doi:10.1006/jmbi.2000.4202 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2001) **305**, 891–904

JMB

Structural Basis for Oligosaccharide Recognition by *Pyrococcus furiosus* Maltodextrin-binding Protein

Artem G. Evdokimov[1]*, D. Eric Anderson[2], Karen M. Routzahn[2] and David S. Waugh[1,2]*

[1]*Protein Engineering Section Macromolecular Crystallography Laboratory*

[2]*Structural Biology Core Facility, Program in Structural Biology, National Cancer Institute - Frederick Cancer Research and Development Center, P.O. Box B, Frederick MD 21702-1201, USA*

**Corresponding authors*

A maltodextrin-binding protein from *Pyrococcus furiosus* (*Pfu*MBP) has been overproduced in *Escherichia coli*, purified, and crystallized. The crystal structure of the protein bound to an oligosaccharide ligand was determined to 1.85 Å resolution. The fold of *Pfu*MBP is very similar to that of the orthologous MBP from *E. coli* (*Eco*MBP), despite the moderate level of sequence identity between the two proteins (27 % identity, 46 % similarity). *Pfu*MBP is extremely resistant to heat and chemical denaturation, which may be attributed to a number of factors, such as a tightly packed hydrophobic core, clusters of isoleucine residues, salt-bridges, and the presence of proline residues in key positions. Surprisingly, an attempt to crystallize the complex of *Pfu*MBP with maltose resulted in a structure that contained maltotriose in the ligand-binding site. The structure of the complex suggests that there is a considerable energy gain upon binding of maltotriose in comparison to maltose. Moreover, isothermal titration calorimetry experiments demonstrated that the binding of maltotriose to the protein is exothermic and tight, whereas no thermal effect was observed upon addition of maltose at three temperatures. Therefore, *Pfu*MBP evidently is designed to bind oligosaccharides composed of three or more glucopyranose units.

Keywords: thermostable protein; carbohydrate-binding; maltose-binding protein; maltodextrin-binding protein; *Pyrococcus furiosus*

Introduction

ATP-binding cassette (ABC) transporters comprise one of the largest superfamilies of proteins (Tatusov *et al*., 1997). They orchestrate the ATP-dependent transport of a wide variety of solutes across biological membranes in both prokaryotes and eukaryotes (Higgins, 1992; Croop, 1998). In bacteria and archae, the majority of ABC transporters consist of a high-affinity substrate-binding protein, two integral membrane proteins, and a pair of membrane-associated ATPase subunits (Ehrmann *et al*., 1998; Horlacher *et al*., 1998). In Gram-negative bacteria (e.g. *Escherichia coli*), the binding protein is a soluble constituent of the periplasm, whereas in Gram-positive bacteria and archae it is a lipoprotein that is anchored in the cytoplasmic membrane and exposed like a receptor on the cell surface (Wu, 1996).

Much of the existing knowledge about microbial ABC transporters has been gleaned from studies of the high-affinity maltose transport system in *E. coli*. The *E. coli* maltodextrin-binding protein (*Eco*MBP), in particular, has been the subject of intensive study. A battery of genetic, biochemical, and biophysical methods have produced a wealth of information about the structure and ligand-binding properties of *Eco*MBP (e.g. see Martineau *et al*., 1990; Spurlino *et al*., 1991; Sharff *et al*., 1992,

Present address: D. E. Anderson, Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0580, USA.

Abbreviations used: ABC, ATP-binding cassette; *Eco*MBP, *Escherichia coli* maltodextrin-binding protein; HIC, hydrophobic interaction chromatography; ITC, isothermal titration calorimetry; MS, mass spectrometry; ORF, open reading frame; *Pfu*MBP, *Pyrococcus furiosus* maltodextrin-binding protein; PMSF, phenylmethylsulphonyl chloride; PCR, polymerase chain reaction; rmsd, root mean square deviation; TCEP, tris-carboxyethyl phosphine hydrochloride; *Tli*MBP, *Thermococcus litoralis* maltose-binding protein; *Tma*MBP, *Thermotoga maritima* maltodextrin-binding protein.

E-mail addresses of the corresponding authors: eudokima@mail.ncifcrf.gov; waughd@ncifcrf.gov

0022-2836/01/040891–14 $35.00/0

1993; Hall *et al*., 1997a,b,c; Novokhatny & Ingham, 1997; Gardner *et al*., 1998; Thomson *et al*., 1998). It is unclear, however, to what degree these observations can be generalized. That is, what are the similarities and differences between *Eco*MBP and its distant relatives? To address this question, we have undertaken a study of an MBP from the hyperthermophilic archeon *Pyrococcus furiosus* (*Pfu*MBP).

Here, we describe the crystal structure of *Pfu*MBP in complex with maltotriose at a resolution of 1.85 Å. Despite the moderate level of sequence identity between *Pfu*MBP and *Eco*MBP (27 % identity, 46 % similarity), the two structures are similar in many respects. The bound oligosaccharide is located in a deep cleft that is lined with hydrophobic side-chains on the sides and charged side-chains on the bottom. *Eco*MBP binds both maltose and maltotriose with similar affinities (Thomson *et al*., 1998). In contrast, *Pfu*MBP appears to bind maltotriose much more tightly than maltose.

Results

Expression, purification, and crystallization of *Pfu*MBP

A BLAST search of the *P. furiosus* genome sequence identified two open reading frames (ORFs) with recognizable similarity to *Eco*MBP. One of these (Utah genome center ID Pf_1616024) encodes a protein that is identical with a previously characterized *Thermococcus litoralis* maltose-binding protein (*Tli*MBP) (Horlacher *et al*., 1998). The amino acid sequence encoded by the other ORF (Utah genome center ID Pf_1788981) is very different from that of *Tli*MBP. However, both ORFs in *P. furiosus* exhibit about the same degree of sequence similarity with *Eco*MBP (27 % identity). Pf_1788981 appears to be part of an operon that includes components of an ABC transporter and an amylopullulanase gene.

To determine whether this uncharacterized ORF in the *P. furiosus* genome (Pf_1788981) encodes a functional MBP, we amplified the coding sequence from genomic DNA by polymerase chain reaction (PCR) and cloned it into the pMal-C2x expression vector (New England Biolabs), effectively replacing the coding sequence of *Eco*MBP with that of the *P. furiosus* ORF. Like *Tli*MBP, the mature domain of this putative *Pfu*MBP is augmented by an N-terminal peptide that presumably serves to anchor the protein to the membrane. Because this peptide bears no similarity to the *Eco*MBP signal peptide, which is removed during translocation of the protein across the inner membrane to generate the mature MBP, it was not clear exactly where the N terminus of the ''mature'' domain of *Pfu*MBP should begin. However, the two sequences could be aligned unambiguously beginning with the fifth residue from the N terminus of mature *Eco*MBP. Therefore, we designed our vector to produce a chimeric polypeptide consisting of the first four residues of mature *Eco*MBP (plus an initiator methionine residue) followed by 376 residues encoded by *P. furiosis* DNA.

Upon induction with isopropyl β-D-thiogalactopyranoside (IPTG) in *E. coli* BL21/DE3 cells, the *Pfu*MBP expression vector produced an abundant protein with the expected molecular mass (ca 43 kDa), as judged by sodium dodecyl sulfate/polyacrylamide gel electrophoresis. This protein accumulated in a soluble form in the crude cell extract, and remained soluble even after endogenous proteins were precipitated by heating the extract at 90 °C for 15 minutes (data not shown). Unlike *Tli*MBP (Horlacher *et al*., 1998), this *Pfu*MBP bound tightly to amylose resin, from which it could be displaced by malto-oligosaccharides, confirming that it is indeed a functional maltodextrin-binding protein. After sequential amylose affinity and hydrophobic interaction chromatography steps, the protein was completely free of detectable contaminants. Ligand-free *Pfu*MBP was prepared by denaturation and refolding after amylose affinity chromatography.

Liganded *Pfu*MBP was crystallized from ammonium sulfate (pH 9) by the hanging drop technique. Cryogenic X-ray diffraction data were collected for the native crystal and data for two heavy-atom derivatives at 1.85 Å and 2.2 Å resolution, respectively. The structure was solved using multiple isomorphous replacement and refined as described in Materials and Methods.

Refolded, ligand-free protein crystallized from polyethylene glycol (4000 or 6000; pH 4-5) in the form of very thin needles that did not lend themselves to crystallographic analysis. When maltotriose was added to the refolded protein, the complex crystallized from ammonium sulfate at high pH. These crystals were identical with those obtained when the liganded protein was purified under native conditions.

