Chapter 5 - Mechanisms contributing to stabilization of bacterial cold shock proteins

5.1 Results

The two *Bs*-CspB variants, whose crystal structures will be analyzed in this chapter, were derived from *Proside in-vitro* selections. Their thermodynamic analysis was published elsewhere [11].

5.1.1 Crystallization and structure solution of two stabilized *Bs*-CspB variants

Two *Bs*-CspB stability mutants were crystallized in space groups $P4_32_12$ (variant M1R/E3K/K65I) and $P3_221$ (variant A46K/S48R) (Figure 5.1, Figure 5.2). Diffraction data were collected at BESSY synchrotron beamline 14.2 [110], and could be used to maximal resolutions of 2.6 Å (*Bs*-CspB M1R/E3K/K65I) and 2.3 Å (*Bs*-CspB A46K/S48R) (Table 5.1, Figure 5.3). The two structures were solved by molecular replacement using a structural model of *Bs*-CspB (1CSP) and refined with REFMAC5 [121] to final R_{work} / R_{free} values better than 23% / 25% in space groups $P4_32_12$ (*Bs*-CspB M1R/E3K/K65I) and $P3_221$ (*Bs*-CspB

Crystallization of Bs-CspB M1R/E3K/K65I
reservoir solution 25% PEG 3350 0.2 M sodium carbonate 0.1 M TRIS HCl pH 8.5
protein buffer 20 mM TRIS (pH 7.5) 50 mM soldium chloride 3 mM magnesium chloride 20.4 mg / ml protein
crystallization setup 0.8 µl protein : 0.8 µl reservoir solution sitting drop crystallization, automated setup

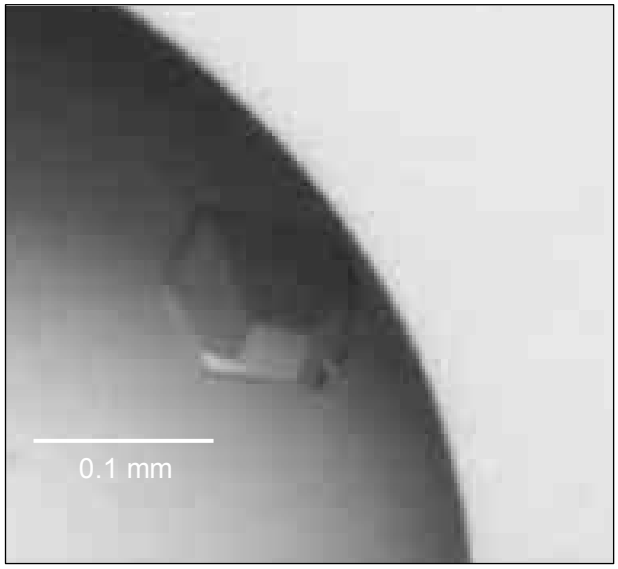


Figure 5.1 Crystallization of *Bs*-CspB M1R/E3K/K65I. The crystal shown on the right was grown according to the given conditions and was used for data collection.

Crystallization of Bs-CspB A46S/S48R
reservoir solution 1.5 M lithium sulfate 0.1 M TRIS pH 7.5 (add. of 15% glycerol for cryoprotection)
protein buffer 20 mM TRIS (pH 7.5) 50 mM NaCl 3 mM $MgCl_2$ 17.4 mg / ml protein
crystallization setup 0.8 µl protein : 0.8 µl reservoir solution sitting drop crystallization, automated setup

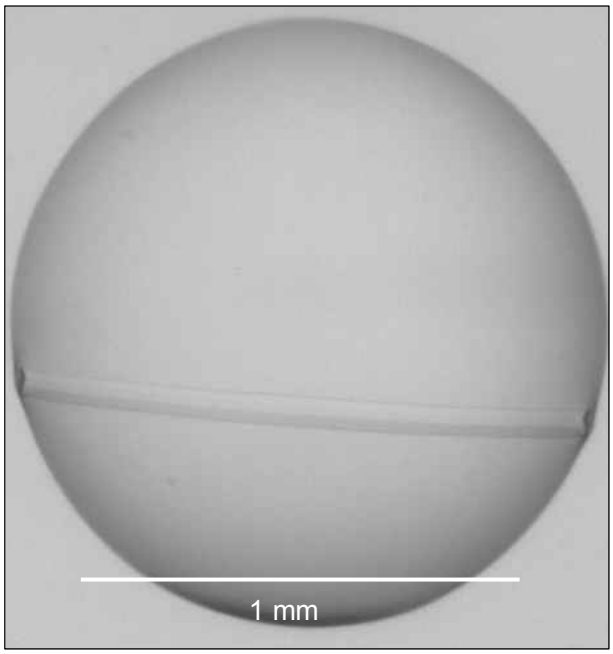


Figure 5.2: Crystallization of *Bs*-CspB A46K/S48R. The crystal shown on the right was grown according to the given conditions and a fragment of it was used for data collection.

Data collection & processing			
***Bs*-CspB variant**		**M1R/E3K/K65I**	**A46K/S48R**
wavelength (Å)		0.9184	0.9184
resolution (Å) resolution (Å)	 (last shell)	18.50 - 2.55 2.62 - 2.55	40.00 – 2.30 2.36 – 2.30
space group		$P4_32_12$	$P3_221$
temperature (K) X-ray source detector		110 BESSY BL14.2 mar165CCD	110 BESSY BL14.2 mar165CCD
unit-cell parameters	*a* (Å) *b* (Å) *c* (Å)	55.50 55.50 55.47	58.63 58.63 46.80
unique reflections	(last shell)	3038 (237)	4231 (1260)
$I/\sigma(I)$	(last shell)	21.40 (5.62)	28.41 (7.24)
% data completeness	(last shell)	98.1 (96.7)	97.2 (95.8)
R_{meas} [a] (%)	(last shell)	7.4 (46.8)	4.8 (37.5)

Table 5.1: X-ray diffraction-data collection of *Bs*-CspB M1R/E3K/K65I and *Bs*-CspB A46K/S48R.
[a] R_{meas} is a redundancy-independent *R* factor, which correlates intensities from symmetry-related reflections [3].

