FEBS Letters 568 (2004) 94–98 FEBS 28477

Binding of phage-display-selected peptides to the periplasmic chaperone protein SurA mimics binding of unfolded outer membrane proteins

Eduard Bitto[1], David B. McKay*

Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

Received 8 April 2004; revised 4 May 2004; accepted 4 May 2004

Available online 20 May 2004

Edited by Stuart Ferguson

Abstract SurA is a periplasmic chaperone protein that facilitates maturation of integral outer membrane proteins (OMPs). Short peptides that bind SurA have previously been characterized. In this work, an enzyme-linked immunoabsorbent assay-based competition assay is utilized to demonstrate that binding of such peptides, presented by peptide-tagged phage, mimics binding of biological substrates. Two representative unfolded OMPs, OmpF and OmpG, bind SurA and a core structural fragment thereof in competition with peptide-tagged phage, and with the same order-of-magnitude affinity as the peptides. Additionally, unfolded OmpF and OmpG bind SurA more tightly than an unfolded water-soluble protein, while folded proteins have no measurable affinity, demonstrating a specificity of SurA for OMP polypeptides.

Keywords: Survival protein A (SurA); Periplasmic molecular chaperone; Outer membrane protein folding; Peptidyl-prolyl isomerase; Enzyme-linked immunoabsorbent assay

1. Introduction

The SurA protein is a bacterial periplasmic molecular chaperone that facilitates correct folding of integral outer membrane proteins (OMPs) [1–3]. Although SurA sequences typically include two (or in a few cases, only one) prolyl isomerase domains, it has been demonstrated with *E. coli* (*Escherichia coli*) SurA that the prolyl isomerase domains are dispensable for biological activity, while the remainder of the protein is essential [4]. The X-ray crystallographic structure of *E. coli* SurA reveals a bipartite structure consisting of core module which includes the first prolyl isomerase domain and a second satellite prolyl isomerase domain (P2) tethered ~25–30 Å distant from the core module [5]. The core domain has an extended, deep crevice which could accommodate segments of unfolded polypeptides, suggesting a scenario in which SurA may bind and stabilize unfolded OMPs at some stage during their transmembrane translocation, folding, and assembly.

Peptide phage display selection experiments with SurA have revealed a binding preference for a consensus peptide of the form aromatic(Ar)-polar–Ar-non-polar-proline. A protein construct in which the satellite prolyl isomerase domain was deleted, SurA(ΔP2), showed a similar preference. Although pentameric sequences fitting this consensus are rare in bacterial OMPs, a less restrictive target of the form Ar–random–Ar (Ar–X–Ar) is common, being prevalent in the "aromatic bands" that encircle the antiparallel β barrel structures of OMPs. These results suggested that SurA and its core module bind preferentially to a peptide sequence motif that is prevalent in OMPs [6].

A question that arises from phage display selection experiments is whether the peptides that are selected are accurate mimics of natural substrates. Since the experiment selects on the basis of peptide affinity for the target protein, sequences that emerge may have artifactually high affinity. Alternatively, if the target protein recognizes extended polypeptide segments in natural substrates, then the 7-mer peptides that were screened in this particular case may under-represent the natural binding activity. In the extreme case, peptides selected through phage display may bind the target protein at a site distinct from the natural substrate binding site. To address the question of whether the SurA-binding peptides selected by phage display are legitimate mimics of biological substrates, we have utilized a competition assay to demonstrate that (a) two representative unfolded OMPs bind SurA and SurA(ΔP2) competitively with, and with affinities that are the same order-of-magnitude as, peptides selected by phage display; and further, (b) unfolded OMPs bind with significantly higher affinity than either a representative unfolded non-membrane protein or folded proteins.

2. Materials and methods

2.1. Protein expression and purification

Mature *E. coli* SurA protein (amino acids 21–421; SwissProt P21202) and SurA lacking the second (P2) parvulin-like domain (SurA(ΔP2), residues 21–281 and 390–428) were expressed and purified as described previously [5,6]. The ATPase fragment of bovine 70 kDa heat shock

* Corresponding author. Fax: +1-650-723-8464.
E-mail address: Dave.McKay@Stanford.edu (D.B. McKay).

