# Analysis of multiple crystal forms of *Bacillus subtilis* BacB suggests a role for a metal ion as a nucleant for crystallization

**M. Rajavel and B. Gopal***

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Correspondence e-mail:
bgopal@mbu.iisc.ernet.in

**PDB Reference:** BacB, triclinic form, 3h9a.

*Bacillus subtilis* BacB is an oxidase that is involved in the production of the antibiotic bacilysin. This protein contains two double-stranded $\beta$-helix (cupin) domains fused in a compact arrangement. BacB crystallizes in three crystal forms under similar crystallization conditions. An interesting observation was that a slight perturbation of the crystallization droplet resulted in the nucleation of a different crystal form. An X-ray absorption scan of BacB suggested the presence of cobalt and iron in the crystal. Here, a comparative analysis of the different crystal forms of BacB is presented in an effort to identify the basis for the different lattices. It is noted that metal ions mediating interactions across the asymmetric unit dominate the different packing arrangements. Furthermore, a normalized $B$-factor analysis of all the crystal structures suggests that the solvent-exposed metal ions decrease the flexibility of a loop segment, perhaps influencing the choice of crystal form. The residues coordinating the surface metal ion are similar in the triclinic and monoclinic crystal forms. The coordinating ligands for the corresponding metal ion in the tetragonal crystal form are different, leading to a tighter packing arrangement. Although BacB is a monomer in solution, a dimer of BacB serves as a template on which higher order symmetrical arrangements are formed. The different crystal forms of BacB thus provide experimental evidence for metal-ion-mediated lattice formation and crystal packing.

## 1. Introduction

The ability to crystallize a protein in multiple crystal forms is often an important factor in protein structure determination. Protein crystals are formed through weak and nearly stochastic packing interactions and most lattice contacts involve polar amino acids rather than hydrophobic interactions and covalent disulfide-bond formation (Dasgupta *et al.*, 1997). This generalization is consistent with the observation that large surface-exposed hydrophobic patches or free cysteines located on the surface lead to oligomerization or aggregation at high concentrations rather than well formed protein crystals (Qiu & Janson, 2004). Lattice engineering is therefore geared to optimize surface properties by using shorter constructs of a given protein, reducing flexible loops and either modifying or mutating flexible surface residues (Lys, Gln or Glu to Ala; Derewenda, 2004). Hydrogen bonds and interactions mediated by divalent metal ions, on the other hand, substantially influence protein crystal formation (Durbin & Feher, 1996). The role of these so-called surface metal ions is well documented, with examples that include the oligomerization of human S100A12 in the presence of $Ca^{2+}$ and $Zn^{2+}$ (Moroz *et al.*, 2009) and of $\alpha$-synuclein in the presence of $Cu^{2+}$ (Paik *et al.*, 1999) and different symmetrical assemblies in MBPC-1 using $Cu^{2+}$ and $Ni^{2+}$ (Salgado *et al.*, 2009).

The systematic engineering of metal-binding sites on the surface of a protein molecule to enable metal-ion-mediated crystal lattice contacts has yielded mixed results. An example of a successful application of this strategy is the case of human ferritin. Human ferritin shares high sequence similarity to rat ferritin, a protein which required $Cd^{2+}$ to form well diffracting crystals. The K86Q mutant of human ferritin was designed to restore the $Cd^{2+}$ site, leading to the successful crystallization of this protein (Lawson *et al.*, 1991). This finding is similar to that for the *Escherichia coli* acyl-carrier protein,

in which the introduction of Glu or Asp residues on the surface led to the crystallization of the protein in the presence of $Zn^{2+}$ (Qiu & Janson, 2004). Another interesting example is that of *Phormidium laminosum* plastocyanin. A condition containing 0.1 *M* sodium cacodylate led to protein crystals in a monoclinic space group, whereas the addition of 0.2 *M* $MgCl_2$ resulted in an orthorhombic lattice and the replacement of $Mg^{2+}$ by $Zn^{2+}$ yielded tetragonal form crystals. This observation led to the hypothesis that different crystal-packing arrangements could be compared with cellular environments and long-range electron-transport and electron-tunnelling mechanisms

and can thus be explained by a model of supramolecular metal-ion-mediated assemblies (Crowley *et al.*, 2008).

