doi:10.1016/j.jmb.2008.03.048 J. Mol. Biol. (2008) 379, 457–470

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Phage Display and Crystallographic Analysis Reveals Potential Substrate/Binding Site Interactions in the Protein Secretion Chaperone CsaA from *Agrobacterium tumefaciens*

Anat R. Feldman[1]†, Yuliya A. Shapova[1]†, Sampson S. T. Wu[1], David C. Oliver[1], Markus Heller[2], Lawrence P. McIntosh[2], Jamie K. Scott[1,3] and Mark Paetzel[1]*

[1]*Department of Molecular Biology and Biochemistry, Simon Fraser University, South Science Building, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6*

[2]*Department of Biochemistry and Molecular Biology, Department of Chemistry, The Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3*

[3]*Faculty of Health Sciences, Simon Fraser University, South Science Building, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6*

Received 14 December 2007; received in revised form 15 March 2008; accepted 20 March 2008
Available online 28 March 2008

Edited by R. Huber

The protein CsaA has been proposed to function as a protein secretion chaperone in bacteria that lack the Sec-dependent protein-targeting chaperone SecB. CsaA is a homodimer with two putative substrate-binding pockets, one in each monomer. To test the hypothesis that these cavities are indeed substrate-binding sites able to interact with other polypeptide chains, we selected a peptide that bound to CsaA from a random peptide library displayed on phage. Presented here is the structure of CsaA from *Agrobacterium tumefaciens* (AtCsaA) solved in the presence and absence of the selected peptide. To promote co-crystallization, the sequence for this peptide was genetically fused to the amino-terminus of AtCsaA. The resulting 1.65 Å resolution crystal structure reveals that the tethered peptide from one AtCsaA molecule binds to the proposed substrate-binding pocket of a symmetry-related molecule possibly mimicking the interaction between a pre-protein substrate and CsaA. The structure shows that the peptide lies in an extended conformation with alanine, proline and glutamine side chains pointing into the binding pocket. The peptide interacts with the atoms of the AtCsaA-binding pocket via seven direct hydrogen bonds. The side chain of a conserved pocket residue, Arg76, has an "up" conformation when the CsaA-binding site is empty and a "down" conformation when the CsaA-binding site is occupied, suggesting that this residue may function to stabilize the peptide in the binding cavity. The presented aggregation assays, phage-display analysis and structural analysis are consistent with AtCsaA being a general chaperone. The properties of the proposed CsaA-binding pocket/peptide interactions are compared to those from other structurally characterized molecular chaperones.

Keywords: molecular chaperone; protein secretion; protein targeting; phage display; CsaA

Introduction

Bacterial proteins targeted for secretion are synthesized in the cytosol where they associate with chaperones that prevent their aggregation and keep them in a translocation-competent form.[1]

In Gram-negative bacteria, newly synthesized secretory proteins interact in the cytoplasm with the chaperone SecB.[2] At the membrane, SecB interacts with the ATPase SecA that catalyzes (along with the

**Corresponding author.* E-mail address: mpaetzel@sfu.ca.

† A.R.F. and Y.A.S. contributed equally to this work.

Abbreviations used: AtCsaA, *Agrobacterim tumefaciens* CsaA; X15peptide–AtCsaA, *A. tumefaciens* CsaA with an amino-terminal peptide fusion derived from phage-display analysis; AtCsaA-biotin, AtCsaA modified with biotin; HRP, horseradish peroxidase; TtCsaA, *Thermus thermophilus* CsaA; BsCsaA, *Bacillus subtilis* CsaA.

proton-motive force across the inner membrane) the translocation of pre-proteins through the SecYEG channel.[3–6] Shortly after translocation, the leader sequence is cleaved off by signal peptidase,[7,8] and mature proteins are released from the membrane into the periplasm. The Sec-dependent protein secretion system is an ancient system,[9] and its components have been universally found in eubacteria, eukaryotes and archaea.[9,10]

Gram-positive bacteria, such as *Bacillus subtilis*, lack a gene for SecB.[11,12] It has been proposed that in *B. subtilis*, the protein CsaA[13] has a function analogous to that of the SecB protein in *Escherichia coli*.[13,14] Although CsaA and SecB share no sequence identity or structural similarity,[15,16] it has been suggested that they have an overlapping functionality. The *B. subtilis* CsaA is able to bind *E. coli* SecA *in vitro* and *in vivo*,[17] and it is able to promote translocation of the pre-protein prePhoB into *E. coli* membrane vesicles supplemented with the *B. subtilis* translocase system.[13] CsaA stimulates protein export in *secB* mutant strains of *E. coli*.[17] CsaA has higher affinity for denatured proteins than for native ones, and the interactions are independent of the signal peptide sequence.[14] A peptide scan using 13-mer peptides based on the sequence of preYvaY, a CsaA substrate, was performed to determine the binding preference of CsaA.[14] Although no consensus sequence was found, CsaA showed preference for peptides with positively charged termini.[14]

The structures of CsaA from *Thermus thermophilus* (TtCsaA)[16] and *B. subtilis* (BsCsaA)[18] revealed that both TtCsaA and BsCsaA are dimeric proteins with each monomer having an oligonucleotide/oligosaccharide binding (OB) fold and having structural similarity to the tRNA-binding protein trbp111.[19]

In contrast to *B. subtilis* and *E. coli*, *Agrobacterium tumefaciens* has homologus genes for both *csaA* and *secB*. *A. tumefaciens* is a Gram-negative plant pathogen that causes crown gall tumors (abnormal tissue growth) in a broad range of plants.[20] Its genome has been sequenced[21,22] and it belongs to the α-subdivision of the *Rhizobiaceae* family.[23] It has been widely used as a tool in plant molecular biology, since it has a complex system for integrating foreign DNA into plant cells.

In order to investigate the putative substrate-binding pocket of CsaA and how CsaA might recognize its substrates, we have selected a peptide from a random library displayed on phage and co-crystallized the selected peptide with *A. tumefaciens* CsaA (AtCsaA) by engineering the AtCsaA gene such that the peptide was fused to the N-terminus of the AtCsaA protein. We present here a detailed analysis of the peptide/pocket interactions in the co-crystal structure and compare these binding pockets to those observed in the *A. tumefaciens* CsaA without peptide, as well as the structures from *B. subtilis* and *T. thermophilus* CsaA. We have analyzed AtCsaA's chaperone activity, using aggregation assays, with and without the fused phage-display peptide. We also compare the binding pocket/peptide interactions present in CsaA to those observed in other structurally characterized molecular chaperones.

Results

The structure of AtCsaA

The structure of CsaA from *A. tumefaciens* was solved to 1.55 Å resolution in space group $P6_1$ with two molecules in the asymmetric unit (Fig. 1). The final refined structure (R_{work} = 17.8%, R_{free} = 20.8%) includes all CsaA residues except for residues 29–31 in chain B and the N-terminal methionine in each chain. The residue Glu28 in chain B was modelled as alanine due to lack of side-chain density (Table 2).

The AtCsaA dimer, surface pockets and comparison to BsCsaA and TtCsaA

The structure of CsaA from *A. tumefaciens* is very similar to those of CsaA from *B. subtilis* (BsCsaA) and *T. thermophilus* (TtCsaA). It is a homodimeric molecule, in which each monomer is 113 amino acids long and consists of 2 α-helices and 10 β-strands (Fig. 1). Strands β1, β2, β3, β4, β7, and helix α2 form the core of the monomer, adopting an OB-fold. The N-terminal helix α1 and the C-terminal strands β8 and β9 contain the majority of residues that participate in the formation of interchain hydrogen bonds important for dimerization. The r.m.s.d. of superposition for the AtCsaA dimer and BsCsaA dimer (2NZO, chains AB) is 1.68 Å over 218 α-carbon atoms, and that of AtCsaA and TtCsaA (1GD7, chains AB) is 1.77 Å over 216 α-carbon atoms. The three structures superimpose very well except for several flexible loop regions (Fig. 1b). Similar to BsCsaA and TtCsaA, AtCsaA contains two large cavities (one in each monomer) that are separated by a rotation of approximately 90° about the long axis of the dimer (Fig. 1C). These cavities are putative substrate-binding sites and are composed of residues 26, 28–31, 33, 49–53, 73–76, 78, 83–90, 97 from one monomer, and residues 4′ and 113′ from the adjacent monomer. The majority of these residues are located on the flexible loops described above, indicating that the binding site may be dynamic. The residues that make up the putative binding site are overall well conserved in CsaA proteins from different species (Fig. 2). It is interesting to note that the pattern of strong positive and negative electrostatic potential in the vicinity of the putative binding site in BsCsaA and TtCsaA also occurs in AtCsaA (Fig. 1c). Due to high sequence variability in that region (Fig. 2), the precise position of the negatively charged area is somewhat different in all three proteins. However, the overall tendency of having both negatively and positively charged regions in the vicinity of the binding cavity is preserved.