[1] Present address: Department of Biochemistry, University of Wisconsin, Madison, WI 53706, USA.

Abbreviations: OMP, outer membrane protein; *E. coli*, *Escherichia coli*; SurA(ΔP2), SurA protein from which residues 282–389, spanning the second prolyl isomerase domain, have been deleted; Ar, aromatic; GST, glutathione-*S*-transferase; RCMLA, reduced carboxymethylated lactalbumin; PMSF, phenylmethyl sulfonyl fluoride; ELISA, enzyme-linked immunoabsorbent assay

0014-5793/$22.00 © 2004 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2004.05.014

cognate protein (Hsc70 ATPase fragment) was expressed and purified as described [7]. Glutathione-*S*-transferase (GST) was expressed from the pGEX-2T vector and purified on a glutathione sepharose essentially following the manufacturer's protocol (Amersham Biosciences). Reduced carboxymethylated lactalbumin (RCMLA) was purchased from Sigma Chemical. The heptapeptide of sequence WEYIPNV was synthesized and purified by the Stanford Protein and Nucleic Acid Facility.

Gene fragments encoding mature *E. coli* OmpF and OmpG proteins (lacking the signal sequence) were PCR-amplified from genomic DNA of *E. coli* strain K-12 MG1655 [8] and cloned into a pET29b vector (Novagen) using the *Nde*I and *Eco*RI restriction sites of the vector. Since trial purification steps on Ni–NTA resin were not particularly advantageous, the hexahistidine coding sequences were subsequently deleted from the vectors and only the native OmpF and OmpG sequences were expressed. Purification of OmpG protein followed a published protocol [9] with modifications. The expression plasmid was transformed into *E. coli* strain BL21(DE3) and maintained with 34 μg/mL kanamycin in Luria–Bertani media. Cells were grown at 37 °C to $A_{600} \sim 0.7$, then protein expression was induced by addition of IPTG to 0.4 mM final concentration, and incubation was continued for an additional 3 h at 37 °C. Cells were pelleted by centrifugation; cell pellets were resuspended in TBS buffer (50 mM Tris–HCl, pH 7.5, 150 mM NaCl, and 0.5 mM phenylmethyl sulfonyl fluoride (PMSF)) and sonicated with two bursts of 8-min duration. Soluble and insoluble fractions were separated by centrifugation at $17\,000 \times g$ for 40 min. The pellet fraction containing ompG in inclusion bodies was suspended in 10 mM Tris–HCl, pH 7.5, 1% Triton X-100, and 1 M urea and re-centrifuged. The pellet was then suspended in 10 mM Tris–HCl, pH 7.5, and 1 M urea and centrifuged again. The remaining pellet was solubilized in loading buffer (10 mM Tris–HCl, pH 7.5, and 8 M urea) and clarified by centrifugation and filtration through 0.2-μm filter. The soluble material was applied on 20 mL DEAE–Sephacel column pre-equilibrated in the same buffer. The column was washed with loading buffer, followed by loading buffer plus 100 mM NaCl. Denatured OmpG protein was eluted from the column by a step gradient of loading buffer plus 300 mM NaCl. A second, essentially pure fraction could be recovered with a high-salt wash of 2 M NaCl. The eluted fraction containing OmpG was dialyzed extensively against 10 mM Tris–HCl, pH 7.5, concentrated, and stored frozen at −20 °C.

Expression of OmpF followed the protocol described above for OmpG, with the minor difference that cells were incubated at 25 °C for 4 h after induction of expression with IPTG. Cell pellet was resuspended in buffer A (20 mM Tris–HCl, pH 8, and 10 mM EDTA) supplemented with 0.5 mM PMSF and sonicated two times for 8 min. Soluble and insoluble fractions were separated by centrifugation. The insoluble fraction containing unfolded ompF in inclusion bodies was washed successively in (a) 50 mL, buffer A supplemented with 0.4% Triton X-100 and 0.4% sodium deoxycholate; (b) 50 mL, buffer A supplemented with 0.8% Triton X-100 and 0.8% sodium deoxycholate; and (c) 50 mL, buffer A with 5 M urea, with soluble and insoluble fractions separated by centrifugation after each wash. The resulting pellet was solubilized in 25 mL of buffer A + 8 M guanidine hydrochloride and clarified by centrifugation and filtration through 0.2-μm filter. 75 mL of acetone was added to last solution and the resulting heavy precipitate was pelleted and re-dissolved in buffer containing either 8 M urea or 1% SDS. The re-solubilized ompF protein that resulted from this procedure could be dialyzed into buffered solutions containing as little as 0.01% SDS.