Here, we describe the crystal-packing arrangement of three crystal forms of BacB. BacB is an oxidase that is involved in the synthesis of bacilysin (Rajavel *et al.*, 2009). A comparison between several divalent cations revealed that $Co^{2+}$ and $Fe^{2+}$ were the most suitable metal cofactors for catalytic activity of this enzyme. An interesting observation based on the crystal structure was on the role(s) of the two cupin domains. The N-terminal domain of BacB is involved in catalysis, whereas the C-terminal cupin domain with a bound phenyl
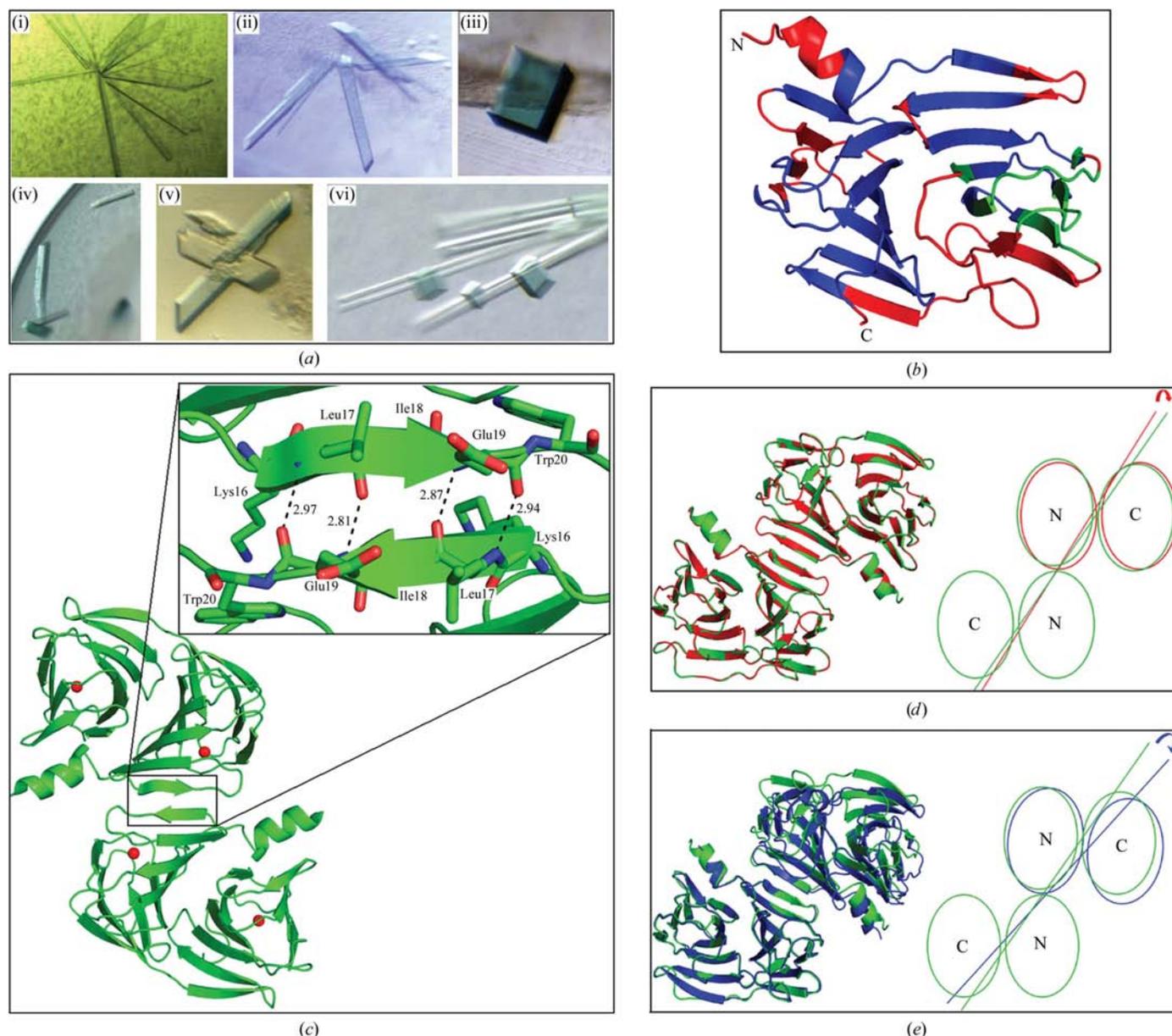