Overall comparison between the structures of *Pfu*MBP and *Eco*MBP

Like *Eco*MBP, *Pfu*MBP is composed of 19 α-helices and 17 β-strands that are distributed between two distinct domains (lobes), both of which are composed of a central pleated β-sheet surrounded by α-helices. (Figure 1(a)). The carbohydrate-binding site is located in the groove between the lobes, which are connected by a hinge region made of two antiparallel β-strands and an α-helix.

Here, we report the structure of the carbohydrate-bound form of *Pfu*MBP. Hence, it is logical to use the structure of a liganded form of *Eco*MBP for comparison. Several structures of oligoscaccharide-*Eco*MBP complexes have been reported in the literature, including the maltotriose-*Eco*MBP complex used as a reference throughout this work (PDB accession code 3MBP, Quiocho *et al*., 1997).

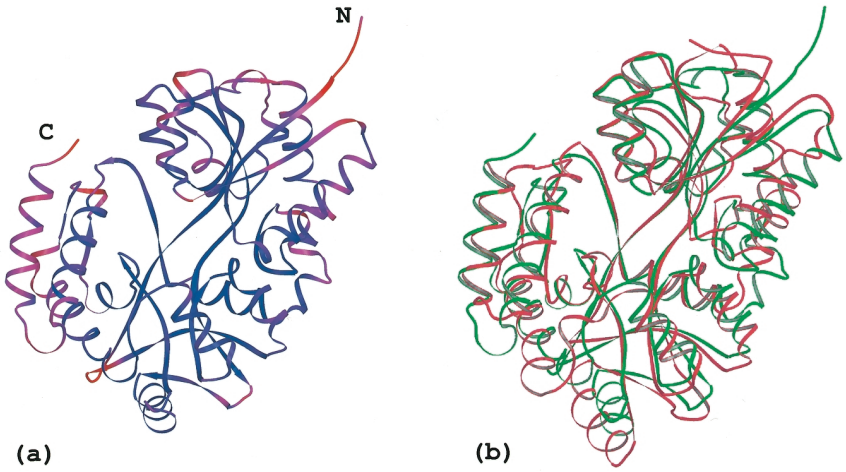

Figure 1. (a) Secondary structure of *Pfu*MBP, colored by atomic displacement factor. Blue corresponds to low values, red to high values. Graphics created with MOLSCRIPT (Kraulis, 1991), BOBSCRIPT (Esnouf, 1997), and Raster3D (Bacon & Anderson, 1988; Merrit & Murphy, 1994). (b) Superposition of the *Pfu*MBP (green) and *Eco*MBP (red) backbone structures.

Alignment of the two structures by LSQMAN (Kleywegt, 1996) resulted in a match of 316 C^{α} atoms with a root-mean-square deviation (rmsd) of 1.8 Å (the overall C^{α} position rmsd is 9.2 Å). Of the aligned amino acid residues, 29% are identical (27% overall sequence identity). In light of the above, it is unclear to us why we did not succeed in our attempts to solve this structure by molecular replacement. A structure-validated sequence alignment of the two proteins is presented in Figure 2. The conserved residues are distributed fairly evenly throughout the two domains with no apparent pattern, except for the residues in the carbohydrate-binding site.

Although *Pfu*MBP and *Eco*MBP have the same fold, the position and orientation of many of the secondary structure elements, especially those on the side of the protein opposing the carbohydrate-binding site, are quite different (Figure 1(b)). The surface of *Pfu*MBP bears 15 more negative charges than that of *Eco*MBP, as evidenced by the electrostatic plots in Figure 3(a). Both proteins have a prominent patch of exposed hydrophobic surface on the side opposite to the carbohydrate-binding site (Figure 3(b)) as well as several smaller hydrophobic patches distributed over their surface.

Bound oligosaccharide and the carbohydrate-binding site

*Pfu*MBP ostensibly was eluted from the amylose column with maltose, which is a disaccharide. To our surprise, however, the shape of the electron density in the carbohydrate-binding site of *Pfu*MBP clearly indicated that three glucopyranoside rings were present. When the three sugar rings are refined independently, the distances between the anomeric oxygen atoms of rings 2 and 3 and the C4 atoms of the neighboring glucopyranose rings 1 and 2 correspond to C-O covalent bond lengths (Figure 4). To further investigate whether the electron density corresponding to the three sugar rings describes a trisaccharide, rather than the superimposition of two disaccharides occupying adjacent

```
          **   *
MKIEEGKVVIWHAMQPNELEVFQSLAEEYMALCPEVEIVFEQKPNLEDALKAAIPTGQG PfuMBP
MKIEEGKLVIWINGDKG-YNGLAEVGKKFEKDT-GIKVTVEHPDKLEEKFPQVAATGDG EcoMBP

    ** *                                              *
PDLFIWAHDWIGKFAEAGLLEPIDEYVTEDLLNEFAPMAQDAMQYKGHYYALPFAAETVA PfuMBP
PDIIFWAHDRFGGYAQSGLLAEIT--PDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALS EcoMBP

                                    ****
IIYNKEMVSEPPKTFDEMKAIMEKYYDPANEKYGIAWPI-NAYFISAIAQAFGGYYFDDK PfuMBP
LIYNKDLLPNPPKTWEEIPALDKE--LKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYE EcoMBP

                                  *                 *
T-----EQPGLDKPETIEGFKFFFTEIW-PYMAPTGDYNTQQSIFLEGRAPMMVNGPWSI PfuMBP
NGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAW EcoMBP

                                  *
NDVKKAGINFGVVPLPPIIKDGKEYWPRPYGGVKLIYFAAGIKNKDAAWKFAKWLTTSEE PfuMBP
SNIDTSKVNYGVTVLPTF----KGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDE EcoMBP

                                          *         *
SIKTLALELG-YIPVLTKVLDDPEIKNDPVIYGFGQAVQHAYLMPKSPKMSAVWGGVDGA PfuMBP
GLEAVNKDKPLGAVALKSYEEE--LAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTA EcoMBP

INEILQDPQNADIEGILKKYQQEILNNMQG                                 PfuMBP
VINAASGRQ--TVDEALKDAQTRITK----                                 EcoMBP
```

Figure 2. Structure-based sequence alignment of *Pfu*MBP and *Eco*MBP. Identical residues are underlined, chimeric residues are italicized, and residues in the sugar-binding site are marked with an asterisk (*).

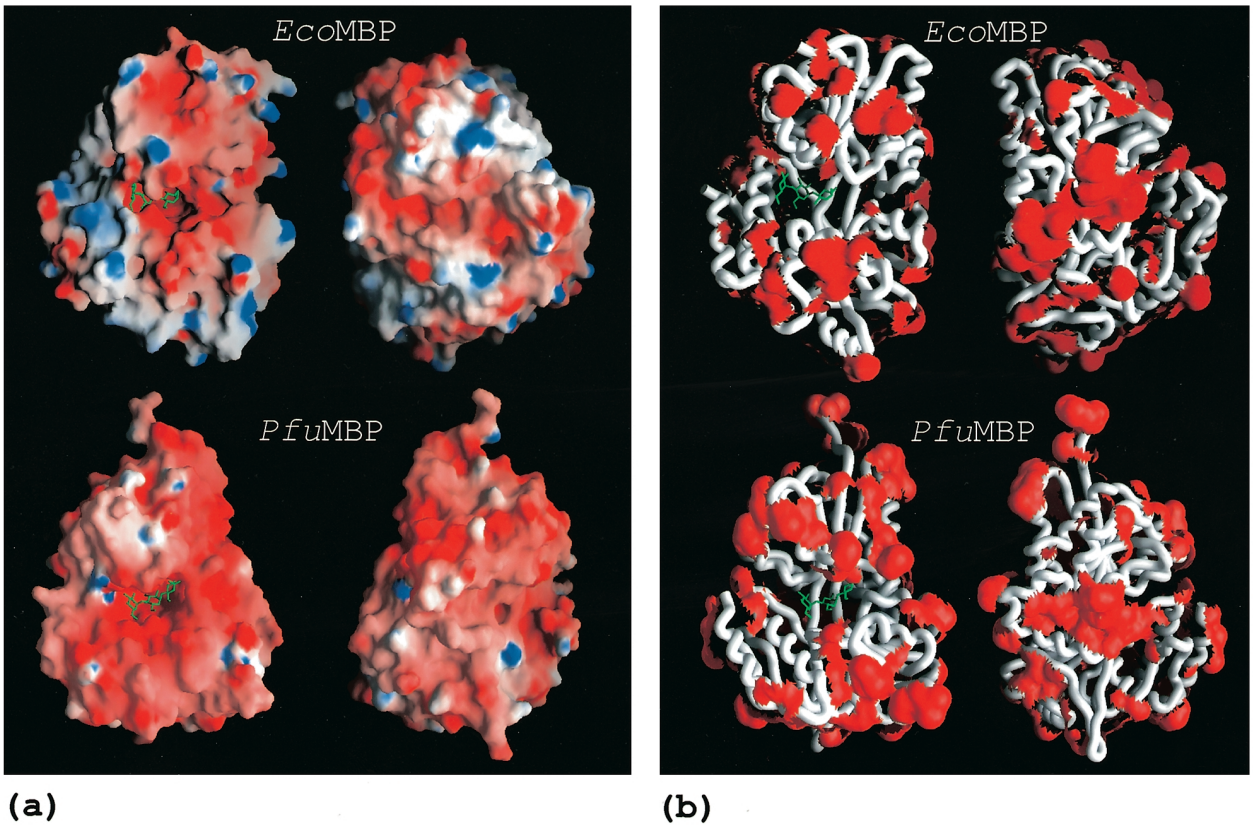

Figure 3. Surface properties of *Pfu*MBP. Graphics generated with GRASP (Nicholls *et al.*, 1991). (a) Electrostatic plots of *Eco*MBP (top) and *Pfu*MBP (bottom) surfaces. Blue areas correspond to positive potential, red to negative potential. Bound carbohydrate is shown in green. (b) Hydrophobic surface plots of *Eco*MBP and *Pfu*MBP. The hydrophobic surface is depicted in red, protein backbone in white, and bound carbohydrate in green.