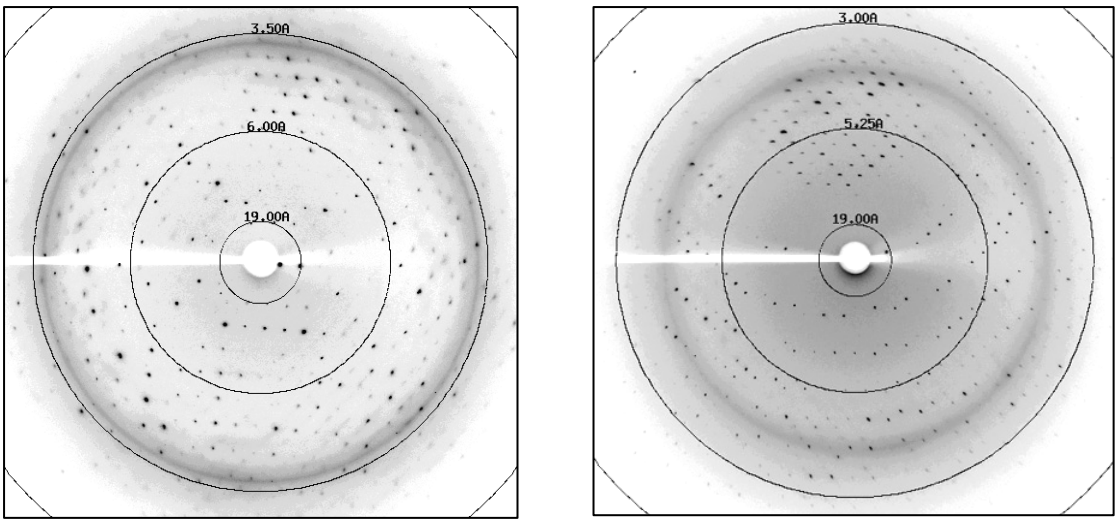
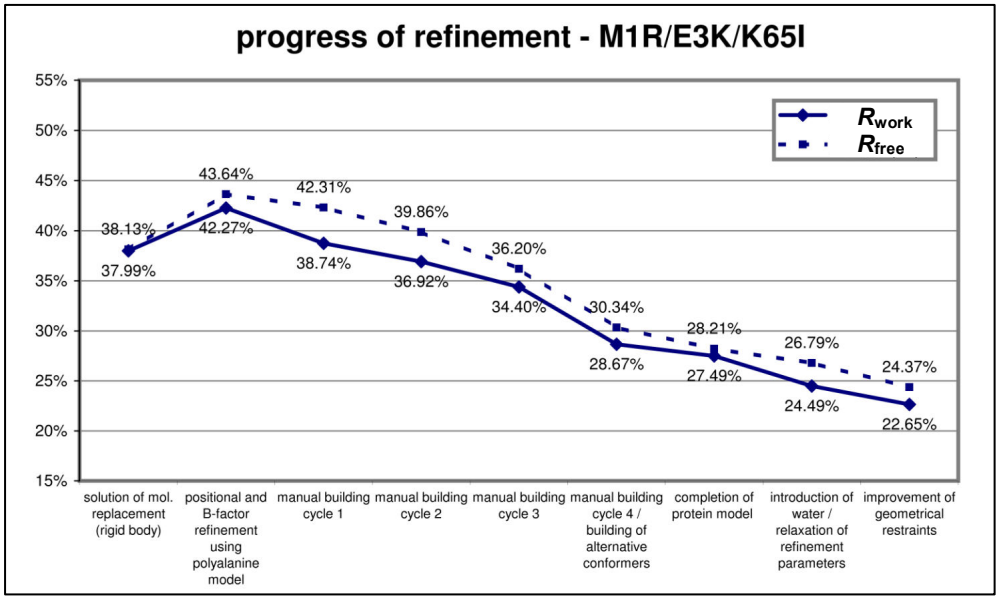


Figure 5.3: Diffraction images of stabilized *Bs*-CspB variants. Left: M1R/E3K/K65I, the outer ring refers to a resolution of 2.5 Å. Right: A46/S48R, the outer ring refers to a resolution of 2.25 Å.



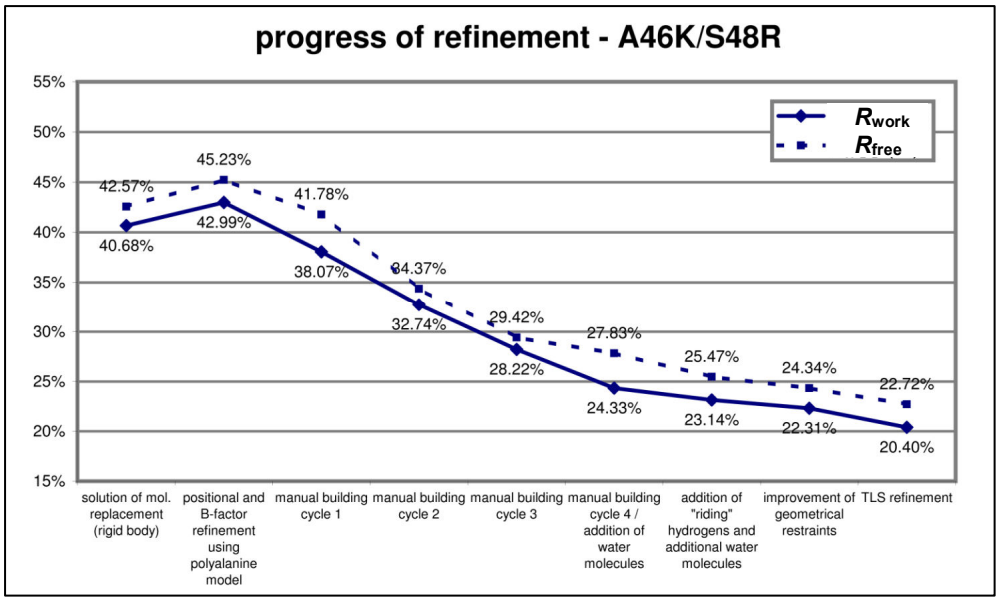


Figure 5.4: Refinement progress of stabilized *Bs*-CspB variants.

A46K/S48R) (Table 5.2). Individual steps in the refinement procedures and their effect on the crystallographic *R* factors are illustrated in Figure 5.4. The two crystal structures are of comparable quality as the two structures of wildtype *Bs*-CspB deposited in the PDB [75]. The crystal used for data collection of mutant A46K/S48R is isomorphous with a crystal of *Bs*-CspB (1CSP). The crystal used for data collection of mutant M1R/E3K/K65I is isomorphous with another crystal of *Bs*-CspB (1CSQ).