The AtCsaA dimer is bridged by 18 hydrogen bonds, 14 of which are formed by the same residues in each monomer due to its 2-fold symmetry. The

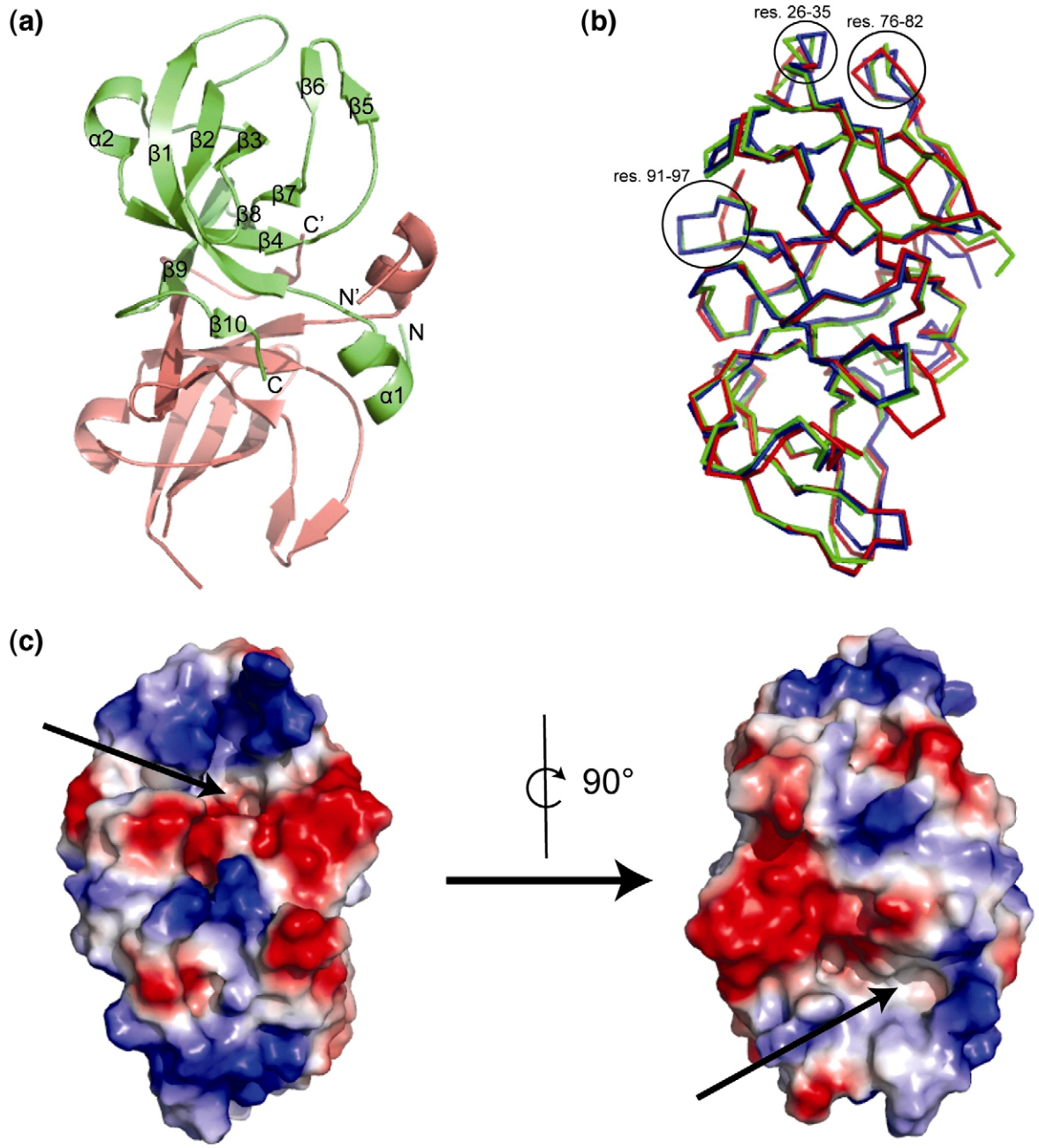

Fig. 1. The structure of AtCsaA. (a) A ribbon diagram of dimeric AtCsaA, with molecule A in green and molecule B in pink. (b) A C^{α}-trace diagram of the superimposed structures of AtCsaA (green), BsCsaA (2NZH, red), and TtCsaA (1GD7, blue). Circles represent areas of the CsaA molecule with high thermal motion. (c) A surface representation of the AtCsaA structure colored according to the negative (red), positive (blue), or neutral (white) electrostatic potential. The location of the putative binding site is indicated with arrows.

pattern of interchain hydrogen bonding is well conserved between the structures of AtCsaA and BsCsaA. The majority of main-chain hydrogen bonds occur between residues that occupy equivalent positions in the sequences of AtCsaA and BsCsaA. Two side-chain-mediated hydrogen bonds are also conserved in the structures of AtCsaA and BsCsaA and occur between the donor–acceptor pairs Tyr 57–Glu103′ and Lys12–Asp9′ (AtCsaA numbering). These residues may therefore be important for dimerization of the CsaA protein.

Selection of a peptide ligand for AtCsaA

We searched for an AtCsaA ligand by screening against a linear peptide library displayed at the amino-terminus of phage coat protein pVIII. Typically, recombinant pVIII molecules bearing peptide compose 0.5–10% of the total coat protein.[24] Three rounds of panning were performed, and the affinity of the pool selected in each cycle was analyzed by ELISA (data not shown).[25] Individual clones were amplified from these pools and showed significant

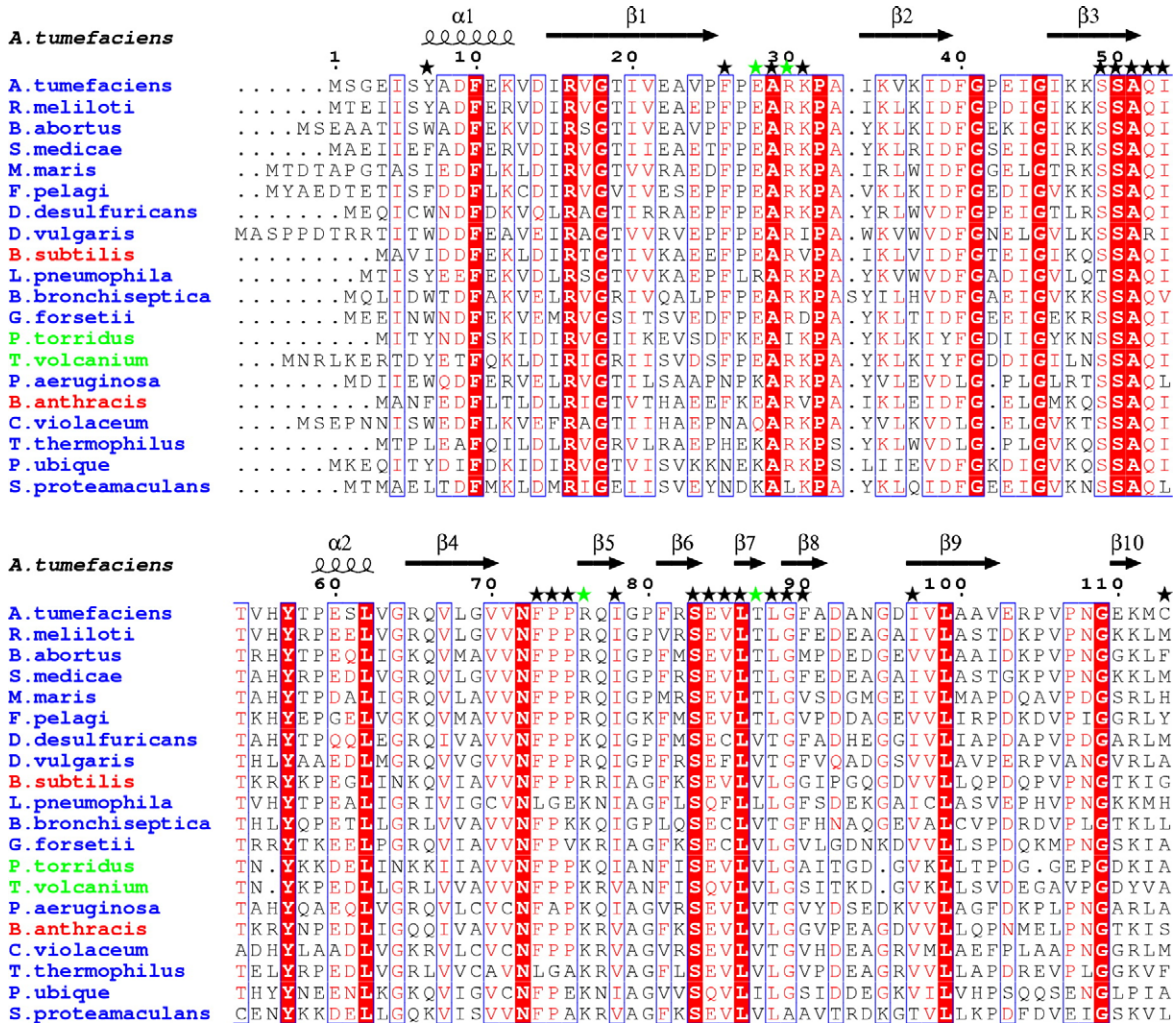

Fig. 2. Sequence alignment of AtCsaA with other annotated CsaA proteins. The secondary structural elements for AtCsaA are shown above the alignment, along with corresponding numbering for AtCsaA. Residues contributing to the binding pockets are labeled with black stars. Residues involved in hydrogen-bonding interactions with the peptide have green stars. The CsaA proteins from Gram-negative bacteria are named in blue, Gram-positive in red, and Archaea in green. Conserved residues are highlighted in red. The sequences are ordered from most conserved to least conserved. Accession numbers for the Swiss-Prot CsaA sequence from each species are as follows: Q8UDB9, *A. tumefaciens* (strain C58/ATCC 33970); Q92N79, *Rhizobium meliloti* (*Sinorhizobium meliloti*); Q57BZ2, *Brucella abortus*; Q0MGG5, *Sinorhizobium medicae*; Q0ARE7, *Maricaulis maris* (strain MCS10); Q0G083, *Fulvimarina pelagi* HTCC2506; Q30UZ8, *Desulfovibrio desulfuricans* (strain G20); Q0EN66, *Desulfovibrio vulgaris* subsp. *vulgaris* DP4; P37584, *B. subtilis*; Q5ZSH0, *Legionella pneumophila* subsp. *pneumophila* (strain Philadelphia 1/ATCC 33152/DSM 7513); Q7WHT9, *Bordetella bronchiseptica* (*Alcaligenes bronchisepticus*); A0M4E5, *Gramella forsetii* KT0803; Q6L208, *Picrophilus torridus*; Q97BA7, *Thermoplasma volcanium*; Q9HZ19, *Pseudomonas aeruginosa*; Q81RI2, *Bacillus anthracis*; Q7NUP4, *Chromobacterium violaceum*; Q9AQH8, *T. thermophilus*; Q4FL08, *Pelagibacter ubique*; A0IMQ4, *Serratia proteamaculans* 568.