sites, we performed refinement using variable occupancy parameters (one for each pyranoside) by assigning the same occupancy to every atom of each carbohydrate residue. The resulting occupancy distribution is 1:1:0.9 for rings 1, 2, and 3, respectively. Although this ratio is also consistent with an occupancy of 90 % maltotriose and 10 % maltose, we consider this to be very unlikely. Finally, the electron density near the anomeric carbon atom of ring 1 clearly shows an α/β (33 % α, 66 % β) disordered oxygen atom, which is typical of reducing carbohydrates. No such feature is observed in the vicinity of the anomeric carbon atoms of rings 2 or 3. Taken together, these results can mean only that the three rings comprise a single molecule of maltotriose rather than two disorder-averaged, overlapping maltose molecules. The results of further experiments supported this view (see below).

In *Pfu*MBP, the carbohydrate is sandwiched between two layers of aromatic rings, formed by the side-chains of Trp345, Tyr161, Trp230, and Trp65 (Figure 5(a), Table 1), which is similar to the arrangement in *Eco*MBP. Ala13, Met14, Ala160, Tyr210, Met335, Ala66, and Phe162 define the boundaries of the binding site and engage in hydrophobic contacts with the ligand. A network of hydrogen bonds involving the carbohydrate moiety, the side-chains of Trp65, Trp69, Asp68, Asn159, Glu116, Lys266, and Glu18, the main chain of residues Ala13(O) and Tyr161(N), and

Table 1. Surface-to-surface contacts between maltotriose and *Pfu*MBP

Carbohydrate residue	Surface contact	
	Residue	Distance (Å)[a]
Glc(1)	Met14	3.4
	Ala13	3.9
	Trp65	3.6-4.3
	Trp230	3.7-4.3
	Tyr210	4.0-4.4
	Phe162	4.5
Glc(2)	Trp65	3.8-4.5
	Ala66	3.3
	Met335	3.8
	Tyr161	3.6-4.4
	Trp230	3.6-3.9
	Phe162	3.6-3.8
Glc(3)	Tyr161	3.1-3.5
	Met335	3.7-3.9
	Trp345	3.8-4.5

[a] Range of interatomic distances between pairs of atoms. Hydrogen atoms are not included.

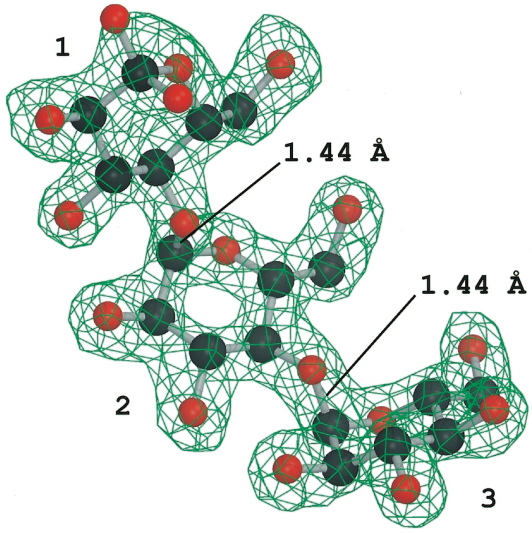

Figure 4. $|F_o - F_c|$ electron density difference map contoured at 3.0σ in the area of the bound maltotriose molecule. This map was calculated before the ligand was added to the model.

nine water molecules is shown in Figure 5(b) and Table 2.

Superimposition of the binding sites of the *Eco*MBP and *Pfu*MBP complexes, using only the carbohydrate atoms of the matching glucopyranoside rings revealed that maltotriose binds to *Pfu*MBP in a radically different manner than the way in which it binds to *Eco*MBP (Figure 5(c)). In the *Pfu*MBP-maltotriose complex, all three glucopyranosides engage in extensive interactions with the protein, whereas the last non-reducing glucose residue of maltotriose makes relatively few contacts with the protein in the *Eco*MBP complex. Compared with *Eco*MBP, the sugar-binding cleft of *Pfu*MBP extends farther toward the core of the protein. The lengthy side-chains of Asn12 and Lys15, which in *Eco*MBP occupy the space necessary to accommodate the reducing pyranoside ring, are transformed into shorter side-chains of Ala13 and Pro16, respectively, in *Pfu*MBP. Additional space is provided by displacement of the (13-19) loop, caused by several residue changes, including Gln15 (from Asp14 in *Eco*MBP), which forms a double hydrogen bond with Asn211 (changed from Ser211 in *Eco*MBP).

Energetics of maltose and maltotriose binding to *Pfu*MBP

No significant heat release or absorption could be detected upon titration of *Pfu*MBP with maltose at 15 °C. As a control, *Eco*MBP was titrated with the same batch of maltose solution at 15 °C. As expected, this reaction was endothermic, with thermodynamic parameters similar to those described earlier (Thomson *et al.*, 1998). Titration of *Pfu*MBP

Table 2. Hydrogen bonds between maltotriose, *Pfu*MBP, and solvent

	Hydrogen bond		
Carbohydrate residue	Atom	To	Distance (Å)
Glc(1)	O1β	Ala13_O	2.96
		W	2.76
		W	2.98
	O1α	Ala13_O	2.74
	O2	Glu18_OE1	2.64
		W	2.67
	O3	Lys266_NZ	2.63
		Glc(2) _O2[a]	2.68
	O5	W	2.90
	O6	W	2.75
		W	2.73
Glc(2)	O2	Glc(1) _O3[a]	2.68
		Glu116_OE2	2.61
	O3	Glc(3) _O2[a]	2.74
		W	2.87
	O6	Asn159_ND2	2.97
		W	3.1
Glc(3)	O2	Trp65_NE1	3.04
		Asp68_OD1	2.64
		Glc(2) _O3[a]	2.74
	O3	Trp69_NE1	2.96
		Asp68_OD2	2.63
		W	2.91
	O4	W	2.78
		W	2.83
	O6	Tyr161_N	3.06
		W	2.74

[a] Intramolecular hydrogen bond between O3 of one residue of maltotriose and O2 of the adjacent residue.