Model building, refinement & evaluation		
***Bs*-CspB variant**	**M1R/E3K/K65I**	**A46K/S48R**
PDB model for mol. replacement	1CSP	1CSP
PDB access. of the refined structure	2I5L	2I5M
reflections – working set	2803	4231
reflections - free set	134 (4.4%)	260 (5.0%)
R_{work} (%)[a]	22.65	20.40
R_{free}[b] (%)[a]	24.37	22.72
Mean *B* factor (Å^2)	49.45	33.46
Number of non-hydrogen atoms	533	554
Number of protein molecules	1	1
Number of water molecules	10	23
root mean square deviations		
bond lengths (Å)	0.016	0.015
bond angles (°)	1.717	1.617
torsion angles (°)	5.794	5.867
planarity (Å)	0.007	0.009
Ramachandran statistics		
residues in allowed regions	92.7	94.4
res. in additionally allowed regions	5.5	5.6
res. in generously allowed regions	1.8	0.0
res. in forbidden regions	0.0	0.0
real-space correlation coefficient[b]	87.8%	88.1%

Table 5.2: Building, refinement and evaluation of atomic models of *Bs*-CspB variants M1R/E3K/K65I and A46K/S48R.

[a] $R_{work,free} = \sum \frac{\left\| F_{obs} \right| - \left| F_{calc} \right\|}{\left| F_{obs} \right|}$ working and free *R* factors were calculated using the working and free reflection sets, respectively. The free reflections were held aside throughout refinement [2].

[b] Real-space correlation coefficients of structural model and experimental were calculated using the SFCHECK software [15].

5.1.2 Global structure of stabilized *Bs*-CspB variants

The asymmetric units of the crystals obtained for the stabilized *Bs*-CspB variants contain single CSP molecules. The mutations M1R/E3K/K65I and A46K/S48R left the backbone structure of *Bs*-CspB virtually unchanged. Superpositions of the structures of the two new variants with the two structures for *Bs*-CspB (1CSP and 1CSQ) reveal α-carbon RMS deviations below 0.7 Å, which indicate a level of conservation as observed between the two structures of *Bs*-CspB (RMSD = 0.70 Å). Regions with deviations larger than 0.80 Å between *Bs*-CspB and mutant variants involve the N- and C-termini as well as short segments (2-3 amino acids) in the loops L_{23} and L_{34}. Deviations of similar magnitude in these regions between the *Bs*-CspB structures suggest that differences in the crystal packing rather than amino acid changes are responsible for these structural deviations. Apart from Lys46 the mutated sidechains in the *Bs*-CspB structures are not involved in intermolecular contacts in the crystals and thus their conformations should not be affected by intermolecular interactions in the crystal. Lys46 in *Bs*-CspB A46K/S48R forms a hydrogen bond with the carbonyl group of Gly44 from a symmetry-related molecule.

In *Bs*-CspB A46K/S48R, no electron density was observed for the C-terminal residue (Ala67), suggesting that this residue remains flexible or adopts multiple conformations in the crystal. The other mutated residues are not involved in intermolecular contacts, and thus their conformations should not be affected by crystal packing. The sidechain of Glu66, whose unfavorable interaction with Glu3 could be observed in a *Bc*-Csp R3E/E46A/L66E mutant [49] and which has been linked to the extensive destabilization of *Bs*-CspB in comparison to *Bc*-Csp, is reoriented in the two novel *Bs*-CspB mutant structures (Figure 5.5).

5.1.3 Experimental sampling of coulombic interactions

In order to determine protein stabilities, Gibbs free energies of denaturation (ΔG_D) were determined experimentally by analyzing thermal unfolding transitions using CD spectroscopy. These data were kindly provided by Michael Wunderlich, University of Bayreuth [11]. Stability differences of the mutant variants with respect to *Bs*-CspB are expressed as differences in their Gibbs free energies of denaturation ($\Delta\Delta G_D$). Using this approach, single mutations and their combinations were analyzed, in order to evaluate if the stabilizing effects associated with individual amino acid changes are additive. The change in stability between 0 M and 2 M NaCl had been used as a lower measure for the change in coulombic interactions caused by a specific mutation [145-147] and served this purpose in this study, too.

Variant	T_M (°C)	ΔH_D (kJ mol^{-1})	$\Delta\Delta G$ (kJ mol^{-1}) 0 M NaCl (total)	0 - 2 M NaCl (coulombic)	2 M NaCl (non-polar)
wt *Bs*-CspB	53.8	193	0	0	0
A46K/S48R	69.7	219	11.1	6.1	5.0
A46K	62.2	216	5.9	2.7	3.2
S48R	62.7	202	6.6	2.6	4.0
M1R/E3K/K65I	83.7	278	20.9	8.2	12.7
M1R/E3K	77.0	262	16.3	10.4	5.9
M1R	64.2	226	7.3	3.9	3.4
E3K	70.4	203	11.5	8.4	3.1
E3K/E66K	64.6	198	8.0	6.3	1.7
K65I	62.4	201	6.4	0.0	6.4
E66K	66.7	226	9.1	5.2	3.9

Table 5.3: Stability data for variants of *Bs*-CspB. T_M is the midpoint of the thermal unfolding transition; ΔH_D is the enthalpy of unfolding at T_M. $\Delta\Delta G_D$ is the change in Gibbs free energy of unfolding at 70 °C relative to the wildtype protein. $\Delta\Delta G_D$ at 0 M NaCl is the total change, $\Delta\Delta G_D$ at 2 M NaCl represents the non-polar contribution, and the difference between 0 M and 2 M NaCl represents the Coulombic contribution to $\Delta\Delta G_D$. The thermodynamic parameters were derived from thermal unfolding transitions. All data in this table were kindly provided by Michael Wunderlich (Lehrstuhl für Biochemie, University of Bayreuth) [11].

5.2 Discussion

5.2.1 Stabilization in *Bs*-CspB M1R/E3K/K65I

Wildtype *Bs*-CspB is only marginally stable. Its T_M is 53.8 °C (Table 5.3) and at 25 °C the Gibbs free energy of denaturation, ΔG_D, is only 11.3 kJ·mol^{-1} [89]. The three selected mutations M1R, E3K, and K65I increased the T_M value by 29.9 °C and ΔG_D by 20.9 kJ·mol^{-1} (at 70 °C). Hence, this triple mutant is even slightly more stabile than *Tm*-Csp from the hyperthermophilic bacterium *Thermotoga maritima*, (T_m = 82.0 °C at pH 7) [148]. The sequences of *Bs*-CspB and *Tm*-Csp differ at 22 positions (approximately 33%).