binding in ELISA, above background binding for F88 phage. Twenty-three clones from the linear library selection were sequenced from the third round of selection. All clones had the following sequence at their randomized positions, (NH_2)-VPGQKQHYVQP-TAAN, and were named X15.1. Binding of the X15.1 phage clone as well as some clones selected from a constrained library (named LX10) to AtCsaA are reported in Table 1. A sequence comparison among selected phage clones from the two libraries revealed no consensus sequence, consistent with CsaA's ability to function as a general chaperone. Moreover, selected peptides on phage showed different affinities to CsaA. The clone with the highest affinity was the dominant clone X15.1, yielding a signal 25-fold above background.

Crystallization of AtCsaA with bound peptide

A peptide based on the selected X15.1 sequence was designed and chemically synthesized. NMR spectroscopic analysis confirmed that this peptide bound AtCsaA *in vitro* (Supplementary Fig. S1). Co-crystallization efforts also resulted in identifiable electron density in the proposed AtCsaA substrate-binding pocket. Unfortunately, attempts to build and refine a peptide into the electron density failed due to low occupancy and the lack of clear side-

Table 1. ELISA on individual phage clones selected from X15 and LX10 libraries

Clone ID	Sequence of displayed peptide	AtCsaA-Bio[a]	Anti-phage antibody[b]
X15.1	(NH_2)-VPGQKQHYVQPTAAN	0.51	1.11
LX10.H1[c]	(NH_2)-QCGIKVATQTQRCQ	0.41	1.12
LX10.D4[c]	(NH_2)-MCRRKPEPLLPNCT	0.22	1.18
LX10.C4[c]	(NH_2)-TCPPVDRRHVAACS	0.26	1.21
LX10.C6[c]	(NH_2)-ACMKRSPNEPHHCW	0.25	1.22
F88[d]		0.02	1.02

Reported values correspond to the absorbance at 405 nm minus the absorbance at 490 nm.

[a] Phage binding was detected with a chromogenic reaction, with AtCsaA-biotin as primary binder and neutravidin-conjugated HRP as secondary binder.

[b] A control experiment was performed to confirm that the number of phage particles was the same for individual clones in the experiment. Phage binding was detected with a chromogenic reaction, with anti-phage antibody as primary binder, and neutravidin-conjugated HRP as secondary binder.

[c] Underlined cysteines are conserved in the loop library.

[d] Phage without displayed peptide.

chain density that made determining the orientation of the peptide impossible (data not shown).

Thus, we cloned the peptide at the N-terminus of AtCsaA, rationalizing that this should improve the occupancy and electron density for the peptide in AtCsaA co-crystals. The design of this experiment was based on the fact that during biopanning the peptide was expressed at the N-terminus of the pVIII phage protein. Therefore, tethering the peptide at the N-terminus of AtCsaA could mimic the peptide-pVIII. In addition, from analysis of the AtCsaA structure (without peptide), we found that the N-terminus of AtCsaA is exposed to solvent. The resulting construct, named X15peptide–AtCsaA, begins with the sequence NH_2-$G_{-19}S_{-18}H_{-17}M_{-16}$ from the expression vector, followed by the X15.1 peptide $V_{-15}P_{-14}G_{-13}Q_{-12}K_{-11}Q_{-10}H_{-9}Y_{-8}V_{-7}Q_{-6}P_{-5}T_{-4}A_{-3}A_{-2}N_{-1}$, fused to Ser_2 of AtCsaA. MALDI-TOF mass spectrometric analysis confirmed that thrombin treatment of the initially His_6-tagged X15peptide–AtCsaA resulted in a monomer with a molecular mass of 14,116 Da (predicted, 14,123 Da).

Interestingly, X15peptide–AtCsaA crystals formed in a different condition from that of the AtCsaA alone, yet retained the same space group ($P6_1$) and approximate unit cell dimensions (Table 2). A molecular replacement solution was found using the structure of AtCsaA as a search model. Clear $F_o - F_c$ difference electron density was obtained for the peptide at the N-terminus of molecule A for residues NH_2-$Q_{-6}P_{-5}T_{-4}A_{-3}A_{-2}N_{-1}$ (Fig. 3). No difference density was observed for the peptide at the N-terminus of molecule B. The first residue to be resolved in the electron density for molecule B is Gly3. Electron density was also missing for a loop region in molecule B (residues 27–31).

Structure of X15peptide–AtCsaA

Overall, the structure of X15peptide–AtCsaA is very similar to that of free AtCsaA, superimposing with an r.m.s.d. of 0.76 Å over 219 α-carbon atoms (Supplemental Fig. S2). The structure of X15peptide–AtCsaA reveals that the peptide, at the N-terminus of molecule A in one dimer, binds to the pocket on molecule A in a symmetry-related dimer (Fig. 4). The peptide binds in an extended conformation, burying 344 $Å^2$ of the AtCsaA surface at the peptide–protein interface. This surface is ~60% non-polar in nature and consists of residues 26, 28–31, 33, 50–52, 76, 78–80, 86 and 87 from chain A and residue 113 from chain B, which contact the peptide through both polar and non-polar interactions. There are seven hydrogen bonds between the atoms of the peptide

Table 2. Data collection and refinement statistics

	AtCsaA	X15peptide–AtCsaA
Data collection		
Crystallization conditions	1.8 M ammonium sulfate, 0.1 M Hepes (pH 7.5), 2% PEG 400, and 5% ethylene glycol	30% PEG 4000, 0.4 M NH_4OAc, 0.1 M sodium citrate (pH 5.5)
Molecular mass (of dimer) (Da)	24,893	28,246
Space group	$P6_1$	$P6_1$
a, b, c (Å)	$a=60.6, b=60.6, c=113.4$	$a=60.5, b=60.5, c=115.3$
Molecules in assymmetric unit	2	2
V_m ($Å^3$/Da)	2.5	2.2
Resolution (Å)	52.53–1.55 (1.61–1.55)	23.86–1.65 (1.71–1.65)
Total observed reflections	369,154	146,729
Unique reflections	34,123	25,681
% Completeness	99.8 (99.6)	89.4 (49.7)
$I/\sigma(I)$	23.7 (7.2)	30.0 (8.6)
R_{merge} (%)	4.0 (34.1)	3.9 (16.2)
Redundancy	10.8 (10.1)	5.7 (4.5)
Refinement		
Protein residues	221	224
Water molecules	350	307
Other solvent molecules	6	3
R_{work} (%)	17.8	16.1
R_{free} (%)	20.8	17.3
r.m.s.d.'s		
Bonds (Å)	0.010	0.008
Angles (°)	1.282	1.183
Overall B ($Å^2$)		
All atoms	7.9	8.2
Protein[a]	4.1/6.1	4.5/7.6
Peptide	–	7.0
Solvent	20.6	19.9
Ramachandran (%)[b]	92.0/8.0	91.7/8.3

$R_{merge} = \sum | I - \langle I \rangle | / \sum I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections after rejections.

$R = \sum || F_o | - | F_c || / \sum | F_o |$, where F_o and F_c are the observed and calculated structure factors, respectively.

R_{free} = defined in Brünger *et al.*[26]

The data collection statistics in parentheses are the values for the highest resolution shell (1.61–1.55 Å) for CsaA crystals and (1.71–1.65 Å) for X15peptide–AtCsaA crystals.

[a] Protein B-factor for chain A/chain B.

[b] Residues in the most favorable region/additionally allowed region.