2.2. Enzyme-linked immunoabsorbent competition assay

Measurement of relative affinities of different ligands for SurA or SurA(ΔP2) was based on an enzyme-linked immunoabsorbent assay (ELISA) used previously to determine relative affinities of phage-presented peptides for the target protein [6]. In initial trials to validate the assay, phage presenting the heptapeptide WEYIPNV at the amino terminus of the minor coat protein PIII of bacteriophage M13 (New England Biolabs) was used as a signal-reporting substrate and the peptide WEYIPNV was used as a competing ligand. For subsequent experiments, phage presenting the peptide FTYMPPV was used. Briefly, including variations on the reported procedure: the wells of 96-well ELISA plates were coated overnight at 4 °C with 100 μL per well of 100 μg/mL SurA(ΔP2) or 129 μg/mL SurA protein (i.e., equal volumes of equimolar protein solutions). The following day, protein solutions were discarded and wells were treated for 2 h at 4 °C with "blocking buffer" that included 5 mg/mL BSA to reduce non-specific binding interactions. ELISA plates were washed six times with TBST buffer (50 mM Tris–HCl, 150 mM NaCl, and 0.5% [v/v] Tween 20, pH 7.5). Dilutions of FTYMPPV-tagged phage, along with a constant concentration of competitor substrates of interest (ompG, ompF, WEYIPNV, RCMLA, or no-competitor control), were made in a separate plate which had been treated with blocking buffer; the first well contained 10^{11} pfu per 200 μL of TBST, and 11 sequential threefold dilutions were made, resulting in a range of $\sim 10^6$–10^{11} pfu per 200 μL. These mixtures were then applied to ELISA plates pre-coated with the target protein (SurA or SurA(ΔP2)) and allowed to incubate for 1 h at room temperature. Unbound ligands were removed from the wells by washing, after which horseradish peroxidase-conjugated anti-M13 antibody (Pharmacia #27-9411-01) was pipetted into the wells. After incubation and washing to remove unbound antibody, freshly prepared horseradish peroxidase substrate (36 μL of 30% H_2O_2, added to 21 mL of "ABTS stock": 22 mg 2,2′-Azino-bis