Two of the three stabilizing mutations (M1R and E3K) are in the N-terminal region of strand β1. This region is in close contact with strand β4 and is thus important for closing the globular β-barrel. (Figure 5.5a, Figure 5.6a)**.** In the *Bs*-CspB, Met1 adopts different conformations and multiple conformations in *Bc*-Csp structures [49, 77]. Its sidechain is not involved in attractive polar interactions with other sidechains and it is therefore expected to be quite flexible in solution. In the triple mutant, the sidechain of Arg1 is well ordered and adopts a single conformation (Figure 5.6a, left). The orientation of its guanidino group suggests that it forms a salt bridge or hydrogen bond with the carboxylate of Glu50, which is located in β-strand 5.

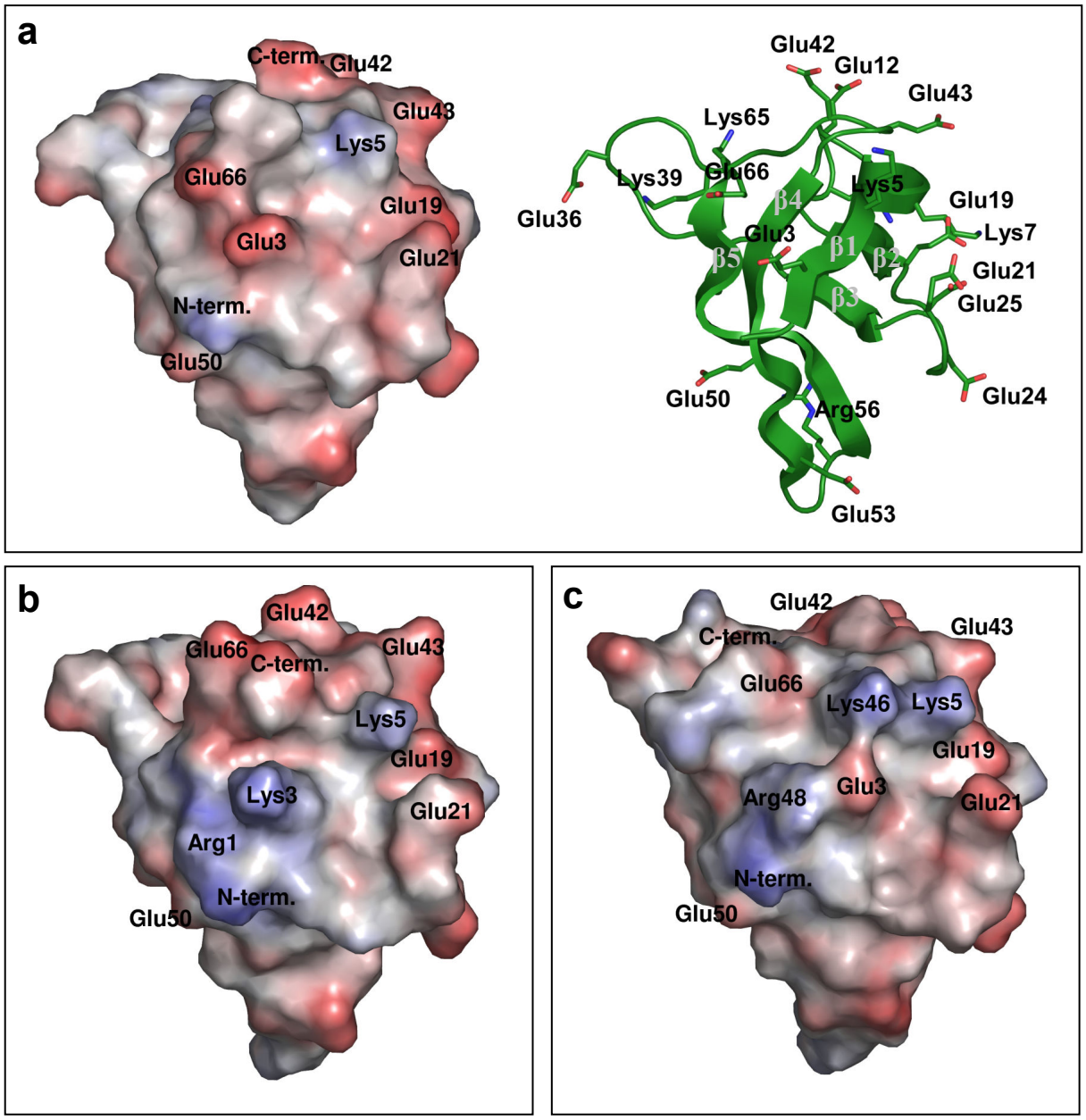
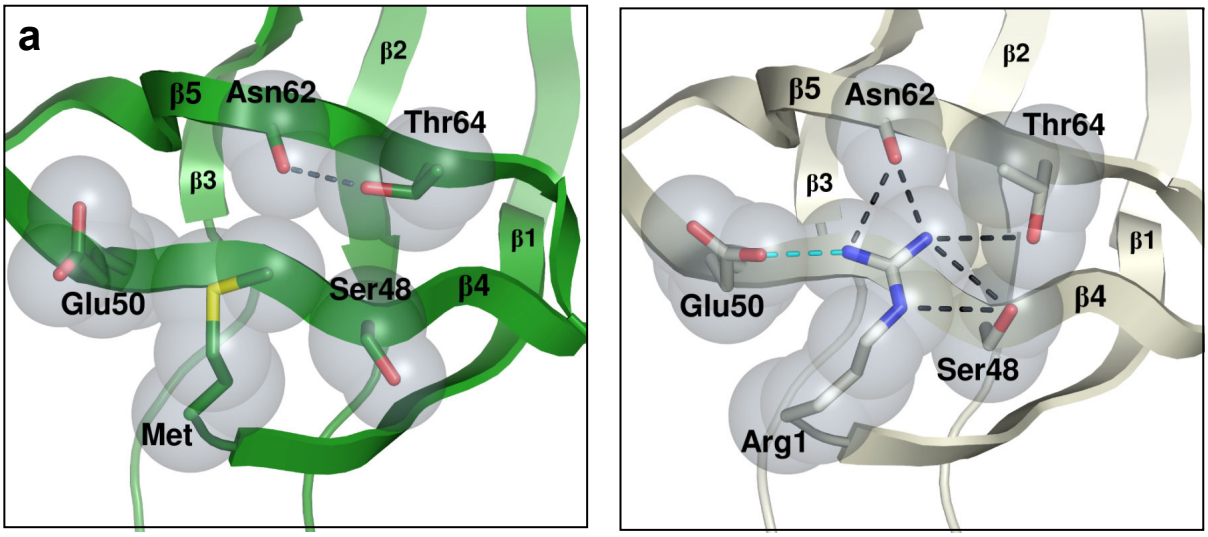