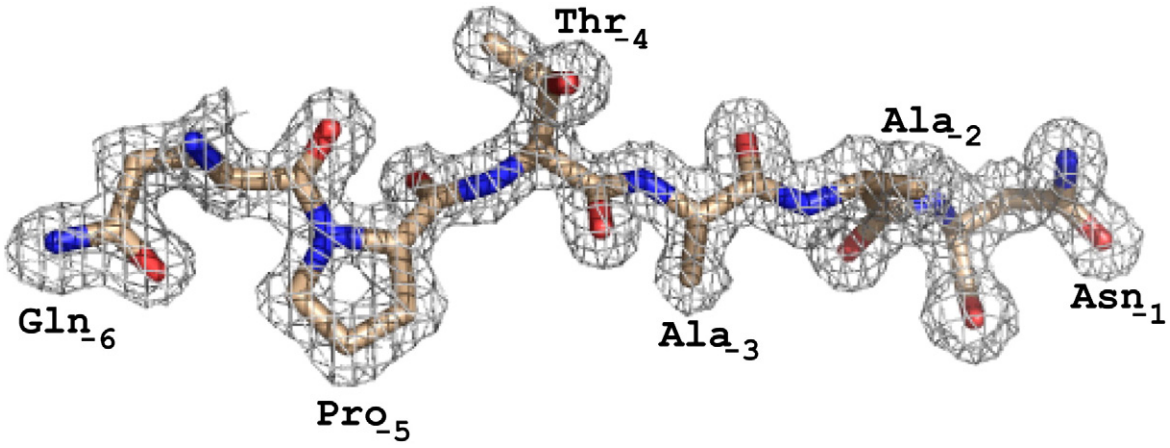

Fig. 3. Structure of the bound peptide from the X15peptide–AtCsaA complex. An electron density omit map calculated for the tethered peptide in the CsaA-binding pocket, contoured at 1 sigma.

and the pocket (Table 3). Two of these hydrogen bonds are strictly main-chain interactions and the other five are side-chain mediated. In addition to the bonds listed in Table 3, there is an indirect hydrogen bond between the peptide and the pocket that occurs via a water molecule. The water W24 is coordinated in its position via hydrogen bonds to the peptide atom Gln(−6) $O^{\varepsilon 1}$ and the pocket atoms Thr87A N, Ser50A O, and Ser49A O^{γ}.

All pocket residues that participate in direct hydrogen-bonding interactions with the peptide do so through main-chain atoms, with the exception of Arg76A. The guanidinium group of this residue forms a bifurcated hydrogen bond, donating hydrogen atoms from both its $N^{\eta 1}$ and its $N^{\eta 2}$ to the same main-chain carbonyl oxygen on the peptide residue Pro(−5). The proposed AtCsaA substrate-binding pocket can fit four peptide residues in an extended conformation, since Asn(−1) and Ala(−2) do not enter the pocket. The side chains of three of these residues [Gln(−6), Pro(−5), and Ala(−3)] point down towards the pocket and possibly act as specificity determinants for the peptide. Gln(−6) is well coordinated in its position via one indirect and three direct hydrogen bonds from its side-chain atoms $N^{\varepsilon 1}$ and $O^{\varepsilon 1}$ to the pocket atoms. The Pro(−5) and Ala(−3) side chains interact with the pocket atoms through non-polar contacts and are constricted in their positions by main-chain hydrogen bonds to AtCsaA. Having large residues in place of Pro(−5) and Ala(−3) in the peptide would be disfavoured due to steric clash with the pocket atoms. Charged residues would also likely be disfavoured due to the hydrophobic character of the pocket at the sites of Pro(−5) and Ala(−3) binding.

The general chaperone activity of AtCsaA and X15peptide–AtCsaA

To compare the general chaperone activity of AtCsaA and X15peptide–AtCsaA we used an aggregation assay where the protein is added to denatured lysozyme. The time course of lysozyme aggregation was monitored by light scattering at 360 nm. Aggregation is slowed by AtCsaA, confirming its general chaperone activity (Fig. 5). In contrast, X15peptide–AtCsaA is less able to prevent aggregation, indicating that the peptide competes with lysozyme for the binding site of this chaperone (Fig. 5).

Discussion

Comparison of the putative substrate-binding pockets in the structures of CsaA from *A. tumefaciens*, *B. subtilis* and *T. thermophilus*

It is possible that the position of the pocket residue Arg76 is a determinant of the ability of AtCsaA to bind peptide ligands. In molecule A of X15peptide–AtCsaA, which has the peptide bound, the guanidinium group of the Arg76 side chain points towards the pocket, donating two hydrogen bonds to the carbonyl oxygen of peptide residue Pro(−5). In molecule B, which lacks the bound peptide, Arg76 points away from the pocket, towards the solvent. Interestingly, in the structure of free AtCsaA, which was crystallized without a peptide, Arg76 also points towards the pocket in molecule A and towards the solvent in molecule B (Fig. 6). A close examination of the pocket in molecule A of free AtCsaA reveals the presence of a sulfate ion occupying the same position as the peptide residue Pro(−5) in X15peptide–AtCsaA and also accepting hydrogen bonds from Arg76. In addition, the sulfate ion is highly coordinated in its position via direct and indirect hydrogen bonds to other residues of the pocket. Both of the structures, with and without the peptide, were crystallized in the space group $P6_1$ and with similar unit cell dimensions. In this particular crystal packing network, steric hindrance from the symmetry-related molecules also restricts the thermal motion of Arg76 from molecule A. The binding of either sulfate or peptide may stabilize Arg76 in the "down" position with a concomitant narrowing of the binding pocket. Since Arg76 sits on a solvent-exposed loop, it is likely that, *in vivo*, its side chain has a much

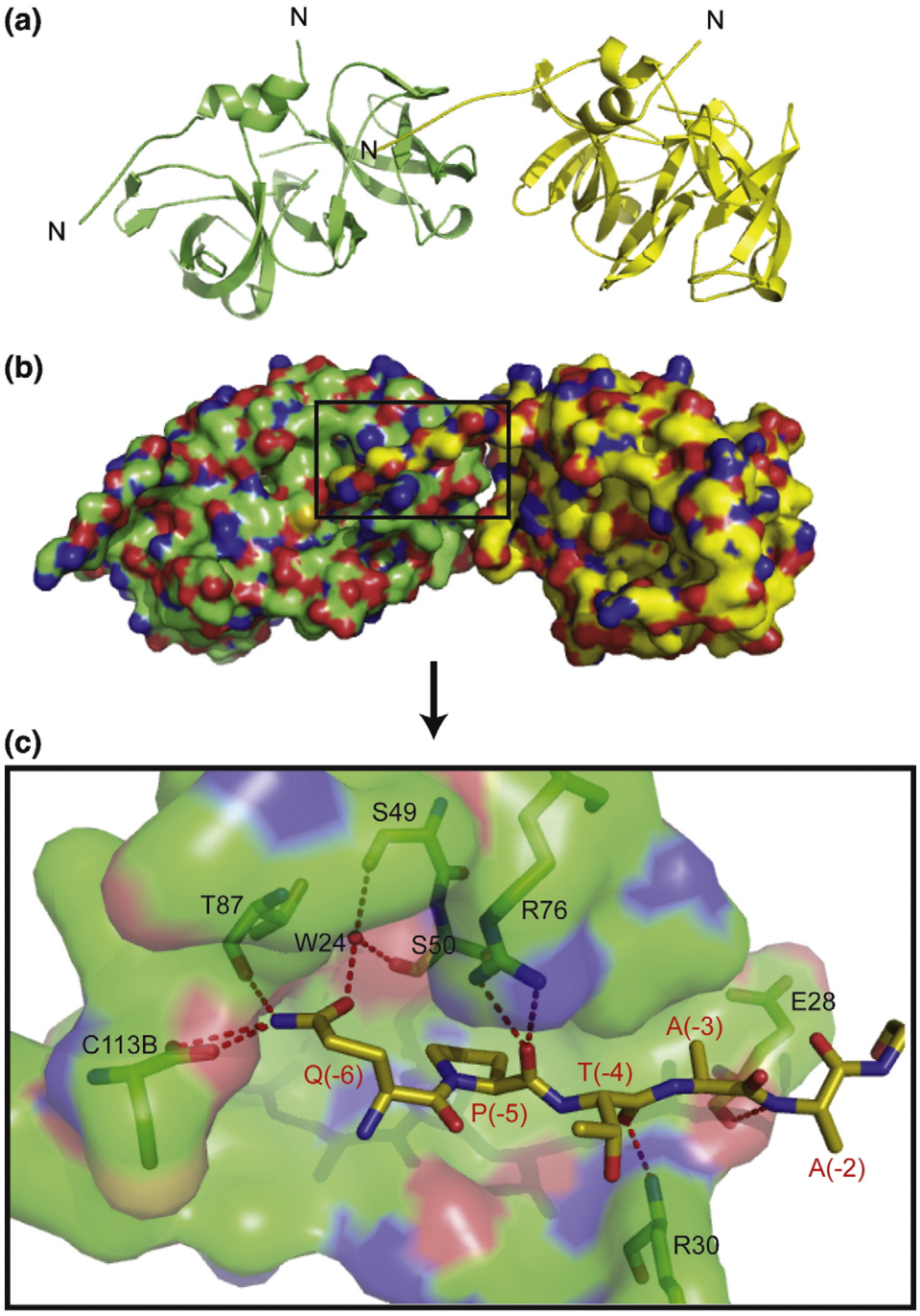

Fig. 4. The structure of AtCsaA in complex with a phage-display-derived peptide (X15peptide–AtCsaA). (a) A ribbon representation of the X15peptide–AtCsaA structure. The phage-display-derived peptide at the N-terminus of molecule A in one dimer (yellow) binds into the pocket of a molecule A from a symmetry-related dimer (green). (b) A surface representation of the molecules depicted in (a). Nitrogen atoms are in blue, oxygen atoms in red and carbon atoms in green or yellow. (c) A close-up view of the region enclosed in the box in (b), showing the interactions between the pocket and the peptide. Peptide residues are shown as yellow sticks (with red labels), and protein residues that make contact with the peptide as green sticks (with black labels). Hydrogen bonds are shown as red dashed lines.