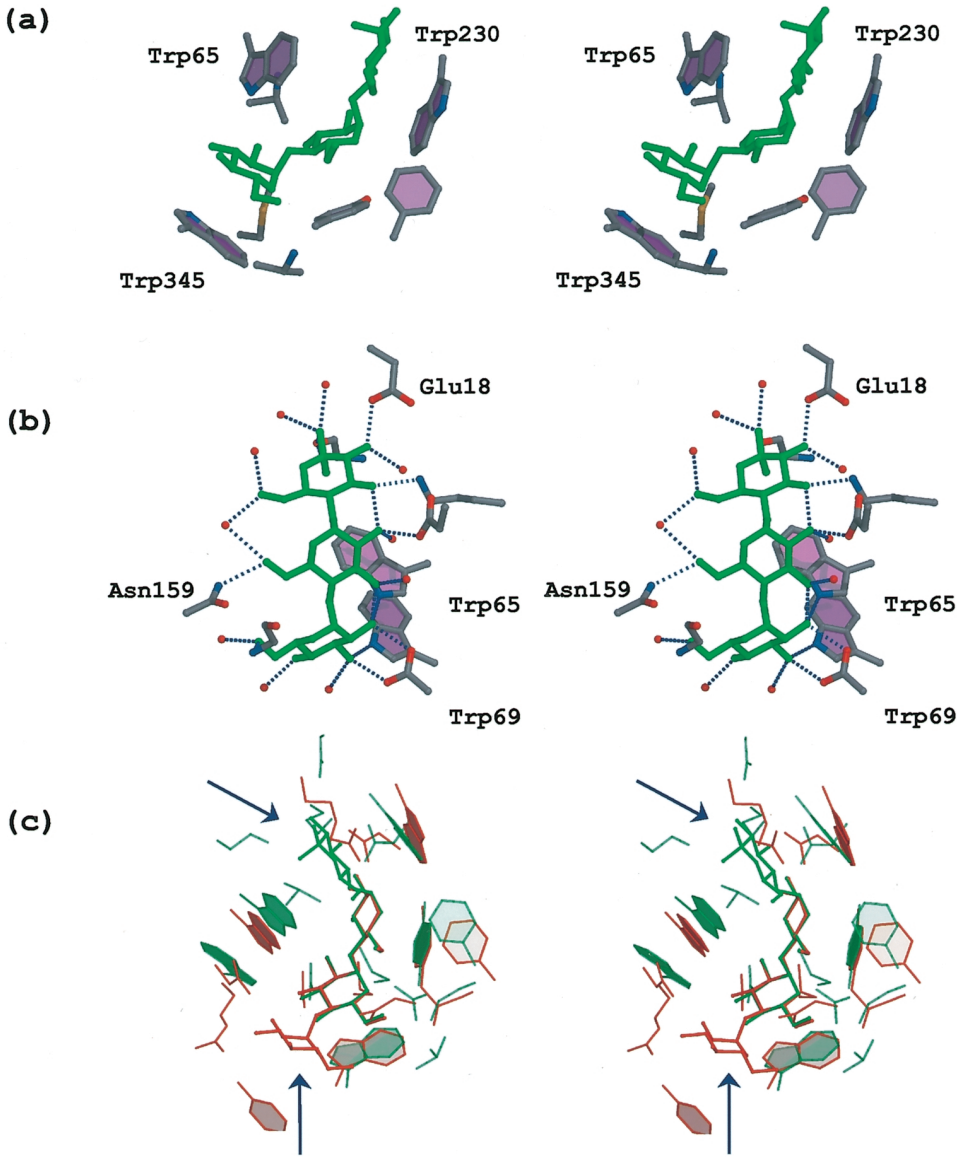

Figure 5. *Pfu*MBP carbohydrate-binding site. The maltotriose ligand is always depicted with its reducing ring pointing upwards. (a) Stereo representation of the *Pfu*MBP residues that participate in hydrophobic interactions with the maltotriose (green) in the binding site. (b) Stereo representation of the *Pfu*MBP residues and water molecules (red spheres) that form the network of hydrogen bonds (blue dotted lines) with maltose (green) in the binding site. (c) Stereo plot of the superposition of *Pfu*MBP (green) and *Eco*MBP (red) carbohydrate-binding sites. Arrows point to the most prominent differences in the arrangement of the bound ligands.

with maltotriose at 15 °C was strongly exothermic (Figure 6), which allowed us to estimate the constant, enthalpy, and entropy of binding to be 2.7×10^7 M^{-1}, −4.98 kcal/mol, and 13.5 cal/mol K, respectively.

Potential determinants of thermostability

*Pfu*MBP is an exceptionally stable protein. No significant change in the circular dichroism spectrum is observed at temperatures up to 85 °C or in 6 M guanidine hydrochloride (J. Fox and D.E.A., unpublished results). In contrast, *Eco*MBP unfolds at approximately 65 °C or in 1 M guanidine hydrochloride (Ganesh *et al.*, 1997). In our opinion, the greater stability of *Pfu*MBP is likely due to a combination of factors, including an abundance of salt-bridges, the presence of proline residues in key positions, and the presence of tightly packed networks of isoleucine residues and other hydrophobic residues in the core of the protein.

Salt-bridges have been proposed to play a crucial role in promoting hyperthermostability in proteins (Perutz & Raidt, 1975; Elcock, 1998). Using a cutoff distance of 3.9 Å, we have manually identified 23 potential ion pairs in *Pfu*MBP and ten potential ion pairs in *Eco*MBP. Using HBPLUS (McDonald & Thornton, 1994) with the same distance cutoff, we

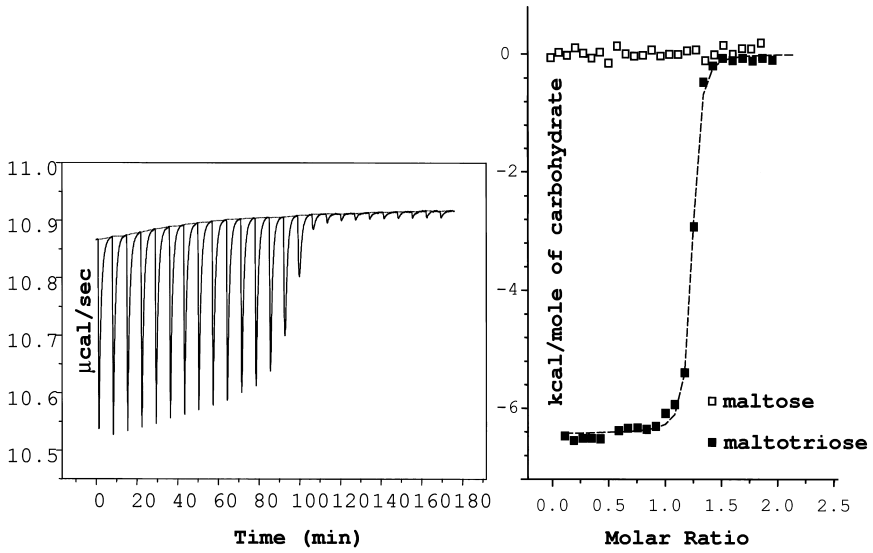

Figure 6. Titration of *Pfu*MBP with maltotriose at 15 °C. On the right side are shown integrated ΔH values (filled squares) plotted against the molar ratio of reagents, along with the fitted curve representing a single-binding-site model (broken line). Integrated ΔH values for titration of *Pfu*MBP with maltose are shown as open squares.

have identified 15 hydrogen bonds between oppositely charged residues of *Pfu*MBP compared to only five such hydrogen bonds in *Eco*MBP. Thus, *Pfu*MBP contains many more salt-bridges than *Eco*MBP. An illustration of the salt-bridge distribution in the *Pfu*MBP structure is presented in Figure 7.

It has been proposed that the presence of proline residues in key positions, such as position 2 of a β-turn, and at the beginning and the end of structural elements, can greatly increase the thermostability of proteins (Matthews *et al*., 1987; Korkhin *et al*., 1998). Out of 21 proline residues (5.7 % of overall composition) in *Eco*MBP, only one is positioned in a β-turn and six are found at the N termini of α-helices. In addition, a short loop (248-254) is stabilized by proline residues at the beginning and the end. In contrast, six out of 29 proline residues in *Pfu*MBP (7.6 % of overall composition) occur in β-turns, five α-helices are capped with proline at both ends, and two more are capped on their N termini. The (248-259) loop in *Pfu*MBP is flanked by proline residues at positions 246 and 261, and contains proline at positions 248, 249, and 259. These proline residues probably help compensate for the destabilizing effect of the five extra residues in the loop.

Isoleucine clusters and tight packing of hydrophobic residues in the core of the protein are often regarded as major determinants of thermal stability (Danson & Hough, 1998). *Pfu*MBP and *Eco*MBP contain 32 and 23 Ile residues, respectively. In *Eco*MBP, the 79-266-59-9 cluster forms a stack across four β-strands. In *Pfu*MBP, the same role can be assigned to the analogous 38-10-64-268 cluster. In addition, the *Pfu*MBP structure contains two smaller Ile clusters, 304-322-316 and 121-232, as well as an extensive network of isoleucine residues, 158, 166, 200, 250, 367, 355, 364, 190, 352 and 163 (Figure 7). A remarkable hydrophobic stack with a high degree of σ-π and π-π stabilization is formed by Val78, Phe73, Trp69, Trp65, the central ring of the ligand, Trp230, Pro 229, Phe325, and Val244 (Figure 8).