Figure 5.5: Electrostatic surface potential of *Bs*-CspB (a) and mutants M1R/E3K/K65I (b) & A46K/S48R (c). The surface charge of *Bs*-CspB is highly negative. In both mutant structures, additional basic sidechains give rise to more balanced surface potentials. A cartoon model of *Bs*-CspB featuring all acidic and basic sidechains is included in a). The surface potential was calculated using APBS [10] for pH 7.0. The depicted potential covers a range from -10kT (red) to +10kT (blue). For mutant A46K/S48R the C-terminal residue, which was missing in the electron density, was included in the atomic model. Likewise, groups missing in the other structures were added.

In the wildtype structures, the sidechain of Glu50 does not interact with basic residues. In addition, the Arg1 guanidino group forms five non-local hydrogen bonds with the sidechains of Ser48 and Thr64 and the backbone carbonyl group of Asn62. These extensive interactions are possible, because the sidechains of Ser48, Glu50, and Thr64 are slightly reoriented in the mutant protein. In the wildtype structures, a hydrogen bond is formed between the sidechain hydroxyl of Thr64 and the carbonyl oxygen of Asn62, which is absent in the triple mutant



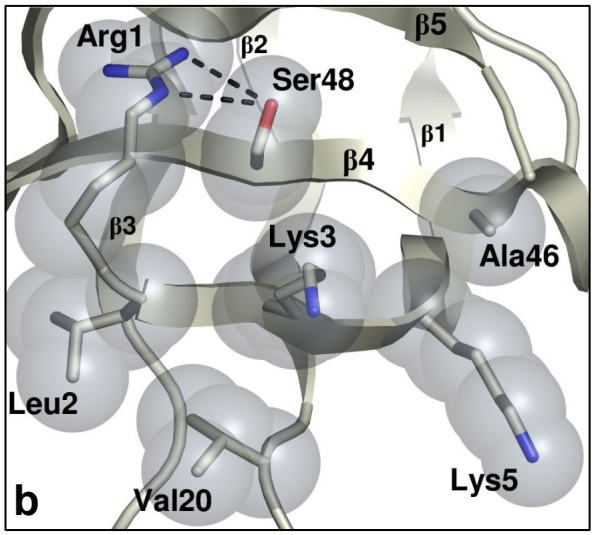
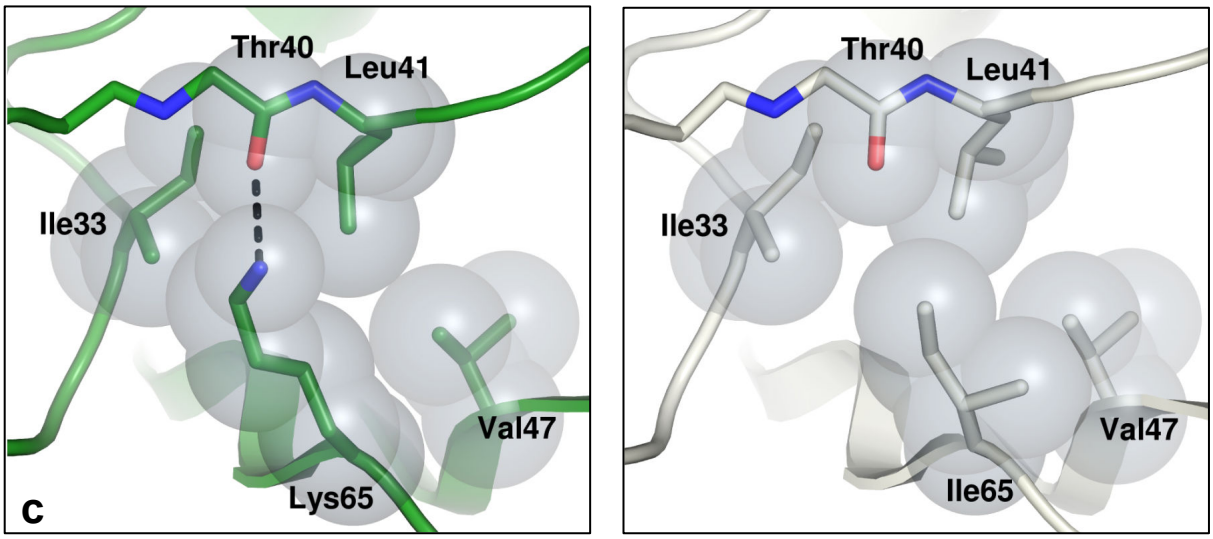


Figure 5.6: Stabilizing effects associated with individual amino acid changes in *Bs*-CspB variant M1R/E3K/K65I. Relevant groups are depicted as sticks. Potential hydrogen bonds and salt bridges are depicted as black and cyan dashed lines, respectively. Carbon atoms of *Bs*-CspB (1CSP) and mutant variants are rendered in green (left images) and gray (right images). Van-der-Waals surfaces of selected groups are shown as semitransparent spheres.
a) Stabilizing effect of M1R: Five additional hydrogen bonds and a salt bridge are formed between the guanidine group of Arg 1 and its surrounding. In the wild-type a hydrogen bond is observed between Thr64 and Asn62.
b) The introduction of a lysine at position three significantly stabilizes *Bs*-CspB. This sidechain does not form a pairwise interaction with surrounding polar residues.
c) Stabilizing effect of K65I: The terminal amino group of the Lys65 sidechain is expected to unfavorably interact with adjacent hydrophobic groups from the protein core. The sidechain of Ile65, like that of a lysine, can form several van-der-Waals contacts with adjacent hydrophobic groups but does not introduce destabilizing charge.



(Figure 5.6a, right). In addition, Arg1 may compensate unfavorable interactions between anionic groups in the surrounding of residue 3 (Figure 5.5a, b). The M1R mutation improves also the local molecular packing as is evident from the comparison of the panels of Figure 5.6a. This might shield the hydrophobic core better from the polar solvent. In summary, the strong thermodynamic stabilization by Arg1 with its strong coulombic part is well explained by the additional salt bridge and H bonds and the improved packing.