**Figure 1**
(*a*) Different crystal forms of BacB. (i) Plate-like crystals of BacB. (ii) Crystals in the monoclinic space group. (iii) Tetragonal form crystals. (iv) Triclinic form crystals. (v) Nucleation of monoclinic BacB crystals in a drop containing monoclinic form crystals. (vi) Nucleation of tetragonal form crystals in a drop containing monoclinic form crystals. (*b*) *Error-inclusive Structure Comparison and Evaluation Tool* (*ESCET*) analysis of six chains (two per crystal form) of BacB. By limiting the error-scaling limits between $2.0\sigma$ and $5.0\sigma$, approximately 63% of the structure was seen to be conformationally invariant. A colour-coded representation of the conformationally invariant (blue), moderately flexible (green) and flexible regions (red) of the BacB structure is shown (Schneider, 2002). (*c*) Inter-subunit contacts through $\beta$-strands. The two subunits of BacB in the asymmetric unit are shown with a magnified view of the interacting backbone $\beta$-strand residues (Lys16, Leu17, Ile18, Glu19 and Trp20; figure drawn using *PyMOL*; DeLano, 2002). (*d*) Superposition of the *A* and *B* monomers of the triclinic form of BacB on the monoclinic form. Subunit *A* of the triclinic form was superposed on subunit *A* of the monoclinic form with an r.m.s.d. of 0.4 Å. A schematic representation of the superposition of the triclinic form on the monoclinic form of BacB is also shown. The difference in the orientation of the *B* subunit is approximately $1.8°$. (*e*) Superposition of the *A* and *B* subunits of the tetragonal form of BacB on the monoclinic form (r.m.s.d. of 0.6 Å over 225 $C^\alpha$ atoms). A schematic representation of the superposition of the tetragonal form on the monoclinic form of BacB is also shown. The difference in the orientation of the *B* subunit is approximately $16.4°$. These calculations were performed using *ALIGN* (Cohen, 1997).

**Table 1**
Summary of crystallization conditions and crystal forms of BacB.

| PDB code | Space group | Matthews coefficient ($\text{Å}^3\,\text{Da}^{-1}$) | Solvent content (%) | Crystallization conditions |
|---|---|---|---|---|
| 3h9a | $P1$ | 2.41 | 49.06 | 0.1 $M$ Tris pH 6.8, 0.2 $M$ NaCl, 10%($w/v$) PEG 8K, 45%($v/v$) MPD (SeMet BacB) |
| 3h7j | $P2_1$ | 2.36 | 47.79 | 0.1 $M$ Tris pH 6.8, 0.2 $M$ NaCl, 10%($w/v$) PEG 8K, 45%($v/v$) MPD |
| 3h7y | $P4_12_12$ | 2.32 | 46.96 | 0.1 $M$ Tris pH 6.8, 0.2 $M$ NaCl, 10%($w/v$) PEG 8K, 60%($v/v$) MPD |

pyruvate plays a regulatory function. Thus, the observation that BacB crystals contain either $Co^{2+}$ or $Fe^{2+}$ is consistent with the function of this enzyme. Structural analysis of the different crystal forms of BacB described here shows that metal ions bind to the active site as well as the surface of this protein. This analysis further suggests that lattice formation involves surface metal-ion-mediated interactions.