greater degree of freedom of movement, allowing it to move "up" and "down" and to transiently interact with and stabilize the peptide bound into the pocket of CsaA. Indeed, in the previously published structures of CsaA from *B. subtilis* and *T. thermophilus*, an arginine and a lysine residue, respectively, occupy positions equivalent to Arg76 in AtCsaA. In the BsCsaA and TtCsaA structures, these residues lack a complete side-chain density and are also likely flexible due to a greater degree of thermal motion. Moreover, all CsaA proteins contain either a lysine or an arginine at this position (Fig. 2). Together, this suggests that the Arg76 side chain (or an equivalent lysine) may function as a molecular clamp that

Table 3. Direct hydrogen bonds between the peptide and AtCsaA

Peptide residue	Atom	X15peptide–AtCsaA residue	Atom	Bond length (Å)
Gln(−6)	Nε2	Thr87A	O	3.1
Gln(−6)	Nε2	Cys113B	OXT	2.7
Gln(−6)	Nε2	Cys113B	O	3.4
Pro(−5)	O	Arg76A	Nη1	3.0
Pro(−5)	O	Arg76A	Nη2	3.2
Thr(−4)	O	Arg30A	N	2.8
Ala(−2)	N	Glu28A	O	2.5

moves down to lock and stabilize the peptide substrate when it is bound into the substrate-binding pocket of CsaA.

The peptide atom Gln(−6) $N^{\varepsilon 1}$ makes hydrogen bonds to three main-chain atoms in the pocket (Thr87A O, Cys113B OXT, and Cys113B O) that together make up a patch of polar surface along the peptide-binding cavity. Comparison of the substrate-binding sites of CsaA structures from *A. tumefaciens*, *B. subtilis* and *T. thermophilus* reveals that the position of this patch of polar surface is also conserved among all three structures (Fig. 6). This conserved patch of polar surface in the pocket may act as another determinant of the chaperone–peptide interaction by providing a possibility for hydrogen bonds between the pocket and proton-donor atoms in the substrate.

Kawaguchi *et al*. suggested that the N-terminal helices in the structure of TtCsaA may be flexible, based on the poor electron density, high *B*-factors and lack of significant interactions between this region and the rest of the protein. Accordingly, it was proposed that the N-terminal helices may adopt an "open" conformation leading to a formation of a single hydrophobic groove unifying the two substrate-binding cavities.[16] However, in the structures of AtCsaA with and without the peptide, the average *B*-factors (8.5 and 6.3 $Å^2$) for the N-terminal α-helices (residues 7–12) are comparable to that of the rest of the protein (see Table 2). In addition, these regions participate in inter- and intrachain hydrogen bonding and van der Waals interactions with the rest of the protein. The N-terminal α-helices also appear to be a stable part of the dimeric AtCsaA structure and thus it is unlikely that they could adopt an open conformation. Based on the overall stability of the AtCsaA dimer, we propose that the substrate wraps around the surface of the AtCsaA particle in going from one binding site to the other. The fact that the peptide forms limited interactions with the CsaA peptide-binding groove and is difficult to co-crystallize with CsaA might also indicate that other parts of the substrate are responsible for providing additio-

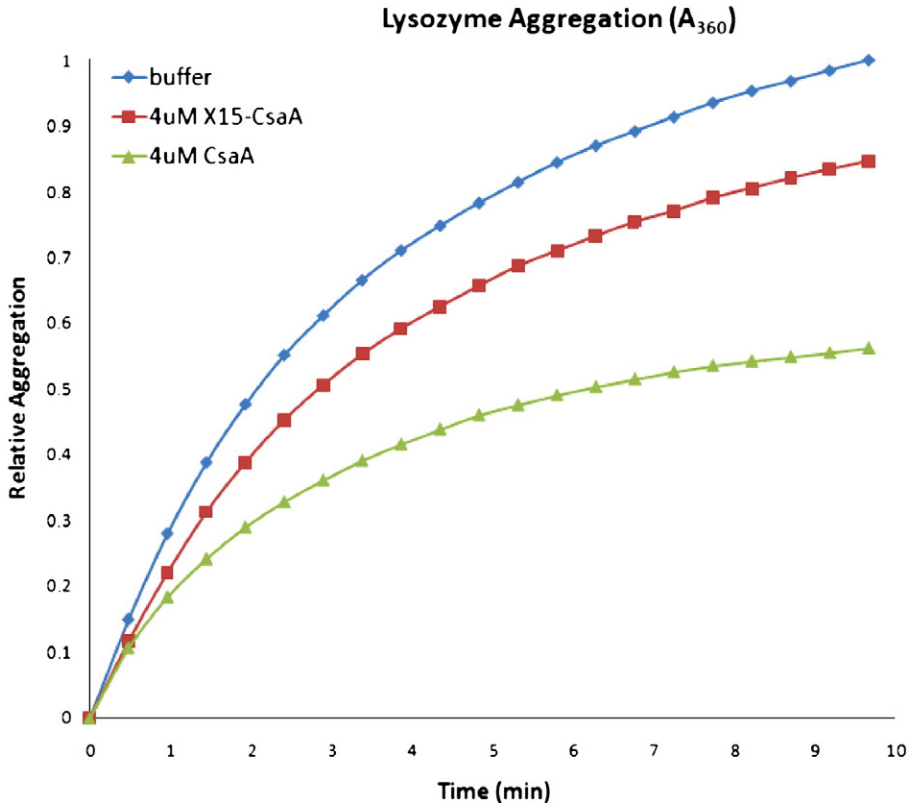

Fig. 5. The general chaperone activity of AtCsaA towards lysozyme is reduced by the presence X15peptide. A time course of denatured lysozyme aggregation in the presence or absence of AtCsaA or X15peptide–AtCsaA as monitored by turbidity at 360 nm. Each sample was run in triplicate and repeated in separate occasions. The data presented are representative trials. See Experimental Procedures for details.

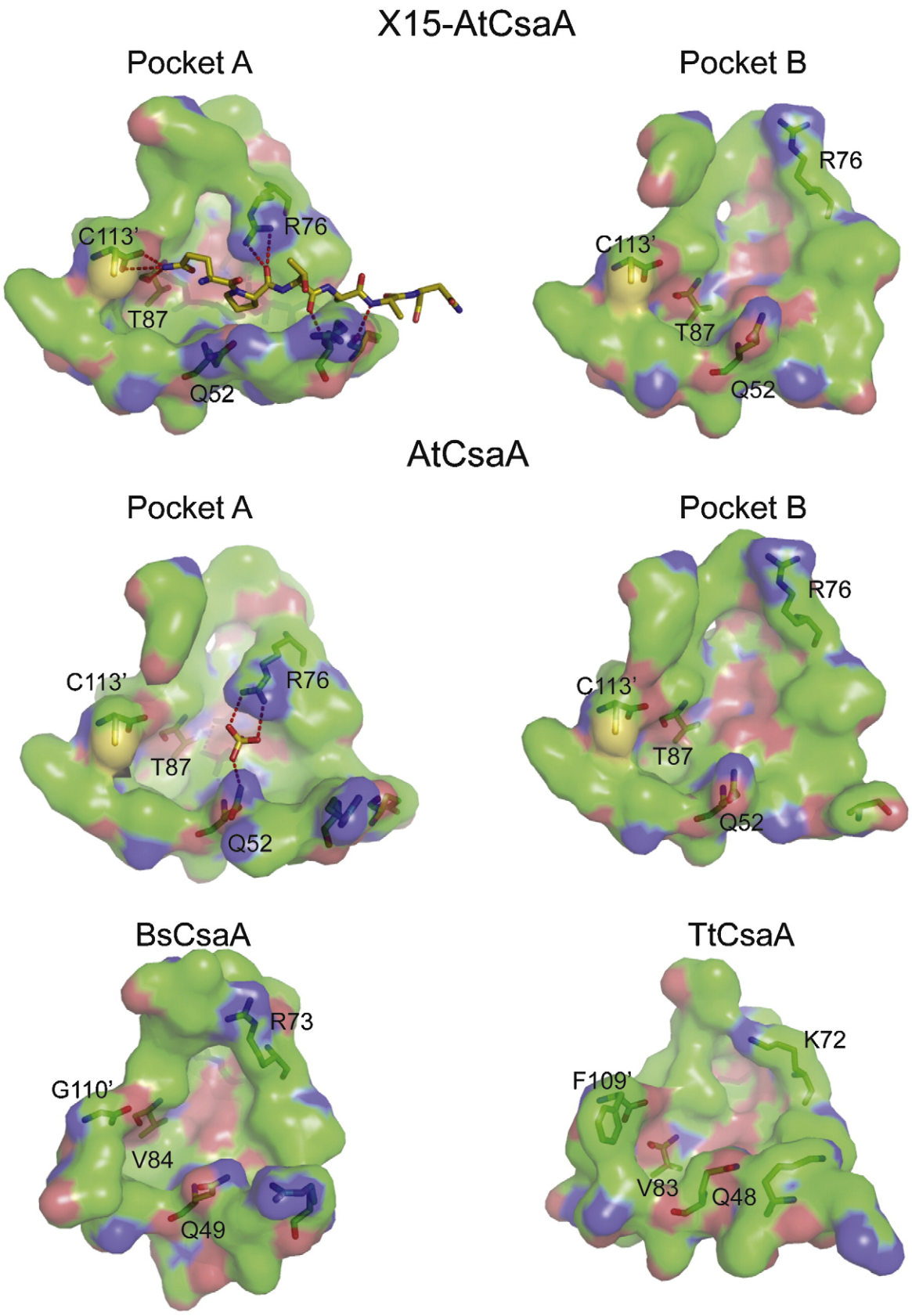

Fig. 6. A comparison of the putative substrate-binding pockets of AtCsaA, X15peptide–AtCsaA, BsCsaA (PDB ID 2NZO_AB), and TtCsaA (1GD7_AB). The binding pocket surfaces are colored (red, oxygen; blue, nitrogen; yellow, sulfur; green, carbon), and the residues important for protein–peptide interactions are shown as sticks (red, oxygen; blue, nitrogen; yellow, sulfur; green, carbon). In the absence of the X15peptide, Arg76, in molecule A of the AtCsaA dimer, binds a sulfate ion.

nal surface contacts with CsaA by wrapping around the chaperone.