The residue at position 3 is very important for the stability of *Bs*-CspB. In the wildtype a glutamate is located at this position, whose carboxyl group is disordered in the two crystal structures. The thermophilic homolog *Bc*-Csp carries an Arg at position3, and, in fact, this variation at position 3 largely accounts for the strong difference in stability between *Bs*-CspB and *Bc*-Csp. The single E3R mutation stabilizes *Bs*-CspB by 11.1 kJ/mol, and most of this increase originates from improved coulombic interactions [91]. The E3K mutation leads to an increase in ΔG_D of 11.5 kJ/mol (Table 5.3). This strong stabilization originates predominantly from the removal of the unfavorable interactions of the wildtype residue Glu3 with anionic groups in its vicinity [49], in particular with Glu66 and the C-terminal carboxyl group. Lys3 in the stabilized variant does not participate in strong short-range charge interactions with surrounding groups (Figure 5.5b, Figure 5.6b). This is an intriguing result and implies that the very strong stabilization by the E3K mutation is not caused by favorable local coulombic interactions of Lys3, but rather by a general improvement of the charge distribution, and in particular by the removal of the strongly unfavorable charge interactions of Glu3 in the wildtype protein (Figure 5.5a). Wildtype *Bs*-CspB is overall destabilized by coulombic interactions and therefore its stability increases strongly when these unfavorable interactions are shielded by the addition of NaCl [91]. In the M1R/E3K/K65I mutant, however, the coulombic interactions are strongly improved and thus overall stabilizing. This variant is therefore destabilized by adding salt [11]. This strong change in the coulombic interactions is not immediately evident in the crystal structures (Figure 5.5a, c). Surprisingly, the side chain of Glu66, whose unfavorable interaction with Glu3 destabilizes wildtype *Bc*-Csp, has moved away from position 3 in the structure of the M1R/E3K/K65I mutant although a positively charged residue (Lys) occupies position 3 in this variant. As a consequence, the distance between the positions 3 and 66 ($d > 9.5$ Å) is too long for a strong interaction. It cannot be excluded that differences in the crystal packing cause these differences in the orientation of Glu66.

The M1R and E3K mutations together with the N-terminus, Lys5, and Lys7 lead to a high local density of positive charges in the N-terminal region. This might destabilize the unfolded protein and thus additionally contribute to the increased the stability of the folded state.

The stabilizing effect of the K65I mutation was unexpected: *Bs*-CspB has a predicted net charge of -6 (at pH 7). Of the know mutations that increases the overall net charge further K65I is the only one that strongly increases the stability [11]. The stabilization of 6.4 kJ·mol^{-1} (Table 5.3) is independent of the salt concentration, suggesting that changes in coulombic interactions do not contribute to the stabilization of this variant. In wildtype *Bs*-CspB the aliphatic moiety of Lys65 is in the van-der-Waals contact range (3.5 Å > *d* > 4.0 Å) with Ile33, Leu41, and Val47 of the hydrophobic core of the protein and thus shields them from contact with the solvent (Figure 5.5c). Its ε-amino group, which is involved in a hydrogen bond with the backbone of Thr40, is neither fully compensated by a group of opposite charge, nor is it solvated in the wildtype structures. The low accessibility may explain why the stabilization by the K65I mutation is independent of salt. In the stabilized variant Ile65 is located about 0.8 Å closer to the hydrophobic core than Lys65 in the wildtype protein. Lys65 also adopts alternate conformations in the two structures of the wildtype protein (now shown). Together, this suggests that part of the stabilization by the K65I mutation originates from a better packing of the Ile side chain with residues of the protein core (Figure 5.6c). In agreement with these results the replacement of residues with partially buried polar sidechains by hydrophobic residues have been shown to improve stability in other proteins [149, 150].

5.2.2 Stabilization in *Bs*-CspB A46K/S48R

In combination, the A46K and S48R mutations increase the stability of *Bs*-CspB by 11.1 kJ·mol^{-1}. Their contributions are approximately additive and of equal magnitude, and in both cases the major contributions to $\Delta\Delta G_D$ originate from screenable coulombic interactions (Table 5.3). The positions 46 and 48 are in the N-terminal region strand β4, which contacts the N-terminal part of strand β1 near Met1 and Glu3.

The repulsion between Glu3 and Glu66 (Figure 2a) provides a major unfavorable contribution to the stability of wildtype *Bs*-CspB [49, 91]. In the A46K/S48R mutant Lys46 and Arg48 are located on two sides of this pair of Glu residues (Figure 5.5, Figure 5.7), and thus they are optimally positioned for favorable interactions with these two negatively charged residues. Mutations at these four positions are strongly interdependent. This is highlighted by the finding that the E3R mutation stabilizes *Bs*-CspB by 16 kJ/mol when the positions 46, 48 and 66 are occupied by the wildtype resides, but destabilized by 6.2 kJ/mol in the presence of the mutations A46K, S48K and E66L [11].