## 2. Materials and methods

### 2.1. Cloning, expression, purification and crystallization of BacB

The purification and crystallization of *Bacillus subtilis* BacB (also referred to as YwfC) has been described previously (Rajavel & Gopal, 2006). Briefly, *E. coli* BL21 (DE3) pLysS cells expressing recombinant BacB were grown in Luria broth. The recombinant BacB protein with a polyhistidine tag at the C-terminus was partially purified using cobalt–NTA resin (Talon, Clontech). The protein was further purified by size-exclusion chromatography using a Sephacryl S-200 column (Amersham Pharmacia). BacB could be crystallized in three different crystal forms by slight modification of the crystallization conditions (Fig. 1a). The concentration of 2,4-methylpentanediol (MPD) appeared to substantially influence crystallization. Thus, plate-like BacB crystals were obtained using a 5%($v/v$) MPD concentration, whereas crystals belonging to monoclinic and tetragonal space groups were obtained on increasing the MPD concentration to 45%($v/v$) and 60%($v/v$), respectively (Fig. 1a). The selenomethionine derivative of BacB crystallized in the triclinic space group. A summary of the crystallization conditions, space groups and solvent contents is provided in Table 1.

### 2.2. Data collection and structure determination of BacB

Diffraction data were collected from flash-frozen BacB crystals using 20% PEG 400 as the cryoprotectant. While monoclinic form crystals diffracted to beyond 1.9 Å resolution, crystals of the tetragonal form diffracted to 2.2 Å resolution at the home source (a MAR imaging-plate detector mounted on a Bruker Microstar Ultra generator with Helios optics). The triclinic crystal form diffracted to 2.0 Å resolution (Table 2). Diffraction data were processed using *MOSFLM* (Leslie, 1999) and scaled using *SCALA* (Collaborative Computational Project, Number 4, 1994). The crystal structure of BacB was determined by the single-wavelength anomalous dispersion (SAD) technique using data collected at the home source (Rajavel *et al.*, 2009). A bound $Co^{2+}$ ion at the active site of BacB served as the anomalous scatterer. The X-ray absorption spectra for BacB crystals suggested the presence of either $Co^{2+}$ or $Fe^{2+}$ (Supplementary Fig. 1[1]).

---

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: MV5031). Services for accessing this material are described at the back of the journal.

**Table 2**
Data-collection, phasing and refinement statistics of the triclinic crystal form of BacB.

Values in parentheses are for the outermost shell.

| Data collection | |
|---|---|
| Wavelength (Å) | 1.5418 |
| Resolution (Å) | 30.67–2.03 (2.14–2.03) |
| Unit-cell parameters (Å, °) | $a = 46.2$, $b = 47.2$, $c = 62.9$, $\alpha = 89.05$, $\beta = 77.27$, $\gamma = 82.28$ |
| Space group | $P1$ |
| Total No. of reflections | 147869 (15924) |
| No. of unique reflections | 30450 (3593) |
| Completeness (%) | 91.4 (73.8) |
| Multiplicity | 2.2 (2.1) |
| $R_{\text{merge}}$† (%) | 6.7 (24.3) |
| $\langle I/\sigma(I) \rangle$ | 8.5 (2.8) |
| **Phasing (molecular replacement with tetragonal form of BacB; PDB code 3h7y)** | |
| *Phaser* statistics | |
| $Z$ score | 40.36 |
| Final LLG | 2178.87 |
| Refinement statistics | |
| $R_{\text{cryst}}$‡ (%) | 26.2 |
| $R_{\text{free}}$§ (%) | 20.6 |
| R.m.s.d. bond lengths (Å) | 0.005 |
| R.m.s.d. angles (°) | 0.96 |

† $R_{\text{merge}} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the $i$th reflection and $\langle I(hkl)\rangle$ is the average intensity. ‡ $R_{\text{cryst}} = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$. § $R_{\text{free}}$ was calculated as for $R_{\text{cryst}}$ but with 5% of the data that were excluded from refinement calculations.