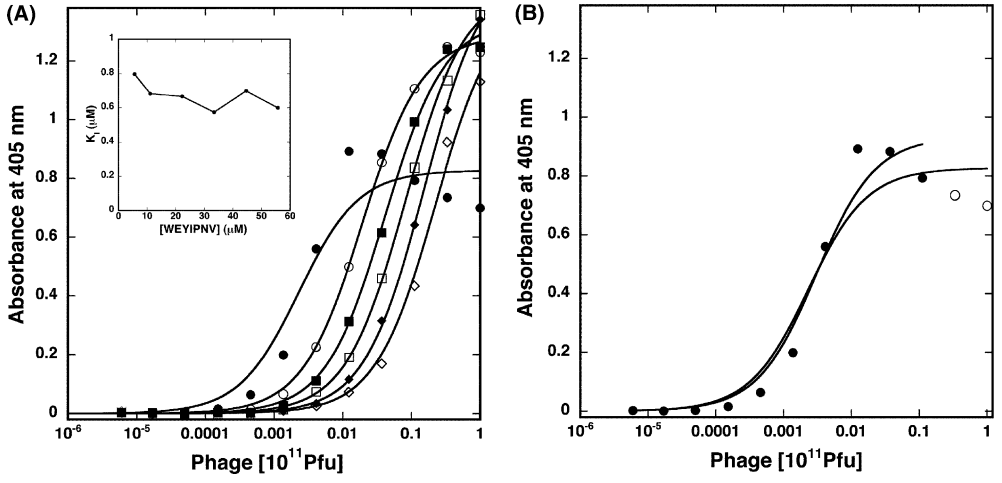

Fig. 1. ELISA competition experiment between WEYIPNV-tagged phage and the peptide WEYIPNV for SurA(ΔP2). Curves are fits of the function $A([\phi]) = A_{max}/(1 + K_{app}/[\phi])$, as described in Section 2. (A) Concentration of WEYIPNV peptide for data points: (●), average of two measurements in absence of competing peptide; (○), 5.6 μM WEYIPNV peptide; (■), 11.1 μM; (□), 22.3 μM; (◆), 44.6 μM; (◊), 55.7 μM. Inset: values for K_I of WEYIPNV peptide computed as described in Section 2. (B) Comparison of fits for data measured in absence of peptide; curve extending further to the right results from fitting function to all data; for curve extending only to phage concentration of 0.1, the two data points shown in open circles were omitted.

(3-ethylbenzthiazoline)-6-sulfonic acid in 100 mL of 50 mM sodium citrate, pH 4.0) was added to the ELISA plate wells, and the reaction was allowed to proceed for approximately 15 min at room temperature. Product of the horseradish peroxidase reaction was monitored as optical absorbance at $\lambda = 405$ nm with a microplate reader. Typically, duplicate rows having FTYMPPV-tagged phage but no additional competitor ligand were included as "no-ligand" controls, while in another row the synthetic peptide WEYIPNV, whose binding to the target proteins has been characterized [6], was often included as a "known ligand" control.

2.3. Data analysis

Data were analyzed using a model of competitive binding between phage-presented peptide and competitor ligand. Data for A_{405} generated by the colorimetric ELISA assay versus phage dilution/concentration [ϕ] were fit with a function of the form:

$$A_{405} = \frac{A_{\max}}{1 + \frac{K_{\text{app}}}{[\phi]}}$$

using MacCurveFit version 1.5 (Kevin Raner Software, Australia), where K_{app} is the apparent relative dissociation constant, corresponding to the phage concentration at half maximal signal $[\phi]_{0.5}$ and $A_{\max}$ is the maximum value of the absorbance. In the absence of competing ligand, K_{app} becomes K_{d}, the relative dissociation constant for phage alone. For assays done in the presence of competing ligand at fixed concentration [I], the ratio $K_{\text{app}}/K_{\text{d}}$, computed as the ratio of the phage concentrations that give half-maximal signal in the presence and absence of ligand, respectively, gives the shift in apparent binding constant of phage-presented peptide due to the competing ligand. For simple competitive inhibition,

$$K_{\text{app}} = K_{\text{d}}\left(1 + \frac{[I]}{K_{\text{I}}}\right)$$

or

$$K_{\text{I}} = \frac{[I]}{\frac{K_{\text{app}}}{K_{\text{d}}} - 1}$$

where K_{I} is the inhibition constant of the competing ligand. For each new ligand, a survey of concentrations was used initially to find assay conditions where $[I] \sim K_{\text{I}}$. Datasets consisting of a series of phage dilutions were taken at several different inhibitor concentrations and computed values of K_{I} from the multiple datasets were averaged.

3. Results

To validate the ELISA assay as a tool for estimating the relative affinity of a competing ligand, WEYIPNV-tagged phage was challenged with a heptapeptide of the same sequence (Fig. 1). As expected, the binding curve shifted to the right (toward higher phage concentration) with increasing concentration of peptide. To compare alternative functions for determining the $[\phi]_{0.5}$ of binding curves, a function of the form

$$A_{405} = \frac{A_{\max}}{1 + \left(\frac{K_{\text{app}}}{[\phi]}\right)^{\gamma}}$$

was used, where γ is an adjustable parameter. For binding of a set of ligands to identical sites, γ equals 1 for non-cooperative binding and differs from 1 for cooperative binding. It was found that in most cases, the function fit the data with a correlation coefficient $R > 0.99$ when $\gamma = 1$, where

$$R^2 = 1 - \frac{n\sum\left(y_{\text{i}} - f\left(x_{\text{i}}\right)\right)^2}{n\sum y_{\text{i}}^2 - \left(\sum y_{\text{i}}\right)^2}$$