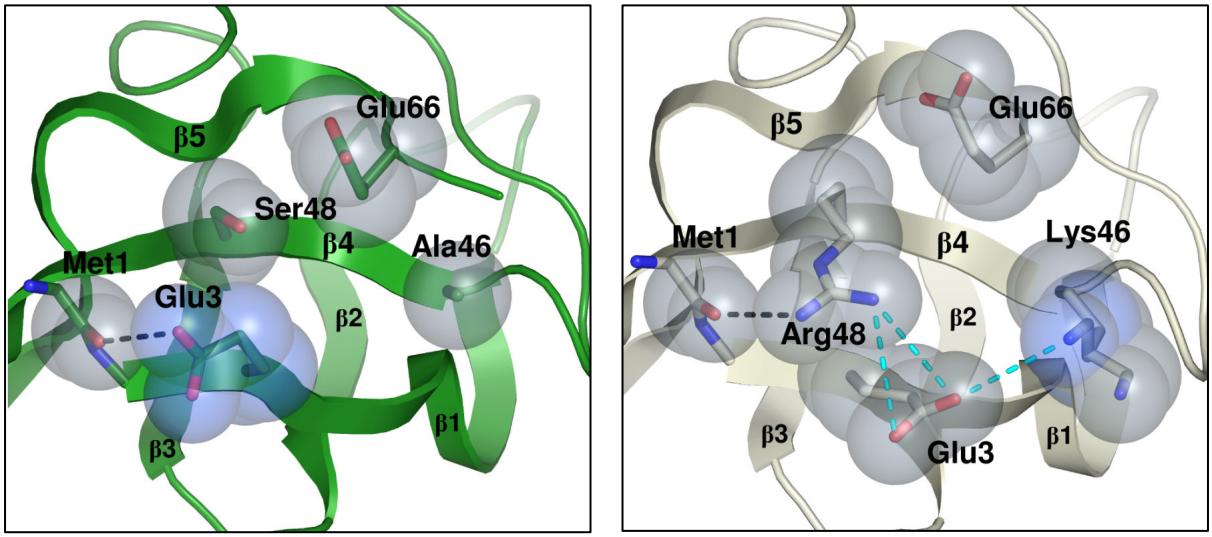


Figure 5.7: Stabilizing effect of the A46K/S48R double mutation. Relevant groups are depicted as sticks. Potential hydrogen bonds and salt bridges are depicted as black and cyan dashed lines, respectively. Carbon atoms of *Bs*-CspB (1CSP) and mutant variants are rendered in green (left images) and gray (right images), respectively. Van-der-Waals surfaces of selected groups are shown as semitransparent spheres. For groups added to the crystallographic models the van-der-Waals sphere is rendered in light blue.
The sidechain of Glu3 is absent in the atomic models of *Bs*-CspB structures. Its orientation derived from molecular dynamics simulation (blue spheres) suggests that it could form a hydrogen bond with Met1 in the 1CSP structure. Right: Arg48 is at a position where it can form two salt bridges with Glu3 and a hydrogen bond with Met1. Lys46 is involved in a hydrogen bond with a symmetry-related molecule (not shown). A slight reorientation of its terminal amino group (blue sphere) would allow the formation of a salt bridge / coulombic interaction with Glu3.

In the *Bs*-CspB wildtype structures, the sidechains of Ala46 and Ser48 are not involved in intramolecular interactions (Figure 5.7). In the double mutant, however, the side chain of Arg48 stretches across the β-sheet surface and, with its guanidino group, forms a hydrogen bond with the backbone carbonyl group of Met1 and electrostatically interacts with the carbxylgroup of Glu3. This additional interaction immobilizes the Glu3 carboxyl group, which is disordered in the structures of the wildtype protein.

The stabilizing contribution of Lys46 is not immediately evident from the structural analysis. In the crystal structure it forms a hydrogen bond with the carbonyl group of Gly44 from a symmetry-related molecule and located in some distance to Glu3 (5.3 Å). In order to form stronger electrostatic interaction ($d \leq 4.5$ Å) between these residues only a minor reorientation of the Lys46 sidechain torsion angles would be required (Figure 5.7).

5.2.3 Prevalence of stabilizing mutations from *Proside* selections in CSP representatives

Several proteins could be stabilized by consensus design [151]. In this strategy, residues are identified that diverge strongly from the consensus residues found in the alignment with a family of homologs. These diverging residues are then replaced by the consensus residue.

In order to evaluate if the mutations generated by *Proside in vitro* selections could have been identified by consensus design, the 250 nearest homologs of *Bs*-CspB were queried from the non-redundant Genbank database and aligned (Figure 5.8). The sequence identity with the most distant members was higher than 42% and the associated expectancy values were above 10^{-12}. Furthermore, residue positions mapped to residues in CSP structures which are part of the nucleic acid binding interface or the hydrophobic core and to glycines with torsion angles in the extended regions of the Ramachandran diagram, are conserved at a level above 75%. Only at two positions, in the center of loop 2 and in β-strand 5, the sequences of some members contain insertions of up to two amino acids.

This high overall conservation asserts that the tertiary structure is conserved for all members of this family, which is a prerequisite for discussing sequence variations on the basis of the three-dimensional structures of *Bs*-CspB and its close homologs.

At position 1, Met is the predominant amino acid, which is not surprising for bacterial proteins. Arg and Lys aligned to position 1 occur in 7.2 % and 8.8 % of the aligned sequences, and all of them are N-terminally extended with respect to *Bs*-CspB. In 52.1 % of all family members with an N-terminal extension, Arg or a Lys is observed at this position.

At position 3, Glu as in *Bs*-CspB is a rare amino acid (3.8 %) in the CSP sequences: Apart from *Bs*-CspB and the closely related *Bacillus licheniformis*, it is found in cold shock proteins from the cryophilic bacterium *Desulfotalea psychrophila*. An uncharged polar residue, Thr, is the predominant amino acid at this position (39.6 %). Lys3 (7.6 %) occurs in some mesophilic and in one thermophilic organism; Arg3 is infrequent (2.8 %) and found only in CSP from thermophilic organisms.

At position 46, Lys (42.8 %) and Arg (20.8 %) are the two most common residues, in agreement with the observations that these two residues were also found most frequently for position 46 in the *Proside* selections.

At position 48 Ser, as in wildtype *Bs*-CspB, is the predominant amino acid. Arg is very rare (0.8 %) and Lys is not observed at all. The frequent selection of these positively charged

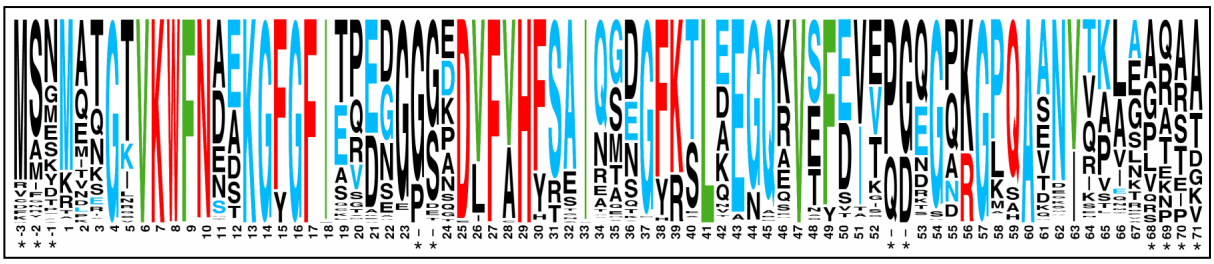


Figure 5.8: Frequency plot based on an alignment of 250 CSP sequences. Amino acids are depicted by their one-letter codes. Relative prevalences of residues at each position in the sequence alignment are indicated by different character heights. The positional numbering scheme (numbers below the plot) refers to positions in the *Bs*-CspB sequence, insertions and extensions are marked with asterisks (*). Residues occurring in *Bs*-CspB are highlighted by colors: Green - residues forming the hydrophobic core, red - surface residues involved in ligand binding, blue - other residues on the protein surface.

residues at position 48 of *Bs*-CspB and the strong stabilization by these residues is possibly related with the presence of Glu3, which is expected to form a coulombic interaction with Arg or Lys48. (Figure 5.7b). As mentioned previously, Glu3 occurs in only very few members of the Csp family.