AtCsaA is a general chaperone

Our aggregation assays with denatured lysozyme as a substrate reveal that AtCsaA is a general chaperone. This is consistent with a previous analysis exploiting a luciferase aggregation study to show that BsCsaA has general chaperone activity.[17] Furthermore, the X15peptide–AtCsaA showed significantly less general chaperone activity than the wild-type AtCsaA. This most likely results from the fused X15peptide occluding a portion of the AtCsaA substrate-binding pockets. More importantly, this indicates that the phage-display selection and the interfacial features uncovered in the resulting X15peptide–AtCsaA crystal structure indeed reflect the key features of the chaperone activity of CsaA.

Comparison of the AtCsaA–peptide interactions to those observed in other chaperone–peptide complexes

Other chaperones that recognize a broad spectrum of substrates and bind short motifs of sequence with few specificity determinants have been previously characterized (a comparison is presented in Table 4). For example, the chaperone DnaK was co-crystallized with a seven-residue phage-display-derived peptide that bound along a groove in an extended conformation.[28] A central leucine residue from the peptide is completely buried in a deep hydrophobic pocket in the floor of this groove, thereby acting as a key specificity determinant for the chaperone–peptide interaction. Five residues of the peptide are responsible for most of the chaperone–substrate interactions that occur through both side-chain-mediated non-polar contacts and main-chain hydrogen bonds.[28] A methionine side chain from the chaperone wraps over the backbone of the central leucine residue in the peptide, creating an archlike structure to further stabilize the bound peptide.[28] Very similar chaperone–peptide interactions were observed for HscA, another hsp70-class molecular chaperone, except the peptide bound with reverse polarity and the central hydrophobic pocket was occupied by a proline residue instead of a leucine.[29] In the structure of Ydj1, an hsp40 homologue from yeast, a 7-mer peptide binds as an antiparallel β-strand into an open hydrophobic groove on the surface of the chaperone, complementing one of the Ydj1 β-sheets.[30] Analogous to DnaK, a central leucine residue is inserted into a hydrophobic depression in the floor of the binding groove and serves as a specificity determinant. SurA, a bacterial periplasmic chaperone, was co-crystallized with a heptameric peptide selected from a phage-display library. The peptide was selected from a random linear library, displayed at the N-terminus of the minor phage coat protein pIII.[31] Similar to the CsaA–peptide complex, the first peptidyl prolyl isomerase domain (P1) of the protein SurA binds its selected peptide in an extended conformation.[32] As in the structures of DnaK, HscA and Ydj1, the peptide-binding groove contained a hydrophobic pocket occupied by a single peptide side chain, which in this case was tryptophan.

The chaperone–peptide complexes described above have several common features. First, the peptide binds in an extended conformation into a groove on the protein surface. Second, the peptide-binding site contains a hydrophobic pocket designed to accommodate a single peptide residue as a key specificity determinant. Third, the chaperone binds to a short stretch of peptide sequence, and the peptide–chaperone interactions occur largely through van der Waals interactions and main-chain hydrogen bonds. The interactions of AtCsaA with a phage-display-derived peptide share these features. AtCsaA binds to a short five-residue motif in the peptide mostly through main-chain hydrogen bonding and van der Waals interactions. The number of hydrogen bonds and the degree of hydrophobicity of the chaperone–peptide interface in AtCsaA are also similar to that of the other chaperones (Table 4). One unique feature of AtCsaA is that it lacks a deep hydrophobic pocket that would accommodate a single hydrophobic side chain. Instead, specificity may arise through accommodating only small uncharged residues at positions −5 and −3. Previous peptide-scanning analysis suggested that CsaA showed preference for peptides with positively charged termini.[14] The interactions seen between the carboxylate of the C-terminal Cys113 B

Table 4. Characteristics of chaperone–peptide interfaces

Chaperone	PDB ID	Peptide sequence	Interacting residues[a]	Hydrogen bonds[b]	ASA[c] (Å^2)	ASA[d] (%)
DnaK	1DKX	NRLLLTG	5	9	472	48
HscA	1U00	ELPPVKIHC	7	7	594	59
Ydj1	1NLT	GWLYEIS	6	7	424	58
SurA	2PV1	WEYIPNV	7	6	450	62
CsaA	2Q2H	QPTAAN	5	7	344	59

The protein–protein interaction server [http://www.biochem.ucl.ac.uk/bsm/PP/server/] was used to analyze the chaperone–peptide interface.[27]

[a] Number of interacting residues between peptide and chaperone.
[b] Number of hydrogen bonds between the peptide and the chaperone.
[c] Accessible surface area (ASA) of the chaperone buried in the interface with the peptide.
[d] Percentage of non-polar ASA of the chaperone buried in the interface with the peptide.

and the bound peptide are consistent with such a preference (Fig. 4c).

It is interesting that the residues Gln(−6) and Asn(−1) occur at the ends of the ordered peptide region within the AtCsaA-binding groove. It is possible these amide side chains could be mimicking main-chain-type interactions that would occur in native substrates. A search of the protein structure and sequence data bases revealed no proteins with a motif matching the sequence of our phage-display peptide. This may not be too surprising in that CsaA supposedly binds preferentially to disordered regions of molten globular secretory proteins, and we may have selected for a peptide that maintains a disordered structure in solution, yet has the features amenable to binding into the CsaA-binding groove. We hypothesize that CsaA acts similar to other chaperones of broad specificity and is able to bind a wide range of substrates that display the above-described characteristics.

Experimental Procedures

DNA manipulating

The *csaA* gene was amplified by PCR with Hot Start Taq Polymerase (Qiagen), using *A. tumefaciens* genomic DNA. Primers 5′catatgagcggcgaaatttcctatgccgatttc and 5′aagctttcagcacatcttctcaccgttcggcacagg containing the NdeI and HindIII restriction sites, respectively, were used for the PCR amplifications. The PCR product was ligated into the cloning vector pCR-TOPO (Invitrogen) to create the plasmid TOPO-CsaA, which was digested with NdeI and HindIII. The AtCsaA insert was ligated into pET28a+ (Novagen) previously digested with NdeI and HindIII to create the plasmid AtCsaA-pET28. This resulted in an open reading frame containing a 6× histidine-tag at the N-terminus, followed by a thrombin cleavage site.

E. coli BL21(DE3) (Novagen) was transformed with AtCsaA-pET28 and used for protein expression. Cells were grown to an OD_{600} of 0.6 and then induced with 0.5 mM IPTG for 2 h at 37 °C with vigorous shaking. The cells were harvested by centrifugation. The cell pellet was resuspended in 50 mM Tris–HCl (pH 8.0) and 0.1 M NaCl and then lysed by pressure disruption with a French pressure cell (for AtCsaA) or an Avestin EmulsiFlex-C3 cell homogenizer (for X15peptide–AtCsaA). For the X15peptide–AtCsaA construct, primers X15peptide_sense: 5′ GCGGCAGCCAT ATGGTT CCTGG GCAGA AGCAGC ATTATG TTCAG CCGACG GCGGC TAATAG CGGCG AAATTTC and T7-terminator 5′ TATGC TAGTTAT TGCTCAG were used. Primer X15peptide_sense has the DNA sequence encoding the X15.1 peptide at its 5′ end, followed by an overlapping region for annealing to the *csaA* gene. A PCR reaction was performed using the pET28a+CsaA vector as template. The new X15peptide–AtCsaA PCR product was cloned as described above, into pET28a+, to create the plasmid X15.1peptide-CsaA-pET28.

Protein purification

The cellular lysate was centrifuged at 27,000*g* and the clarified supernatant was loaded onto a Ni-NTA column (Qiagen) equilibrated with 20 mM Tris–HCl (pH 8.0) and 0.1 M NaCl. The column was washed with 10 column volumes of a buffer containing 20 mM Tris–HCl (pH 8.0), 0.1 M NaCl and 30 mM imidazole, and then the protein was eluted in a stepwise fashion with 5-ml fractions of 50 mM Tris–HCl (pH 8.0), 0.1 M NaCl, supplemented with 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 or 1 M imidazole. Collected fractions were dialyzed against 20 mM Tris–HCl (pH 8.0), 0.1 M NaCl. The proteins were treated with thrombin to remove the 6× His affinity tag. CsaA proteins were applied to a Sephacryl S-100 HiPrep 26/60 column on an AKTA Prime system (Pharmacia Biotech) and run at 1 ml/min, with a buffer containing 20 mM Tris–HCl (pH 8.0) and 0.1 NaCl. Fractions containing pure CsaA, as analyzed by SDS-PAGE, were concentrated for crystallization using an Amicon ultracentrifuge filter (Millipore). The protein concentration was determined by the bicinchoninic acid protein assay (Pierce).