The function in which γ was an adjustable parameter rather than a constant set equal to 1 fit only slightly better and the

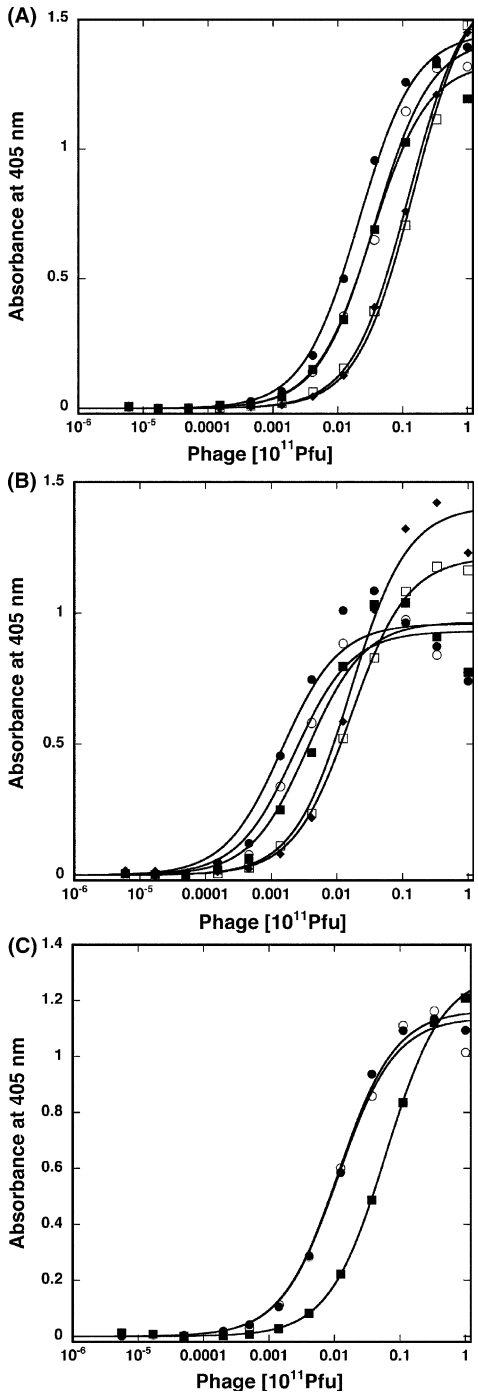

Fig. 2. Representative ELISA data for competition experiments between FTYMPPV-tagged phage and various ligands. Curve fits as described in Fig. 1. (A) Target protein, SurA. (●), average of two measurements with FTYMPPV-tagged phage in absence of competing ligand; competing ligands for other data: (○), 31.4 μM RCMLA; (■), 3.79 μM OmpF; (□), 2.44 μM OmpG; (◆), 3.4 μM peptide WEYIPNV. Note that the RCMLA and OmpF curves nearly overlap, indicating that the ratio of their K_{I} values is approximately equal to the ratio of their concentrations (31.4/3.79 = 8.3). Similarly with the OmpG and WEYIPNV curves. (B) Target protein, SurA(ΔP2). (●), average of two measurements with FTYMPPV-tagged phage in absence of competing ligand; competing ligands for other data: (○), 28.0 μM RCMLA; (■), 3.79 μM OmpF; (□), 2.44 μM OmpG; (◆), 3.4 μM peptide WEYIPNV. (C) Target protein, SurA(ΔP2). (●), FTYMPPV-tagged phage in absence of competing ligand; (○), 24.8 ΔM Hsc70 ATPase fragment; (■), 3.4 ΔM peptide WEYIPNV.

difference between the $[\phi]_{0.5}$ of the two alternative fitting procedures was generally small. Hence, the former function was used throughout for data analysis.