The presence of $Co^{2+}$ and $Fe^{2+}$ metal ions in BacB samples was also confirmed by inductively coupled plasma atomic emission spectroscopy (ICP-AES) analysis (Supplementary Table 1[1]). The location(s) of the anomalous scatterers were determined using *PHENIX* (Adams *et al.*, 2002). The structure of BacB was solved using this *de novo* phase information, with subsequent iterations of model building and refinement using *Coot* (Emsley & Cowtan, 2004), *ARP/wARP* (Lamzin & Wilson, 1993) and *REFMAC*5 (Collaborative Computational Project, Number 4, 1994). The structures of BacB in the monoclinic and triclinic space groups (PDB codes 3h7j and 3h9a) were determined by molecular replacement using the tetragonal BacB structure (PDB code 3h7y). The crystal structures of BacB solved in the monoclinic and tetragonal crystal forms have been reported previously (Rajavel *et al.*, 2009). This manuscript describes the crystal structure of the triclinic form and its comparison with the other BacB structures.

## 3. Results and discussion

### 3.1. The structure of and subunit interactions in BacB

BacB is a bicupin protein with two $\beta$-barrel domains fused together in a tightly packed arrangement. The N- and C-terminal cupin domains are remarkably similar, with an r.m.s.d. of 0.95 Å over 95 overlapping $C^{\alpha}$ atoms. Fig. 1(b) highlights the conformationally invariant, moderately flexible and flexible regions of BacB. The N-terminal cupin domain has higher $B$ factors, especially for the residues that coordinate the active-site metal ion (45–55). An *Error-inclusive Structure Comparison and Evaluation Tool* (*ESCET*) analysis further suggested that the C-terminal domain was more rigid (Schneider, 2002). Biochemical studies on BacB and a point mutant revealed that the N- and C-terminal domains, although similar in structure, perform different functional roles. The N-terminal domain was shown to be involved in catalysis, while the presence of a bound phenyl pyruvate (PPY) in the C-terminal cupin domain suggested that this domain was probably involved in feedback regulation of catalytic activity (Rajavel *et al.*, 2009).

The crystal structures of BacB solved in the triclinic and tetragonal forms were superposed on that of the monoclinic form. The peptide backbone of residues 16–20 located on a β-strand was seen to be involved in contacts across asymmetric units (Fig. 1c). An additional salt bridge involving Arg74–Asp124 was seen in the structure of BacB solved in the tetragonal space group (Supplementary Fig. 1). The A chain of the triclinic form superposes on the corresponding chain from the monoclinic form with an r.m.s.d. of 0.4 Å (r.m.s.d. of 0.6 Å with the tetragonal form; Figs. 1d and 1e). Superposing the A chain also revealed variations in the location of the noncrystallographic symmetry-related molecule. Chain B is oriented differently, with a 16.4° tilt between the monoclinic and tetragonal forms, whereas this packing in the monoclinic form is closer to the triclinic arrangement with a tilt of 1.8°. An interesting aspect to note in this context is that when crystallization drops containing the monoclinic form crystals

were disturbed (for example, to harvest crystals for data collection), tetragonal form crystals appeared in the same drop within 2–3 d. Mechanical disturbance was not always necessary to obtain crystals in the tetragonal space group. In a few instances, crystals belonging to the tetragonal form appeared in the same drop several days after the initial batch of monoclinic form crystals appeared or upon increasing the MPD concentration to 60%(v/v).