A number of other factors, such as surface-to-volume ratios, overall hydrogen bonding, helix capping, and the proportion of thermolabile residues, probably do not contribute signifiantly to the stabilization of *Pfu*MBP. The cavity volumes and surface-to-volume ratios of the two proteins are similar. The largest cavity found for *Pfu*MBP has a volume of 28 $Å^3$, with the total volume of cavities being only 33 $Å^3$. The two largest cavities in *Eco*MBP have volumes of 25 and 16 $Å^3$, the total volume of the cavities being 54 $Å^3$. *Pfu*MBP has a volume of 51,553 $Å^3$ and a surface area of 13,044 $Å^2$. The corresponding values for *Eco*MBP are 49,328 $Å^3$ and 13,352 $Å^2$, respectively. The statistics of hydrogen bonding are also similar for the two proteins: *Pfu*MBP and *Eco*MBP have an average of 0.88 and 0.90 hydrogen bond per residue, respectively. No significant difference is observed in the distribution of main-chain *versus* side-chain hydrogen bonds. Negatively charged amino acid residues are found at the N termini of α-helices in only two cases for *Pfu*MBP and in one case for *Eco*MBP. Consequently, we do not believe that helix capping contributes significantly to the stability of either protein. Asn, Gln, Met, and Cys can be classified as thermolabile because of their tendency to undergo chemical modification at high temperatures (Russell *et al*., 1997). Surprisingly, *Pfu*MBP contains almost twice as many of these residues as does *Eco*MBP (65 and 36, respectively). Moreover, *Pfu*MBP has four Asp-Pro bonds, which are potentially susceptible to hydrolysis at high

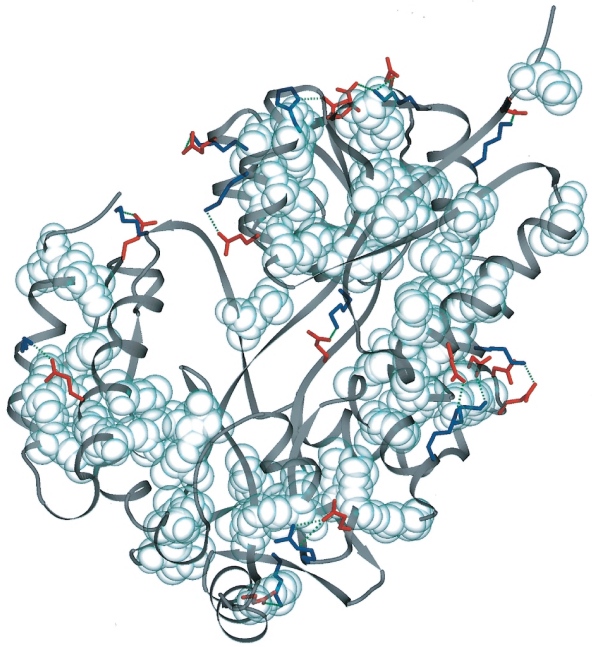

Figure 7. Salt-bridges and leucine/isoleucine residues within the *Pfu*MBP structure. Salt-bridges are shown as pairs of oppositely charged residues (acidic in red, basic in blue) joined by hydrogen bonds (green). Leucine and isoleucine residues are shown as pale blue spherical models.

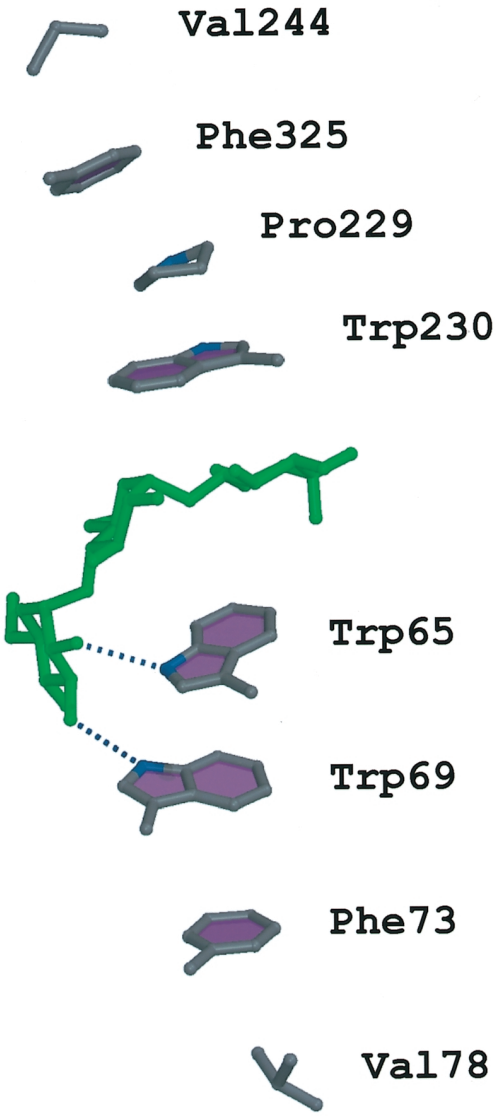

Figure 8. A prominent hydrophobic stack in the *Pfu*MBP structure. The bound ligand is shown in green.

temperature (Robinson *et al.*, 1988; Rittenhouse & Marcus, 1984), whereas *Eco*MBP has only one. Thus, there is no indication that *Pfu*MBP has been under selective pressure to minimize the number of thermolabile amino acids and Asp-Pro bonds that it contains.

Discussion

The overall folds of *Pfu*MBP and *Eco*MBP are very similar. The extreme thermal stability of *Pfu*MBP is most likely due to a combination of factors, mainly efficient hydrophobic interactions, such as isoleucine clustering, an increased number of salt-bridges, and the presence of proline residues in many key positions. The similarity of folds does not, however, result in identical carbohydrate-binding sites. Before discussing the carbohydrate-binding site in detail, it is important to clarify the nature and origin of the ligand observed in the structure, because the carbohydrate-binding site of the protein clearly contains maltotriose rather than maltose, even though the latter sugar was used for elution of the protein and formation of the complex.

There are at least two explanations for this phenomenon. One possibility is that *Pfu*MBP has enzymatic properties that enable it to convert maltose into maltotriose. Alternatively, *Pfu*MBP may have preferentially accumulated this trisaccharide from the solutions used in elution and dialysis. The evidence presented here strongly supports the second hypothesis. The elution of *Pfu*MBP from the amylose column was accomplished using commercial D-maltose (ICN, biotechnology grade). Commercial maltose is known to contain small quantities of maltotriose and higher oligoglucosides. The particular batch of maltose we used for chromatography was 97% pure, according to the manufacturer. NMR and MS measurements, as well as paper chromatography (data not shown), indicate that this batch of maltose contained 0.5 to 1% of maltotriose, along with some higher oligosaccharides. Normally, such a small fraction of contaminants is of no consequence, especially when the saccharide is used as a carbon source for

bacterial culture or as a reagent for eluting *Eco*MBP fusion proteins from amylose affinity resin during purification. In this case, however, it is appears that the much greater affinity of *Pfu*MBP for maltotriose than for maltose resulted in the preferential binding of maltotriose from solutions containing both sugars. Maltose was used in ~1000-fold molar excess during elution, so there was more than enough maltotriose present to bind all of the available protein. We would like to emphasize that it was not necessary to add any carbohydrate to the protein in order to crystallize the complex, which further illustrates the point that the ligand bound to *Pfu*MBP during elution from the amylose affinity column and did not dissociate during subsequent purification and dialysis steps. Only after denaturation, size-exclusion chromatography and refolding could we produce a batch of *Pfu*MBP that was consistently free of ligand.

Conceptually, the binding of maltotriose to *Pfu*MBP resembles the mode of binding of oligosaccharides to *Eco*MBP. In both cases, the sugar is sandwiched between two layers of hydrophobic residues and is held in position by a number of hydrogen bonds (Figure 5). A striking difference emerges, however, when the geometry of the *Pfu*MBP-maltotriose compex is compared with that of the *Eco*MBP-maltotriose complex. The reducing glucopyranoside ring of the maltotriose in complex with *Eco*MBP corresponds to the middle glucopyranoside ring of the maltotriose in the *Pfu*MBP complex, in which the reducing ring is accommodated by an extension of the binding site toward the core of the protein (Figure 5). This enlargement of the cavity is caused by the concerted change of *Eco*MBP amino acids in the (12-19) loop, resulting in the inward motion of the whole loop, which is accompanied by multiple changes in other parts of the protein. As a result, all three pyranoside rings of maltotriose engage in multiple interactions with *Pfu*MBP. In contrast, the terminal glucopyranoside ring of maltotriose makes relatively few contacts with *Eco*MBP, as can be seen from the X-ray structure of the complex (Quiocho *et al*., 1997).