The limitations in accuracy inherent to the fitting of a hyperbolic function to the data were explored using some of the more poorly fitting data (Fig. 1B). At phage concentrations substantially in excess of tenfold higher than $[\phi]_{0.5}$, the ELISA assay signal saturates and decreases from its maximum value for unknown reasons. This consequently affects both of the adjustable parameters in the curve fit. In the example shown in Fig. 1B, when data for the two highest phage concentrations are omitted, the computed value for $[\gamma]_{0.5}$ shifts from 2.3 ± 0.8 μM ($R = 0.97$) using all data to 3.0 ± 0.9 μM ($R = 0.98$). Although these two values are equal within computed error, the second is 30% larger than the first. Hence, the assay is sufficiently robust to distinguish severalfold differences in affinities of two ligands measured under identical conditions, but is not precise enough to distinguish twofold differences. In this work, the method is used to parameterize binding affinities that differ by an order of magnitude.

Values of computed K_I at each concentration of WEYIPNV peptide are shown in Fig. 1 (inset). There is only modest variation over the range of peptide concentrations tested (5.6–56 μM). Averaging all data points gives an estimated value of $K_I = 0.67 \pm 0.08$ μM. This can be compared to the value 1.69 ± 0.15 μM measured by calorimetry at pH 7.3, 20 °C [6]. The difference of roughly twofold in the values may be due to the differences in assay conditions (for example, the ELISA assay uses 0.5% v/v Tween 20 detergent, while no detergent was present in the calorimetry experiment), the difference in the nature of the two assays, and the limitation on accuracy of absolute values of binding constants discussed above. In the work that follows, conclusions are based on the relative affinities of different ligands when compared to the WEYIPNV peptide as a reference standard, rather than on the absolute values of the apparent inhibition constants.

To monitor apparent binding inhibition by candidate SurA ligands, FTYMPPV-tagged phage was used. Representative data are shown in Fig. 2 and results are summarized in Table 1. The K_I values for the heptapeptide WEYIPNV are 0.5 ± 0.3 μM for SurA and 0.9 ± 0.5 μM for SurA(ΔP2), equal within experimental error to the value determined with the WEYIPNV-presenting phage. Binding of two unfolded OMPs solubilized in aqueous solution, *E. coli* OmpG and OmpF, was measured. The apparent affinity of OmpG was similar to that of WEYIPNV for binding to both SurA and SurA(ΔP2), while the binding of OmpF was severalfold weaker. Binding of RCMLA, which is often taken as a representative unfolded soluble protein [10], was measured and found to have an order of magnitude lower affinity for both SurA or SurA(ΔP2) than either of the OMPs. Further, two native soluble proteins were tested, the ATPase fragment of bovine Hsc70 and GST. Neither showed any significant affinity for SurA or SurA(ΔP2) (Fig. 2C and data not shown).

4. Discussion

Two different unfolded OMPs stabilized in aqueous solvent, OmpF and OmpG, have been used to characterize in interactions of OMPs with SurA and SurA(ΔP2). OmpF is a trimeric OMP [11] whose in vivo maturation is impaired in *surA*$^-$ strains [1]. OmpG is a momomeric OMP whose in vitro folding has been characterized extensively [9,12].

The data presented here establish several points. First, OmpF and OmpG both compete with FTYMPPV-tagged phage for binding to SurA and SurA(ΔP2), demonstrating that a peptide selected by phage display mimics OMP binding by competing for the same binding site.

Second, the relative affinities of unfolded OmpG and OmpF for SurA and SurA(ΔP2) are similar in magnitude to the affinity of the peptide WEYIPNV, arguing that the phage display experiment selected peptides that are representative of OMP affinity; the experiment did not select peptides with anomalously high or low affinity. Difference in affinity between OmpG and OmpF may be intrinsic to the proteins, or it may reflect differences in the method of solubilization (for example, OmpF preparations have residual SDS present, while OmpG preparations are devoid of detergent). It is notable in this context that among the peptides selected by phage display having the Ar–X–Ar motif, the sequence WEYIPNV had the highest affinity for SurA and SurA(ΔP2), and others spanned approximately an order of magnitude in affinity, as measured by the ELISA assay [6]. This illustrates an apparent variation in affinities of peptides carrying the Ar–X–Ar motif; a similar level of variation for natural substrates would not be unexpected.