At position 65, Lys occurs most frequently (35.6 %), although it destabilizes *Bs*-CspB. The next frequent residues are Ala (20.4 %) and Pro (18.0 %). Aliphatic residues such as Ile (as found in the selections) occur less frequently (10.4 %).

In summary, this analysis indicates that the strongly stabilizing potential of the E3K and the A46K mutations would also have been identified by a consensus analysis. The strongly stabilizing mutations M1R, S48R and K65I would have been missed.

5.3 Conclusions and closing remarks

The polypeptide chain of *Bs*-CspB folds into two β-sheets, β1-β2-β3 (β-sheet 1) and β4-β5 (β-sheet 2), which are separated by a long loop (L_{34} - residues 33 to 45). Both sheets assemble forming a β1-β2-β3-β5-β4-β1 barrel (see cartoon model in Figure 5.5a and Figure 4.10). Stabilizing polar interactions are not equally distributed in this assembly: Many interactions are observed within β-sheet 1 (1 – 2 salt bridges and about 20 hydrogen bonds) and 2 (17 – 19 hydrogen bonds), whereas only a limited number of interactions is observed between them (11 – 12 hydrogen bonds, numbers vary for different structures of wildtype CspB, see also Figure 4.11). A limiting factor for stability in CSP may be the extent of non-local interaction, as most of the stabilizing interactions described so far strengthen the contact between β1 and β4 (M1R, E3K: Figure 5.6a, A46K, S48R: Figure 5.7**,** *Bs*-CspB/*Bc*-Csp recombinants: [10]), which form one interface between the two β-sheets. This is quite intriguing, because φ-value

analyses of *Bs*-CspB [94] have indicated that β-strands 1 and 4 are already native-like and participate in stabilizing interactions at a very early stage of protein folding. Further stabilizing interactions are formed between residues from strands β1 and β5 (M1R), but none of the stabilizing mutations identified so far strengthen the interface between β-strands 3 and 5, which forms the second interface between the two β-sheets that closes the β barrel. These two β-strands are highly curved and their polar contacts are limited to a single backbone hydrogen bond, only. Two additional backbone hydrogen bonds are observed between ß-strand 3 and L3, which is expected to be flexible in solution [132]. β5 (residues 46 - 54) contains polar surface-based residues. In contrast, β-strand 3 (residues 25-30) contains three aromatic solvent-exposed sidechains (Phe 25, Phe27, His29) (Figure 5.8), which are part of the CSP ligand-binding platform (see Figure 3.15 & [132, 144]). Although this explains; why these residues are highly conserved in Nature (Figure 5.8), variants with polar residues at these positions which are engaged in stabilizing non-local interactions with residues from β-strand 5 should have been selected in *Proside*, if they result in an overall stability increase. As has been demonstrated by replacing His29 by glutamine, the introduction of polar residues into the hydrophobic platform leads to a significant destabilization of the protein (see Table 3.2), presumably due to unfavorable interactions with surrounding hydrophobic groups and a deteriorated packing of sidechains in the ligand-binding platform. This indicates that the second interface of the β-barrel is inadequate for generating further stabilizing interactions between the β -sheets.

It is noteworthy that apart from their role in CSP stabilization the two β-sheets are also involved in the structural reorganization of a *Bc*-Csp swapped dimer (see Chapter 4.2.2). In addition, a partitioning of the peptide chain into similar entities has been reported for amyloid fibers of *Ec*-CspA [93]. Further research is therefore advisable, to analyze the importance of these secondary structure elements in the folding and misfolding of the CSP.

Protein stabilization in CSP is achieved by increasing attractive non-local charge interactions (M1R, A46K, S48R) and by attenuating unfavorable non-local charge interactions (E3K, E3R [49], E66K [49]). Only the K65I mutation, located in β-sheet 2, contributes to a higher stability by removing local unfavorable charge interactions. Cooperativity was not observed for any combination of alterations, their stabilizing effects were additive at best [11]. Strong non-additivity is observed when the mutated side chains are in close contact, as observed for the mutations at the positions 3, 66, 46 and 48 [11]. The mutation K65I is largely independent of other mutations because it is spatially separated from them. For cooperative stabilization involving at least two mutations I would expect that the individual alterations would either be

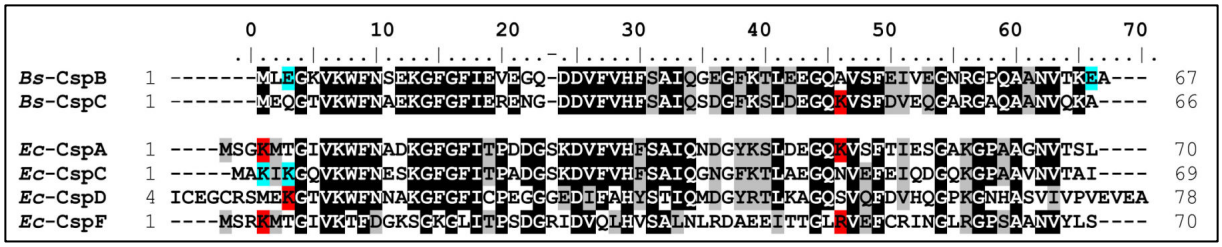


Figure 5.9: Prevalence of stabilizing residues in selected CSP paralogs from *Bacillus subtilis* (above) and *Escherichia coli* (below). Regions which are conserved at levels above 75% sequence identity and similarity within each species are shaded in black and gray, respectively. Residues which were shown to significantly increase or decrease stability in *Bs*-CspB and as well as analogous residues which are expected to be stabilizing or destabilizing are shaded in red and blue.

destabilizing or neutral and might therefore have been eliminated from the pool of mutants, due to the high selection pressure for stability in *Proside* selections. In Nature, neutral mutations may coexist and accumulate in the gene pool. Most of the alterations derived from *Proside* selections can be found in naturally CSP quite frequently, however. This may suggest that their selection in natural evolution is governed by similar principles. In addition, it is quite intriguing to see that the occurrence of stabilizing and destabilizing mutations and can vary considerably for CSP paralogs within one species (Figure 5.9), despite the fact that most other parts, including their nucleotide binding interface, are highly conserved (Figure 3.19). Although further experimental proof will be required, this could indicate that CSP paralogs which differ in stability, promote equivalent functions at different temperatures.