The equilibrium binding constant of the *Pfu*MBP-maltotriose complex (2.7×10^7 M^{-1}) is 20 times greater than that of the *Eco*MBP-maltotriose complex (1.3×10^6 M^{-1}) (Thomson *et al*., 1998), which is not surprising in light of the differences between them. It is likely that tight binding is a requirement for high-fidelity recognition and transport of oligosaccharides by *P. furiosus* at its ambient temperature. On the other hand, maltotriose in the carbohydrate-binding site of *Pfu*MBP has considerable exposure to the solvent, whereas in *Eco*MBP the sugar is buried much more deeply (Figure 5). These results suggest that *Pfu*MBP is not as flexible as *Eco*MBP, or at least that the equilibrium between the open and closed forms of the protein is shifted toward the closed form. If this hypothesis is correct, then the high solvent exposure of the sugar can be explained as a result of the binding mode being more like a ''lock and key'' than a conventional periplasmic transporter mechanism in which the protein closes around the bound ligand. The reduced flexibility of *Pfu*MBP may contribute to its thermostability. It is also worth mentioning that *Pfu*MBP accommodates both anomeric forms of maltotriose, unlike *Eco*MBP and other proteins of this family, which exhibit a preference for specific anomers (Gehring *et al*., 1991).

While this manuscript was in preparation, we learned that the protein corresponding to the second MBP-like ORF in *P. furiosus* (Utah genome center ID Pf_1616024) has been expressed and crystallized in complex with trehalose by the group of Wolfram Welte (University of Konstanz, Germany). As discussed above, the sequence of this protein is identical with *Tli*MBP and, thus, we will use the same abbreviation in order to avoid confusion. An exchange of coordinates prior to publication has allowed us to draw some conclusions from the structure derived by Welte and co-workers. The overall fold of the protein is very similar to the fold of *Eco*MBP and *Pfu*MBP, but again, the proteins have significant differences in the position and orientation of many individual secondary structure elements. The sugar-binding site of *Tli*MBP cannot accommodate oligosaccharides larger than maltose or trehalose (W. Welte *et al*., personal communication). Our preliminary calorimetric studies indicate that *Pfu*MBP does not bind maltose, and the structure of the protein strongly suggests that binding of trehalose would also be thermodynamically unfavorable. Because both *Pfu*MBP and *Tli*MBP are found in *P. furiosus*, it is logical to assume that the maltose/trehalose recognition and transport, and the recognition and transport of higher oligoglucosides are for some reason carried out by two separate systems in this organism. The presence of *Tli*MBP in *P. furiosus* may be explained by a lateral transfer of genetic material (Pennisi, 1999), which could have allowed its native MBP to evolve into the protein targeted toward longer oligosaccharides.

Remarkably, *Eco*MBP, *Pfu*MBP, and *Tli*MBP all have prominent hydrophobic patches on their surfaces in approximately the same location, on the opposite face from the ligand-binding site (Figure 4(b) and data not shown). Thus far, no biological function has been ascribed to this region of *Eco*MBP. The identity and precise arrangement of these solvent-exposed, hydrophobic side-chains are not highly conserved, but the existence of such an unusually large hydrophobic patch on all three of these very distantly related proteins is unlikely to be a coincidence. Perhaps this part of the surface constitutes a heretofore unrecognized protein-interaction site. It might be fruitful to investigate the impact of amino acid substitutions in this region of *Eco*MBP on its biological activity *in vivo*.

Interestingly, an MBP from another thermophilic bacterium, *Thermotoga maritima* (*Tma*MBP), exhibits saccharide-binding thermodynamic properties similar to those of *Eco*MBP (Wassenberg *et al.*, 2000), suggesting that there is a greater degree of structural similarity between *Tma*MBP and *Eco*MBP than between *Tma*MBP and *Pfu*MBP or *Tli*MBP. The binding of both maltose and maltotriose to *Eco*MBP and *Tma*MBP is endothermic and exergonic, driven primarily by entropy changes (Thomson *et al.*, 1998; Wassenberg *et al.*, 2000). In contrast, no detectable thermal effect accompanied the addition of maltose to *Pfu*MBP at several temperatures (Figure 6 and data not shown), whereas the binding of maltotriose was decidedly exothermic. Analysis of the binding curves indicates that this reaction is both enthalpically and entropically driven. A full explanation of the thermodynamics of *Pfu*MBP-carbohydrate interactions requires an in-depth study, which will be described elsewhere. We are attempting to obtain a structure of the ligand-free form of *Pfu*MBP in order to analyze the conformational changes that accompany ligand binding.

Materials and Methods

Vector construction

To construct pKM800, the ORF encoding the putative *Pfu*MBP (Utah gene center ID Pf_1788981) was amplified from genomic DNA by PCR with primers PE-472 (5′-CCTCCCATATGAAAATCGAAGAAGGAAAAGTT GTTATTTGGCATGCAATG-3′) and PE-473 (5′-CAGC CTGGATCCATTATCCTTGCATGTTGTTAAGGATTTC TTG-3′). The PCR product was digested with *Bam*HI and *Nde*I, and then ligated with the *Bam*HI/*Nde*I vector backbone of pMal-C2x (New England Biolabs, Beverly, MA). The nucleotide sequence of the insert was verified experimentally.

Protein expression

pKM800 was introduced into *E. coli* CT14 cells, which do not produce any endogenous MBP (Scheffler *et al.*, 1994). The cells also contained pDC952 (Calderone *et al.*, 1996), a plasmid that constitutively overproduces the cognate tRNA for AGA and AGG codons (*argU*). Cells from a single drug-resistant colony were grown to saturation in LB broth (Miller, 1972) supplemented with 100 μg/ml ampicillin and 30 μg/ml chloramphenicol at 37 °C. The saturated culture was diluted 66-fold in the same medium and grown in shake-flasks to mid-log phase (A_{600} 0.5-0.7), at which time IPTG was added to a final concentration of 1 mM. After three to four hours, the cells were recovered by centrifugation and frozen at −80 °C.

Protein purification

Approximately 20 g of wet cell paste was resuspended in 200 ml of 20 mM Tris-HCl (pH 7.4), 200 mM NaCl. Just prior to lysis, phenyl methylsulphonyl fluoride (PMSF), EDTA, and benzamidine were added to 1, 2, and 2.5 mM, respectively. The cells were lysed by sonication (Sonics and Materials Vibracell, 1/2'' SM0896 horn) in 50 ml aliquots for 45 seconds at a power level of 65 %. The sonication was repeated three times with swirling of the sample in a large beaker in an ice-water bath to keep it cold. Polyethylenimine (Sigma, P3143) was then added to 0.1 %, and the lysate was apportioned into six Oakridge tubes (Nalgene, 3114-0050) and centrifuged in an SA-600 rotor at 37,000 ***g*** for ten minutes. Solid ammonium sulfate was added to the supernatant to bring it to 35 % saturation. Once the salt had dissolved, the sample was clarified by centrifugation again, as above. Next, more ammonium sulfate was added to the supernatant to bring it to 75 % saturation. After 15 minutes on ice, the precipitate was recovered by centrifugation, as above, and resuspended in 200 ml of 20 mM Tris (pH 7.4), 200 mM NaCl. The solution was clarified again by centrifugation and then passed through a 0.45 μm filter prior to chromatography.

The sample was next applied to an amylose column (5 cm × 12 cm, amylose-NEB, XK-50, Pharmacia), washed with 1 l of 20 mM Tris-HCl (pH 7.6), 200 mM NaCl, 1 mM EDTA, and eluted with 10 mM maltose (ICN Biochemicals, biotechnology grade). Further purification of the ligand-bound form of *Pfu*MBP was accomplished by hydrophobic interaction chromatography. Solid ammonium sulfate was added to 1.0 M, and then the sample was filtered before applying it to a butyl-Sepharose column (1.6 cm × 10 cm, Pharmacia). The column was washed with three volumes of 1.0 M ammonium sulfate, 30 mM sodium phosphate (pH 6.2). *Pfu*MBP was eluted from the column with 0.8 M ammonium sulfate, 30 mM sodium phosphate (pH 6.2).

To prepare ligand-free *Pfu*MBP, after amylose affinity chromatography (above) the protein was dialyzed in 20 mM sodium acetate (pH 5.2), and then solid guanidinium thiocyanate (which is a considerably stronger denaturing agent than guanidinium chloride) was added to approximately 5.5 M. Dithiothreitol (DTT) was added to 2 mM, and the solution was left at room temperature for three days. Denaturation was suggested by the fact that rapid dilution into water at room temperature led to nearly complete aggregation of the protein. The material was then concentrated by ultrafiltration in a stirred cell (Amicon) and applied in 2.5 ml aliquots to a stack of three desalting columns (G-25SF Sephadex, 5-ml Hitrap Desalting, Pharmacia) using a running buffer of 20 mM sodium acetate (pH 5.4), 5.5 M guanidinium thiocyanate at a flow-rate of 1 ml/minute at room temperature. The protein peak was collected for each run, pooled, and used for refolding.