Third, the order-of-magnitude weaker binding of RCMLA, as compared to OmpF and OmpG, demonstrates substrate selectivity for OMPs among unfolded polypeptides. This apparent selectivity is consonant with studies showing a

Table 1
Apparent binding constants derived from ELISA assay data using FTYMPPV-tagged phage

Competing ligand	Inhibitor concentrations at which ELISA datasets were collected (μM)	Average K_I (μM)	K_I/K_I(WEYIPNV)
For SurA protein			
WEYIPNV	0.15, 0.38, 0.70, 0.70	0.5 ± 0.3	1.0
OmpG	0.32, 0.42, 0.50, 0.50	0.44 ± 0.09	0.9
OmpF	2.9, 3.6, 4.8, 6.1, 6.7, 7.0	5.2 ± 1.7	10
RCMLA	16.0, 34.9, 39.1, 40.7	33 ± 11	66
For SurA(ΔP2) protein			
WEYIPNV	0.33, 0.50, 0.77, 0.80, 0.86, 1.6, 1.6	0.9 ± 0.5	1.0
OmpG	0.22, 0.24, 0.27, 0.30, 0.60, 1.0	0.4 ± 0.3	0.4
OmpF	1.9, 2.3, 2.6, 2.6, 2.7, 3.4	2.6 ± 0.5	2.9
RCMLA	28.3, 35.7, 50.7	38 ± 11	42

preference for membrane proteins in vivo, where in a *surA*$^-$ mutant strain, several OMPs, including OmpF, become preferentially sensitive to trypsin digestion (presumably reflecting impaired folding), while several soluble periplasmic proteins are unaffected (suggesting that participation of SurA in their folding is negligible) [1]. Notably, bovine lactalbumin does not have any Ar–X–Ar motifs in its amino acid sequence, while OmpF has eight such motifs and OmpG has 14. These data corroborate the suggestion that the Ar–X–Ar motif is a primary binding target of SurA.

Fourth, SurA and SurA(ΔP2) specifically interact with unfolded polypeptides; they show no measurable affinity for two native folded proteins that were tested.

Hence, further studies of the interaction of SurA and SurA(ΔP2) with short peptides derived from phage display experiments can be pursued with the assurance that their interactions mimic those natural substrates.

Acknowledgements: This work was supported by Grant GM-39928 from the National Institutes of Health to D.B.M.

References

[1] Lazar, S.W. and Kolter, R. (1996) J. Bacteriol. 178, 1770–1773.
[2] Missiakas, D., Betton, J.M. and Raina, S. (1996) Mol. Microbiol. 21, 871–884.
[3] Rouvière, P.E. and Gross, C.A. (1996) Genes Dev. 10, 3170–3182.
[4] Behrens, S., Maier, R., de Cock, H., Schmid, F.X. and Gross, C.A. (2001) EMBO J. 20, 285–294.
[5] Bitto, E. and McKay, D.B. (2002) Structure 10, 1489–1498.
[6] Bitto, E. and McKay, D.B. (2003) J. Biol. Chem. 278, 49316–49322.
[7] Wilbanks, S.M., DeLuca-Flaherty, C. and McKay, D.B. (1994) J. Biol. Chem. 269, 12893–12898.
[8] Blattner, F.R. et al. (1997) Science 277, 1453–1474.
[9] Conlan, S., Zhang, Y., Cheley, S. and Bayley, H. (2000) Biochemistry 39, 11845–11854.
[10] Palleros, D.R., Shi, L., Reid, K.L. and Fink, A.L. (1994) J. Biol. Chem. 269, 13107–13114.
[11] Cowan, S.W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R.A., Jansonius, J.N. and Rosenbusch, J.P. (1992) Nature 358, 727–733.
[12] Behlau, M., Mills, D.J., Quader, H., Kuhlbrandt, W. and Vonck, J. (2001) J. Mol. Biol. 305, 71–77.