### 3.2. Surface metal ions and metal-ion-mediated interactions in BacB crystals

Additional metal ions to the metal ions at the active site(s) could be modelled in the electron-density maps of BacB in all crystal forms (Figs. 2a and 2b). These metal ions are located on the surface of the protein. It is likely that these metal ions were incorporated either
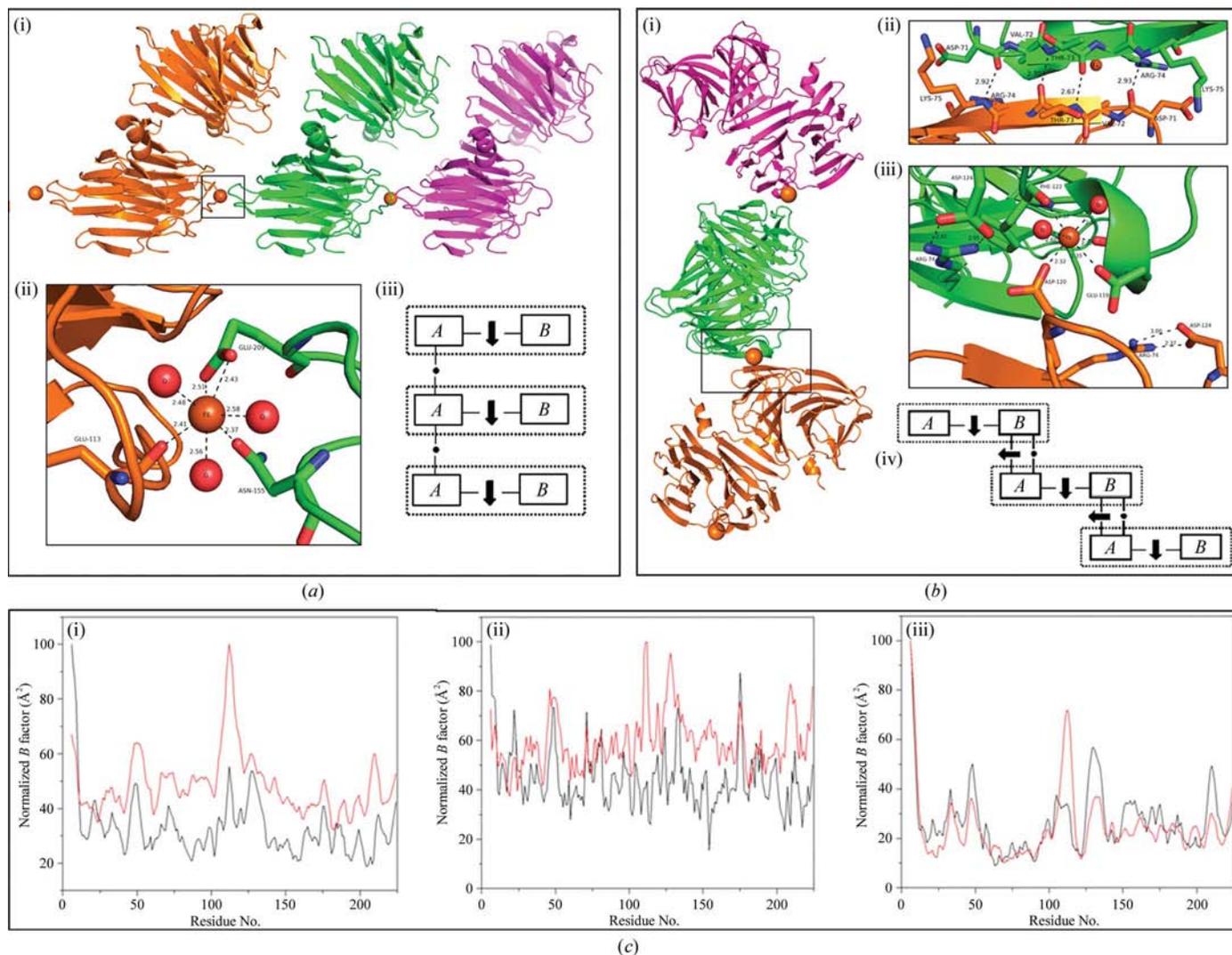