Protein refolding was carried out on a heat block with an indicated temperature of 50 °C. The starting concentration of protein was 2.5 mg/ml. A 1.5 ml aliquot of the protein solution was added to a clean 16 mm × 125 mm glass tube and incubated at 50 °C for ten minutes. Deionized, filtered water was equilibrated at this temperature for at least 20 minutes before use. To initiate refolding, 5 ml of 50 °C water was slowly added to the denaturant solution with vigorous stirring. The diluted protein solution was incubated for 10-15 minutes at 50 °C before another 8 ml of 50 °C water was added and the tube was allowed to cool slowly to room temperature. The refolded protein was concentrated to 0.5 mg/ml and dialyzed against water at room temperature to remove any residual thiocyanate. Removal of the ligand was confirmed by calorimetric experiments and by adsorption of the protein to amylose resin.

The refolded *Pfu*MBP was adjusted to 15 mM Tris-HCl (pH 7.4), 25 mM NaCl, and approximately 2 mM

(tris)carboxhyethyl phosphine hydrochloride (Pierce) was added to reduce any intermolecular disulfide bonds that may have formed. After 30 minutes at room temperature, this material was applied to a MonoQ column (1.6 cm × 60 cm, Pharmacia) and a linear gradient from 25 mM to 750 mM NaCl in 20 mM Tris (pH 7.4) was applied. The peak fractions were next adjusted to 1.5 M ammonium sulfate and applied to an 8 ml column of Source ISO HIC resin (1.6 cm × 4 cm, HR16/5, Pharmacia). Using the same starting buffer described above for the butyl separation, a gradient of 1.0 M to 0.5 M ammonium sulfate (pH 6.2, 30 mM sodium phosphate) was run in 13 column volumes, followed by ten column volumes of 0.5 M ammonium sulfate to elute the *Pfu*MBP. The protein obtained in this way had the correct mass (as measured by electrospray MS) within the error of the instrument. Amino acid composition analysis did not reveal any amino acid derivatization.

Crystallization

Several alkaline (pH 8-9), high-salt (2.0-2.4 M ammonium sulfate) crystallization conditions were found for complex of *Pfu*MBP with maltotriose (non-refolded) using a sparse-matrix approach (Jancarik & Kim, 1991). At that time, we thought the ligand-binding site was occupied by maltose, and so the protein was maintained in a solution containing 10 mM maltose. However, after the structure was solved and it became clear that the ligand was actually maltotriose, we established that the addition of maltose was not necessary for crystallization. In all cases, the crystals were orthorhombic prisms, which appeared within two to four days. Diffraction-quality crystals were obtained by hanging-drop equilibration of 6 mg/ml protein-carbohydrate complex mixed with an equal volume of reservoir solution against 2.4 M ammonium sulfate, 0.1 M bicine (pH 9.0), 0.1% (v/v) β-mercaptoethanol. Crystals appeared within two to four days and reached an average size of 0.5 mm × 0.3 mm × 0.3 mm in two weeks.

In contrast, the carbohydrate-free (refolded) protein crystallized in the form of very thin needles from 30% (w/v) PEG 4000 (pH 5.0) or from 2 M sodium phosphate (pH 4.0). If this sugar-free protein was incubated with 5 mM maltotriose, then crystals of the complex could be obtained under the same conditions as outlined above for the non-refolded protein.

Data collection, derivatives and structure solution

A single crystal of the *Pfu*MBP complex, measuring 0.3 mm × 0.2 mm × 0.2 mm, was immersed in Paratone-N oil, mounted in a monofilament loop, and flash-frozen in a cryogenic nitrogen stream (Oxford Cryostream) at 100 K. X-ray diffraction was recorded using a MAR-345 image plate mounted on a Rigaku X-ray generator. The crystal belonged to space group $P2_12_12_1$. A crystal-to-detector distance of 100 mm and an exposure time of 60 minutes/degree were used to collect data over 90° in 0.25° oscillations. The data were reduced and scaled using HKL software (Otwinowski & Minor, 1997) (see Table 3 for details of data collection and processing).

An attempt to solve the structure by molecular replacement, using both the sugar-bound and sugar-free forms of *Eco*MBP, was unsuccessful. Therefore, heavy-atom derivatives were prepared by soaking the *Pfu*MBP crystals in artificial mother liquor containing 2.2 M ammonium sulfate, 100 mM bicine (pH 9.0), and 1 mM of the heavy-atom reagent. Two derivatives were used: MeHgOAc (soaked for 40 minutes) and K_4PtI_6 (soaked for 25 minutes). Derivative data collection procedures were the same as for the native crystal; a summary is given in Table 3.

The heavy-atoms were located in isomorphous and anomalous Patterson maps, calculated with the program CNS (Brünger *et al.*, 1998), and verified using the program SOLVE (Terwilliger & Berendzen, 1999). Two-derivative isomorphous/anomalous phasing of the native dataset to 3.0 Å was performed with CNS (Table 3) and was followed by density modification using the same program. At this point, continuous electron density was observed for the protein monomer and secondary-structure elements were clearly visible in the electron density map.

Model building and refinement

The first model of the protein was manually built into the electron density using the program O (Jones *et al.*, 1991). Several cycles of simulated annealing were performed with CNS, after which all further refinement was performed using Shelx-97 (G. M. Sheldrick (1997) University of Grottingen; Sheldrick & Schneider, 1997). A total of 380 amino acid residues were built and refined (the C-terminal glycine residue was not observed in the density, most likely due to disorder).

Table 3. Statistics of data collection and structure solution

	Native	MeHgOAc	K_4PtI_6
a,b,c (Å)	129.07, 68.57, 61.26	128.12, 69.43, 61.22	130.30, 68.20, 61.23
Resolution (Å)	100-1.85	100-2.2	100-2.2
Completeness (%)	91.88	91.5	96.0
>2σ (%)	80	85	83
Redundancy	4.5	3.8	4.2
Unique reflections	42,983	26,199	27,407
$I/\sigma I$	20.08	17.8	10.6
R_{merge} (%)	5.3	9.1	10.1
$R_{derivative/native}$ (%)	-	19.3	21.2
Number of sites	-	2 strong + 2 weak	1 strong + 1 weak
Phasing power	-	0.56	0.93
Figure of merit	0.43[a]	0.2	0.3

[a] Both derivatives to 3.0 Å resolution, before density modification.

The electron density in the sugar-binding site improved with successive refinement cycles, until it reached the state depicted in Figure 4. At this point, it was apparent that three D-glucopyranoside rings were occupying the site. Three glucopyranoside units were built into the density, using coordinates from Hic-Up (Kleywegt & Jones, 1998), and were refined independently. The anomeric oxygen atom of ring 1 was built as disordered between α and β configurations, since electron density was observed for both. Disordered anomeric oxygen components were refined with the sum of their occupancies restrained to unity. Carbohydrate rings 2 and 3 were built as α anomers. Cysteine 33 had undergone covalent modification by β-mercaptoethanol into Cys-S-S-CH_2-CH_2-OH. An appropriate residue was built into the density using coordinates from Hic-Up (Kleywegt & Jones, 1998). At the final stage of the refinement, 386 solvent molecules were added to the structure, using difference density and standard proximity criteria. A summary of refinement details is given in Table 4.

Electrostatic and hydrophobicity plots were prepared using the program GRASP (Nicholls *et al.*, 1991). The detection of cavities capable of accommodating a water molecule was undertaken using VOIDOO (Kleywegt & Jones, 1994) with default settings. In both cases, the ligand was included in the calculations. The number of hydrogen bonds was calculated using HBPLUS (McDonald & Thornton, 1994). Only protein-protein hydrogen bonding was considered.

Isothermal titration calorimetry

Ligand-free (refolded) *Pfu*MBP was dialyzed against three changes of 100 volumes of 25 mM Tris-HCl (pH 7.5), 30 mM NaCl. The protein concentration was adjusted to 0.05 mM, on the basis of the specific absorbance at 280 nm, established by quantitative amino acid analysis. An aliquot of this solution was degassed and loaded into the 1.4915 ml chamber of a MicroCal VP-ITC calorimeter and allowed to equilibrate to 15 °C. The protein was titrated (25 injections of 10 μl) with the same buffer containing 0.35 mM carbohydrate. The same titrations were performed at 40 °C and at 75 °C (data not shown) to confirm that binding of maltose to *Pfu*MBP was not masked by the fact that the ΔH happens to be close to zero at one particular temperature. The maltose and maltotriose used in this part of the study (Sigma, analytical grade) were of much higher quality than the maltose that was used for protein chromatography, as gauged by NMR and paper chromatography (data not shown). A single-site binding model in the Origin VI software package (MicroCal Inc.) was used to obtain thermodynamic parameters from the experimental results.

Table 4. Statistics of the refinement

R_{all} (%)	19.6
$R_{>4\sigma}$ (%)	18.4
R_{free} (%)[a]	22.9
Data/parameter	4.1
Ramachandran preferred (%)	94
Ramachandran allowed (%)	6
rmsd bond lengths (Å)	0.07
rmsd bond angles (deg.)	1.2

[a] Reflections for R_{free} were a randomly selected 5 % (2253) of the working dataset.