Mass spectra analysis of X15peptide–AtCsaA

Maldi-TOF analysis of the thrombin-digested form of X15peptide–AtCsaA was analyzed at the Proteomics Core Facility, University of British Columbia.

Biotinylation of AtCsaA

AtCsaA modified with biotin (AtCsaA-biotin) was prepared according to the procedure described by Zwick *et al.*[33] Briefly, 100 μg of AtCsaA lacking the N-terminal His-tag was incubated in 180 mM $NaHCO_3$ buffer in the presence of 0.4 mg/ml NHS-LC-Biotin (Pierce). Reactions were quenched with 1 M Tris–HCl (pH 7.4). Excess biotin was removed with several washes with TBS (50 mM Tris base, 0.9% NaCl, pH 7.6) using a Millipore concentrator. AtCsaA-biotin was stored in the presence of 50% glycerol at −20 °C.

Screening for AtCsaA substrates in phage-display peptide libraries

For biopanning,[34] two peptide libraries were used for screening: X15, containing 15 random amino-acid positions, and LX10, containing 10 random amino-acid positions constrained by two conserved cysteines. Selections were performed with 10 nM AtCsaA-biotin in the presence of 10^{10} phage particles during the first and second rounds, and 1 nM AtCsaA-biotin with 10^9 phage particles during the third round. In an attempt to remove streptavidin-binding phage from the initial libraries,[35] a pre-adsorption step was utilized, where the initial library was incubated overnight with the streptavidin magnetic beads (Invitrogen). Unbound phage were then subjected to solution screening with AtCsaA-biotin at room temperature (3 h for the first two rounds and 1 h for the last round). CsaA/phage complexes were captured with streptavidin beads and washed thoroughly to remove unbound phage. Bound phage was acid-eluted and used to infect freshly made K91 starved cells[36] and amplified in NZY tetracycline (15 μg/ml), 1 mM IPTG. Phage purification was performed with polyethylene glycol (PEG)/NaCl, and purified phage was used for DNA sequencing and ELISA assays.[25] In every round of biopanning, a control selection was performed with F88 phage and AtCsaA-biotin to monitor the background binding of AtCsaA-biotin to wild-type F88 phage. The percentage of phage produced from each cycle was calculated by spot titering of diluted infected pools. Individual phage clones from the second and third round

of biopanning were analyzed for their displayed peptide sequence.

Sequencing

The DNA encoding the displayed peptide of individual phage clones was PCR-amplified with primers F88-4 pVIII forward (5′GCT CTA AAT CGG GGG AGCT) and F88-4 pVIII reverse (5′GCA GAA CCA ATG CAT AAG CTA GCT TAA AAA AAA GCC CGC). PCR-amplified phage DNA was sequenced with primer F88 (5′ CGATAACAACCAC-CATAGC).

Phage ELISA binding assay

For the binding assays,[36] 1×10^{10} phage particles (in a 35-μl volume) were incubated overnight in a microtiter plate at 4 °C. Plates were washed three times with TBS/0.1% Tween 20. Plates were blocked with 200 μl of 2% bovine serum albumin in TBS for 1 h at 37 °C. Plates were washed again (six times) with TBS/Tween 20. Next, 35 μl of 50 nM CsaA-Bio was added, and the mixture was incubated at room temperature for 2 h. Plates were washed six times, and a 1:1000 dilution of neutravidin-conjugated horseradish peroxidase (HRP, Pierce) was added to the plate. The plate was washed again six times, and the assay was developed with 0.03% (v/v) H_2O_2 and 400 μg/ml of 2,2′-azino-bis(3-ethylbenzthiazoline-6-sulfonic), according to the methods of Menendez *et al.*[36] The absorbance was measured at 405 and 495 nm in a VersaMax tunable microplate reader (Molecular Devices, Sunnyvale, CA). Reported values are the absorbance at 405 nm minus the absorbance at 490 nm after 60 min incubation. To verify that each well had similar amounts of phage, parallel ELISA assays were performed in which the primary binder used was rabbit polyclonal anti-phage antibody and the secondary binder was protein-A-conjugated HRP. The assay was developed as described above.

Crystallization

Sitting-drop vapor diffusion at room temperature was used to crystallize AtCsaA. The final refined reservoir conditions were 1.8 M ammonium sulfate, 0.1 M Hepes, pH 7.5, 2% PEG 400 and 5% ethylene glycol. Drops consisted of 1 μl protein (13 mg/ml) and 1 μl reservoir. The cryosolvent was the same as in the growth conditions with 15% of the water replaced with ethylene glycol. Crystals of X15peptide–AtCsaA were produced by the sitting-drop vapor-diffusion technique at 19 °C. Drops consisted of 1 μl protein (9 mg/ml) and 1 μl reservoir. The optimized reservoir solution for the X15peptide–AtCsaA crystal was 30 PEG 4000, 0.4 M NH_4OAc and 0.1 M sodium citrate (pH 5.5). The mother liquor was used as the cryosolution, and crystals were mounted directly from the drop into the cryostream.

Data collection

Diffraction data were collected with an RAXIS IV++ image plate detector mounted on a 007 Rigaku X-ray generator with VariMax CuHF optics. Data for AtCsaA and X15peptide–AtCsaA crystals were collected with a crystal-to-detector distance of 120 and 150 mm, respectively. Data were collected at 100 K using an X-stream 2000 cryosystem. A total of 192 and 186 frames were collected for the AtCsaA and X15peptide–AtCsaA data sets, respectively, using 0.5° oscillations. Each image was exposed for 2 min. Data were indexed, integrated and scaled with the program Crystal Clear.[37]

Phasing, model building and refinement

Molecular replacement solutions were found with the program Phaser 1.2.[38] The search model used was a homology model of AtCsaA constructed with CPH[39] based on coordinates from TtCsaA [Protein Data Bank (PDB) 1DG7, chain A; AtCsaA and TtCsaA have 48% sequence identity]. The structures were refined using restrained refinement in REFMAC5,[40] and simulated annealing, energy minimization and *B*-factor refinement in CNS.[26] Manual adjustments to the atomic coordinates were performed with the program COOT.[41] The final models were obtained by running restrained refinement in REFMAC5 with TLS restraints obtained from the TLS motion determination server.[42] Crystal and data collection and refinement statistics are shown in Table 2.

Structural analysis

The program PROCHECK[43] was used to analyze the quality of the final refined model. The program BAVERAGE within CCP4i[44] was used to calculate the average *B*-factor of the final refined model. The programs SUPERIMPOSE[45] and SuperPose[46] were used to compare CsaA monomers. The program CASTp[47] was used to analyze the surface of CsaA. Hydrogen-bonding distances were measured with the WhatIF server[48] and the program HBPLUS.[49] Intermolecular interactions were measured using the protein–protein interaction server‡.

Figure preparation

Figures were made with PYMOL.[50] The sequence alignment was made with the program ClustalW[51] and ESPript.[52]

Prevention of protein aggregation assays

The *in vitro* chaperone activities of AtCsaA and X15peptide–AtCsaA were determined essentially as described by Leroux *et al.*[27] Briefly, lysozyme (chicken egg white, Sigma L7651, 200 μM) was denatured in 6 M guanidine-HCl, 100 mM NaCl, 20 mM sodium phosphate and 50 mM DTT (pH 8.0). Denatured lysozyme was diluted 100-fold into buffer [100 mM NaCl, 20 mM sodium phosphate and 50 mM DTT (pH 8.0)] or buffer containing 4 μM CsaA or X15peptide–CsaA. Lysozyme aggregation was monitored as increased light scattering at 360 nm for 10 min. A Varian Cary 100-bio UV–vis spectrophotometer was used. All experiments were performed at 25 °C. Each sample was run in triplicate and repeated in separate occasions. The data presented are representative trials.

Protein Data Bank accession numbers

The atomic coordinates and structure factors (PDB codes 2Q2H for the AtCsaA structure with the bound peptide and 2Q2I for the AtCsaA structure without the peptide) have been deposited in the PDB, Research Collaboratory

‡ http://www.biochem.ucl.ac.uk/bsm/PP/server

for Structural Bioinformatics, Rutgers University, New Brunswick, NJ§.

