysozyme plasmid.

References

[1] C. Ostermeier, H. Michel, Curr. Opin. Struct. Biol. 7 (1997) 697–701.
[2] H. Michel, *Crystallization of membrane proteins*, in: M.G. Rossmann, E. Arnold (Eds.), *International Tables for Crystallography*, Kluwer, Dordrecht, 2001, pp. 94–99.
[3] S. Penel, E. Pebay-Peyroula, J. Rosenbusch, G. Rummel, T. Schirmer, P.A. Timmins, Biochimie 80 (1998) 543–551.
[4] T.M. Iverson, C. Luna-Chavez, G. Cecchini, D.C. Rees, Science 284 (1999) 1961–1966.
[5] E. Pebay-Peyroula, R.M. Garavito, J.P. Rosenbusch, M. Zulauf, P.A. Timmins, Structure 3 (1995) 1051–1059.
[6] I.T. Paulsen, L. Nguyen, M.K. Sliwinski, R. Rabus , M.H. Saier Jr., J. Mol. Biol. 301 (2000) 75–100.
[7] H.R. Kaback, M. Sahin-Toth, A.B. Weinglass, Nat. Rev. Mol. Cell. Biol. 2 (2001) 610–620.
[8] M. Sahin-Toth, A. Karlin, H.R. Kaback, Proc. Natl. Acad. Sci. U. S. A. 97 (2002) 10729–10732.
[9] G.G. Privé, G.V. Verner, C. Weitzman, K. Zen, D. Eisenberg, H.R. Kaback, Acta Crystallogr. Sect. D 50 (1994) 375–379.
[10] G.G. Privé, H.R. Kaback, J. Bioenerg. Biomembr. 28 (1996) 29–34.
[11] C. Ostermeier, S. Iwata, B. Ludwig, H. Michel, Nat. Struct. Biol. 2 (1995) 842–846.
[12] G. Kleymann, C. Ostermeier, B. Ludwig, A. Skerra, H. Michel, Biotechnology 13 (1995) 155–160.
[13] C. Ostermeier, A. Harrenga, U. Ermler, H. Michel, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 10547–10553.
[14] E. Padan, M. Venturi, H. Michel, C. Hunte, FEBS Lett. 441 (1998) 53–58.
[15] C. Hunte, J. Koepke, C. Lange, T. Rossmanith, H. Michel, Struct. Folding Des. 8 (2000) 669–684.
[16] Y. Zhou, J.H. Morais-Cabral, A. Kaufman, R. MacKinnon, Nature 414 (2001) 43–48.
[17] J. Ay, F. Gotz, R. Borriss, U. Heinemann, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 6613–6618.
[18] G.J. Turner, L.J. Miercke, A.K. Mitra, R.M. Stroud, M.C. Betlach, A. Winter-Vann, Protein Expression Purif. 17 (1999) 324–338.
[19] B. Byrne, J. Abramson, M. Jansson, E. Holmgren, S. Iwata, Biochim. Biophys. Acta 1459 (2000) 449–455.
[20] J. Zhuang, G.G. Privé, G.E. Werner, P. Ringler, H.R. Kaback, A. Engel, J. Struct. Biol. 125 (1999) 63–75.
[21] C.K. Engel, L. Chen, G.G. Privé, Stability of the lactose permease in detergent solutions, Biochim. Biophys. Acta 1564 (2002) 47–56 (this issue).
[22] D.M. Hoover, M.L. Ludwig, Protein Sci. 6 (1997) 2525–2537.
[23] T. Alber, D.P. Sun, J.A. Nye, D.C. Muchmore, B.W. Matthews, Biochemistry 26 (1987) 3754–3758.
[24] D.K. Wilson, F.B. Rudolph, F.A. Quiocho, Science 252 (1991) 1278–1284.
[25] M.C. O'Brien, D.B. McKay, J. Biol. Chem. 268 (1993) 24323–24329.
[26] N.C. Strynadka, H. Adachi, S.E. Jensen, K. Johns, A. Sielecki, C. Betzel, K. Sutoh, M.N. James, Nature 359 (1992) 700–705.
[27] R.M. Teather, J. Bramhall, I. Riede, J.K. Wright, M. Furst, G. Aichele, U. Wilhelm, P. Overath, Eur. J. Biochem. 108 (1980) 223–231.
[28] P. Viitanen, M.J. Newman, D.L. Foster, T.H. Wilson, H.R. Kaback, Methods Enzymol. 125 (1986) 429–452.
[29] T.G. Consler, B.L. Persson, H. Jung, K.H. Zen, K. Jung, G.G. Privé, G.E. Verner, H.R. Kaback, Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 6934–6938.
[30] J.L. Popot, D.M. Engelman, Annu. Rev. Biochem. 69 (2000) 881–922.

[31] E. McKenna, D. Hardy, H.R. Kaback, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 11954–11958.
[32] W. Wrubel, U. Stochaj, U. Sonnewald, C. Theres, R. Ehring, J. Bacteriol. 172 (1990) 5374–5381.
[33] A.B. Weinglass, H.R. Kaback, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 8938–8943.
[34] X.J. Zhang, J.A. Wozniak, B.W. Matthews, J. Mol. Biol. 250 (1995) 527–552.
[35] A. Jacobo-Molina, A.D. Clark Jr., R.L. Williams, R.G. Nanni, P. Clark, A.L. Ferris, S.H. Hughes, E. Arnold, Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 10895–10899.
[36] P.D. Kwong, R. Wyatt, J. Robinson, R.W. Sweet, J. Sodroski, W.A. Hendrickson, Nature 393 (1998) 648–659.
[37] L.C. Kovari, C. Momany, M.G. Rossmann, Structure 3 (1995) 1291–1293.