Protein Data Bank accession number

The coordinates and structure factors were deposited with the RCSB Protein Data Bank (Berman *et al.*, 2000) under accession code 1ELJ .

Acknowledgments

A.E. gratefully acknowledges a Cancer Research Training Award from NIH. We thank Dr Jocelyn DiRugierrio (University of Maryland) for her kind gift of *P. furiosus* genomic DNA, Dr Sergey Tarasov for help with the ITC experiments, Dr Tad Guszczynski for protein sequencing and amino acid analysis, Anne Arthur for expert editorial assistance, and our colleagues in the Macromolecular Crystallography Laboratory, especially Dr Alexander Wlodawer, for their support and thoughtful comments on the manuscript.

References

Bacon, D. & Anderson, W. F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219-220.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. *Nucl. Acids Res.* **28**, 235-242.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.

Calderone, T. L., Stevens, R. D. & Oas, T. (1996). High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in *Escherichia coli*. *J. Mol. Biol,* **262**, 407-412.

Croop, J. M. (1998). Evolutionary relationships among ABC transporters. *Methods Enzymol.* **292**, 101-116.

Danson, M. J. & Hough, D. W. (1998). Structure, function and stability of enzymes from the Archaea. *Trends Microbiol.* **6**, 307-314.

Ehrmann, M., Ehrle, R., Hofmann, E., Boos, W. & Schlosser, A. (1998). The ABC maltose transporter. *Mol. Microbiol.* **29**, 685-694.

Elcock, A. H. (1998). The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* **284**, 489-502.

Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 132-134.

Ganesh, C., Shah, A. N., Swaminathan, C. P., Surolia, A. & Varadarajan, R. (1997). Thermodynamic characterization of the reversible, two-state unfolding of maltose binding protein, a large two-domain protein. *Biochemistry,* **36**, 5020-5028.

Gardner, K. H., Zhang, X., Gehring, K. & Kay, L. (1998). Solution NMR studies of a 42-kDa *Escherichia coli* maltose binding protein/β-cyclodextrin complex:

chemical shift assignments and analysis. *J. Am. Chem. Soc,* **120**, 11738-11748.
Gehring, K., Williams, P. G., Pelton, J. G., Morimoto, H. & Wemmer, D. E. (1991). Tritium NMR spectroscopy of ligand binding to maltose-binding protein. *Biochemistry,* **30**, 5524-5531.
Hall, J. A., Gehring, K. & Nikaido, H. (1997a). Two modes of ligand binding in maltose-binding protein of *Escherichia coli*. Correlation with the structure of ligands and the structure of binding protein. *J. Biol. Chem.* **272**, 17605-17609.
Hall, J. A., Thorgiersson, T. E., Liu, J., Shin, Y.-K. & Nikaido, H. (1997b). Two modes of ligand binding in maltose-binding protein of *Escherichia coli*. Electron paramagnetic resonance study of ligand-induced global conformational changes by site-directed spin labeling. *J. Biol. Chem.* **272**, 17610-17614.
Hall, J. A., Ganesan, A. K., Chen, J. & Nikaido, H. (1997c). Two modes of ligand binding in maltose-binding protein of *Escherichia coli*. Functional significance in active transport. *J. Biol. Chem.* **272**, 17615-17622.
Higgins, C. F. (1992). ABC transporters: from microorganisms to man. *Annu. Rev. Cell Biol.* **8**, 67-113.
Horlacher, R., Xavier, K. B., Santos, H., DiRuggiero, J., Kossmann, M. & Boos, W. (1998). Archaeal binding protein-dependent ABC transporter: molecular and biochemical analysis of the trehalose/maltose transport system of the hyperthermophilic archaeon *Thermococcus litoralis*. *J. Bacteriol.* **180**, 680-689.
Jancarik, J. & Kim, S. H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallog.* **24**, 409-411.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110-119.
Kleywegt, G. J. (1996). Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallog. sect. D,* **52**, 842-857.
Kleywegt, G. J. & Jones, T. A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallog. sect. D,* **50**, 178-185.
Kleywegt, G. J. & Jones, T. A. (1998). Databases in protein crystallography. *Acta Crystallog. sect. D,* **54**, 1119-1131.
Korkhin, Y., Kalb(Gilboa), A. J., Peretz, M., Bogin, O., Burstein, Y. & Frolow, F. (1998). NADP-dependent bacterial alcohol dehydrogenases: crystal structure, cofactor-binding and cofactor specificity of the ADHs of *Clostridium beijerinckii* and *Thermoanaerobacter brockii*. *J. Mol. Biol.* **278**, 967-981.
Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
Martineau, P., Saurin, W., Hofnung, M., Spurlino, J. C. & Quiocho, F. A. (1990). Progress in the identification of interaction sites on the periplasmic maltose binding protein from *E. coli*. *Biochimie,* **72**, 397-402.
Matthews, B. W., Nicholson, H. & Becktel, W. J. (1987). Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc. Natl Acad. Sci. USA,* **84**, 6663-6667.
McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777-793.
Merrit, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0: a program for photorealistic molecular graphics. *Acta Crystallog. sect. D*, **50**, 869-873.
Miller, J. H. (1972). *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281-296.
Novokhatny, V. & Ingham, K. (1997). Thermodynamics of maltose binding protein unfolding. *Protein Sci.* **6**, 141-146.
Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276A**, 307-326.
Pennisi, E. (1999). Is it time to uproot the tree of life? *Science,* **284**, 1305-1307.
Perutz, M. F. & Raidt, H. (1975). Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. *Nature,* **255**, 256-259.
Quiocho, F. A., Spurlino, J. C. & Rodseth, L. E. (1997). Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure,* **5**, 997-1015.
Rittenhouse, J. & Marcus, F. (1984). Peptide mapping by polyacrylamide gel electrophoresis after cleavage at aspartyl-prolyl peptide bonds in sodium dodecyl sulphate-containing buffers. *Anal. Biochem.* **138**, 442-448.
Robinson, D. R., Watts, N. R. & Coombs, D. H. (1988). Heat cleavage of bacteriophage T4 gene 23 product produces two peptides previously identified as head proteins. *J. Virol.* **62**, 1723-1729.
Russell, R. J. M., Ferguson, J. M. C., Hough, D. W., Danson, M. J. & Taylor, G. L. (1997). The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry,* **36**, 9983-9994.
Scheffler, J. E., Waugh, D. S., Bekesi, E., Kiefer, S. E., LoSardo, J. E., Neri, A., Prinzo, K. M., Tsao, K.-L., Wegrzynski, B., Emerson, S. D. & Fry, D. C. (1994). Characterization of a 78-residue fragment of c-Raf-1 that comprises a minimal binding domain for the interaction with Ras-GTP. *J. Biol. Chem.* **269**, 22340-22346.
Sharff, A. J., Rodseth, L. E., Spurlino, J. C. & Quiocho, F. A. (1992). Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry,* **31**, 10657-10663.
Sharff, A. J., Rodseth, L. E. & Quiocho, F. A. (1993). Refined 1.8-Å structure reveals the mode of binding of beta-cyclodextrin to the maltodextrin binding protein. *Biochemistry,* **32**, 10553-10559.
Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high resolution refinement. *Methods Enzymol.* **277**, 319-343.
Spurlino, J. C., Lu, G. Y. & Quiocho, F. A. (1991). The 2.3-Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* **266**, 5202-5219.

Tatusov, R. L., Koonin, E. V. & Lipman, D. J. (1997). A genomic perspective on protein families. *Science,* **278**, 631-637.

Terwilliger, T. C. & Berendzen, J. (1999). Automated structure solution for MIR and MAD. *Acta Crystallog. sect. D*, **55**, 849-861.

Thomson, J., Liu, Y., Sturtevant, J. M. & Quiocho, F. A. (1998). A thermodynamic study of the binding of linear and cyclic oligosaccharides to the maltodextrin-binding protein of *Escherichia coli*. *Biophys. Chem.* **70**, 101-108.

Wassenberg, D., Liebl, W. & Jaenicke, R. (2000). Maltose-binding protein from the hyperthermophilic bacterium *Thermotoga maritima*: stability and binding properties. *J. Mol. Biol.* **295**, 279-288.

Wu, H. C. (1996). Biosynthesis of lipoproteins. In Escherichia coli *and* Salmonella: *Cellular and Molecular Biology* (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M. & Umbarger, E., eds), pp. 1005-1014, American Society for Microbiology, Washington, DC.

Edited by D. Rees

(Received 26 June 2000; received in revised form 29 September 2000; accepted 30 September 2000)