Acknowledgements

We thank Drs. Alfredo Menendez and Andrew L. Lovering for helpful discussions and Drs. Michel R. Leroux and Victor F. Lundin for advise on the aggregation assays and use of their spectrophotometer. This work was supported in part by a Canadian Institute of Health Research operating grant (to M.P.), National Science and Engineering Research Council of Canada discovery grants (to M.P. and L.P.M.), a Michael Smith Foundation for Health Research Senior Scholar award (to M.P.), a Canadian Foundation of Innovation grant (to M.P.), NIH operating grant AI049111 (to J.K.S.) a Canadian Cystic Fibrosis Foundation Postdoctoral Fellowship Award (to A.R.F.), a Michael Smith Foundation for Health Research Post-doctoral Fellowship Award (to D.C.O.) and a Feodor Lynen Fellowship from the Alexander von Humboldt Foundation (to M.H.). NMR instrument support was provided by the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, the British Columbia Knowledge Development Fund, the UBC Blusson Fund, and the Michael Smith Foundation for Health Research.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.03.048

References

1. de Keyzer, J., van der Does, C. & Driessen, A. J. (2003). The bacterial translocase: a dynamic protein channel complex. *Cell Mol. Life Sci.* **60**, 2034–2052.
2. Driessen, A. J., Manting, E. H. & van der Does, C. (2001). The structural basis of protein targeting and translocation in bacteria. *Nat. Struct. Biol.* **8**, 492–498.
3. Zhou, J. & Xu, Z. (2003). Structural determinants of SecB recognition by SecA in bacterial protein translocation.*Nat. Struct. Biol.* **10**, 942–947; Epub 2003 Sep 28.
4. Vrontou, E. & Economou, A. (2004). Structure and function of SecA, the preprotein translocase nanomotor. *Biochim. Biophys. Acta*, **1694**, 67–80.
5. Bessonneau, P., Besson, V., Collinson, I. & Duong, F. (2002). The SecYEG preprotein translocation channel is a conformationally dynamic and dimeric structure. *EMBO J.* **21**, 995–1003.
6. Veenendaal, A. K., van der Does, C. & Driessen, A. J. (2004). The protein-conducting channel SecYEG. *Biochim. Biophys. Acta*, **1694**, 81–95.
7. Paetzel, M., Dalbey, R. E. & Strynadka, N. C. (1998). Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. *Nature*, **396**, 186–190.
8. Paetzel, M., Dalbey, R. E. & Strynadka, N. C. (2002). Crystal structure of a bacterial signal peptidase apoenzyme: implications for signal peptide binding and the Ser–Lys dyad mechanism. *J. Biol. Chem.* **277**, 9512–9519.
9. Pohlschroder, M., Dilks, K., Hand, N. J. & Wesley Rose, R. (2004). Translocation of proteins across archaeal cytoplasmic membranes. *FEMS Microbiol. Rev.* **28**, 3–24.
10. Bolhuis, A. (2004). The archaeal Sec-dependent protein translocation pathway. *Philos. Trans. R. Soc. London, Ser. B*, **359**, 919–927.
11. Yamane, K., Bunai, K. & Kakeshita, H. (2004). Protein traffic for secretion and related machinery of *Bacillus subtilis*. *Biosci. Biotechnol. Biochem.* **68**, 2007–2023.
12. van Wely, K. H., Swaving, J., Freudl, R. & Driessen, A. J. (2001). Translocation of proteins across the cell envelope of Gram-positive bacteria. *FEMS Microbiol. Rev.* **25**, 437–454.
13. Muller, J. P., Ozegowski, J., Vettermann, S., Swaving, J., Van Wely, K. H. & Driessen, A. J. (2000). Interaction of *Bacillus subtilis* CsaA with SecA and precursor proteins. *Biochem. J.* **348**(Pt 2), 367–373.
14. Linde, D., Volkmer-Engert, R., Schreiber, S. & Muller, J. P. (2003). Interaction of the *Bacillus subtilis* chaperone CsaA with the secretory protein YvaY. *FEMS Microbiol. Lett.* **226**, 93–100.
15. Xu, Z., Knafels, J. D. & Yoshino, K. (2000). Crystal structure of the bacterial protein export chaperone secB. *Nat. Struct. Biol.* **7**, 1172–1177.
16. Kawaguchi, S., Muller, J., Linde, D., Kuramitsu, S., Shibata, T., Inoue, Y. *et al.* (2001). The crystal structure of the ttCsaA protein: an export-related chaperone from *Thermus thermophilus*. *EMBO J.* **20**, 562–569.
17. Muller, J. P., Bron, S., Venema, G. & van Dijl, J. M. (2000). Chaperone-like activities of the CsaA protein of *Bacillus subtilis*. *Microbiology*, **146**, 77–88.
18. Shapova, Y. A. & Paetzel, M. (2007). Crystallographic analysis of Bacillus subtilis CsaA. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **63**, 478–485.
19. Swairjo, M. A., Morales, A. J., Wang, C. C., Ortiz, A. R. & Schimmel, P. (2000). Crystal structure of trbp111: a structure-specific tRNA-binding protein. *EMBO J.* **19**, 6287–6298.
20. Escobar, M. A. & Dandekar, A. M. (2003). *Agrobacterium tumefaciens* as an agent of disease. *Trends Plant Sci.* **8**, 380–386.
21. Wood, D. W., Setubal, J. C., Kaul, R., Monks, D. E., Kitajima, J. P., Okura, V. K. *et al.* (2001). The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science*, **294**, 2317–2323.
22. Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B. *et al.* (2001). Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science*, **294**, 2323–2328.
23. Van Sluys, M. A., Monteiro-Vitorello, C. B., Camargo, L. E., Menck, C. F., Da Silva, A. C., Ferro, J. A. *et al.* (2002). Comparative genomic analysis of plant-associated bacteria. *Annu. Rev. Phytopathol.* **40**, 169–189.
24. Zwick, M. B., Bonnycastle, L. L., Noren, K. A., Venturini, S., Leong, E., Barbas, C. F., 3rd *et al.* (1998). The maltose-binding protein as a scaffold for monovalent display of peptides derived from phage libraries. *Anal. Biochem.* **264**, 87–97.
25. Menendez, A., Bonnycastle, L. L. C., Pan, O. O. C. & Scott, J. K. (2001). *Chapter 17 in Phage Display: A Laboratory Manual*. (Barbas, C. F., III, Burton, D. R., Scott, J. K. & Silverman, G., eds), Cold Spring Harbor Laboratories Press, Plainview, NY.

§ http://www.rcsb.org/

26. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54**, 905–921.
27. Leroux, M. R., Fandrich, M., Klunker, D., Siegers, K., Lupas, A. N., Brown, J. R. *et al.* (1999). MtGimC, a novel archaeal chaperone related to the eukaryotic chaperonin cofactor GimC/prefoldin. *EMBO J.* **18**, 6730–6743.
28. Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E. & Hendrickson, W. A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science*, **272**, 1606–1614.
29. Cupp-Vickery, J. R., Peterson, J. C., Ta, D. T. & Vickery, L. E. (2004). Crystal structure of the molecular chaperone HscA substrate binding domain complexed with the IscU recognition peptide ELPPVKIHC. *J. Mol. Biol.* **342**, 1265–1278.
30. Li, J., Qian, X. & Sha, B. (2003). The crystal structure of the yeast Hsp40 Ydj1 complexed with its peptide substrate. *Structure*, **11**, 1475–1483.
31. Bitto, E. & McKay, D. B. (2003). The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J. Biol. Chem.* **278**, 49316–49322.
32. Xu, X., Wang, S., Hu, Y. X. & McKay, D. B. (2007). The periplasmic bacterial molecular chaperone SurA adapts its structure to bind peptides in different conformations to assert a sequence preference for aromatic residues. *J. Mol. Biol.* **373**, 367–381.
33. Zwick, M. B., Menendez, A., Bonnycastle, L. L. C. & Scott, J. K. (2001). *Chapter 18 in Phage Display: A Laboratory Manual*. (Barbas, C. F., III, Burton, D. R., Scott, J. K. & Silverman, G., eds), Cold Spring Harbor Laboratories Press, Plainview, NY.
34. Bonnycastle, L. L., Mehroke, J. S., Rashed, M., Gong, X. & Scott, J. K. (1996). Probing the basis of antibody reactivity with a panel of constrained peptide libraries displayed by filamentous phage. *J. Mol. Biol.* **258**, 747–762.
35. Menendez, A. & Scott, J. K. (2005). The nature of target-unrelated peptides recovered in the screening of phage-displayed random peptide libraries with antibodies. *Anal. Biochem.* **336**, 145–157.
36. Menendez, A., Chow, K. C., Pan, O. C. & Scott, J. K. (2004). Human immunodeficiency virus type 1-neutralizing monoclonal antibody 2F5 is multispecific for sequences flanking the DKW core epitope. *J. Mol. Biol.* **338**, 311–327.
37. Pflugrath, J. W. (1999). The finer things in X-ray diffraction data collection. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 1718–1725.
38. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **61**, 458–464.
39. Lund, O., Nielsen, M., Lundegaard, C. & Worning, P. (2002). *Abstract at the CASP5 Conference A102.* Asilomar, CA, USA Dec. 1–5, 2002.
40. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **57**, 122–133.
41. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
42. Painter, J. & Merritt, E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 439–450.
43. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Crystallogr.* **26**, 283–291.
44. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). A graphical user interface to the CCP4 program suite. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **59**, 1131–1137.
45. Diederichs, K. (1995). Structural superposition of proteins with unknown alignment and detection of topological similarity using a six-dimensional search algorithm. *Proteins*, **23**, 187–195.
46. Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart, D. S. (2004). SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* **32**, W590–W594.
47. Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **34**, W116–W118.
48. Hooft, R. W., Sander, C. & Vriend, G. (1996). Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. *Proteins*, **26**, 363–376.
49. McDonald, I. K. & Thornton, J. M. (1994). Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793.
50. DeLano, W. L. (2002). The PyMOL User's Manual. DeLano Scientific, Palo Alto, CA, USA.
51. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
52. Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305–308.