**Figure 2**
(a) A schematic representation of the packing in the triclinic and monoclinic crystal forms of BacB. (i) Packing of BacB in either the triclinic or monoclinic lattice. (ii) The metal ion involved in inter-asymmetric unit interactions is coordinated by Asn155, Glu209 and Glu113 of a symmetry-related molecule. (iii) A schematic representation of crystal packing in the triclinic and monoclinic forms of BacB. Closed boxes containing A and B denote two monomeric units, whereas the dotted box depicts the asymmetric unit. Interactions between noncrystallographic symmetry-related monomers are through β-strands, whereas the inter-asymmetric unit contacts are mediated by the surface metal ions. Only the A chain of adjacent subunits is involved in this packing arrangement. (b) The tetragonal form of BacB. (i) BacB in a tetragonal lattice with its symmetry-related molecules. (ii) Stacking of β-strands between symmetry-related molecules. (iii) The metal ion involved in inter-asymmetric unit contact is coordinated by Glu119 and Phe122 of one monomer and Asp120 from a symmetry-related molecule. (iv) A schematic representation of the packing in the tetragonal form of BacB. Closed boxes containing A and B denote two monomers and the dotted box depicts the asymmetric unit. (c) Temperature-factor plots corresponding to BacB structures solved in the triclinic, monoclinic and tetragonal crystal forms. (i) Chains A (black) and B (red) of the triclinic form of BacB. (ii) The presence of the second metal ion (coordinated by Asp33, Glu167 and Glu197) is reflected in a plot of temperature factors for residues in BacB (the monoclinic form). (iii) In the tetragonal form, chain A and B crystal contacts mediated through a β-strand as well as a surface metal ion are reflected by the lower temperature factors for Glu119, Asp120 and Phe122.

during the expression of recombinant BacB in *E. coli* or from the purification buffers. Interestingly, the average coordination distance for these surface metal ligands was around 2.4 Å, with a distorted octahedral coordination geometry. The residues that coordinate the metal ion are the same in the triclinic and monoclinic forms but are different in the tetragonal form. As the data were not specifically collected at the Fe or Co edges for the three data sets examined here, anomalous difference Fourier maps were not conclusive in establishing that the surface metal ions were $Fe^{2+}$ as opposed to the $Co^{2+}$ ions at the active site. A comparison of the coordination distances and geometry with those reported by Dokmanić and coworkers, however, suggests that the surface metal ions in BacB are more likely to be $Fe^{2+}$ (Dokmanić *et al.*, 2008). The coordination distances of a metal ion with different functional groups (main chain, side chain and solvent molecules) compiled by these authors were used as a reference in this analysis. In particular, we note that the interaction distances of $Fe^{2+}$ with the side chains of Asp or Glu (reported as 2.45 Å) are in close agreement with those of the surface metal ions of BacB (Supplementary Table 2).

All crystal forms of BacB contain two protein monomers in the asymmetric unit. In the triclinic and monoclinic forms the interaction between symmetry-related molecules is mediated by a metal ion bound on the surface (Fig. 2*a*). The role of the surface metal ions in mediating contacts between BacB molecules is thus consistent with the observations from dynamic light-scattering (DLS) measurements that suggest protein aggregation in the presence of $Co^{2+}$ or $Fe^{2+}$ ions (Supplementary Table 3). The coordinating residues of this surface metal ion are Asn155 and Glu209 of one monomer and Glu113 of a symmetry-related molecule (Fig. 2*a*). Of the two surface metal ions that could be modelled in the electron-density map of the monoclinic form of BacB, only one of them is involved in inter-asymmetric unit contacts. The second metal ion is coordinated by residues Asp33, Glu167 and Glu197 of the same polypeptide chain and does not participate in inter-asymmetric unit contacts. A schematic representation of crystal packing in the triclinic and monoclinic forms of BacB is shown in Fig. 2(*a*). The residues coordinating the surface metal ion in the tetragonal form of BacB are different in the triclinic and monoclinic crystal forms. Glu119 and Phe122 of one monomer and Asp120 of a symmetry-related molecule coordinate this metal ion. In this case, the backbone of the $\beta$-strand comprising residues 71–75 is also involved in hydrogen bonding to its symmetry-related counterpart, leading to a compact packing arrangement of molecules with higher symmetry (Fig. 2*b*). The conformational features that contribute to intramolecular and intermolecular packing are consistent with the variation in the temperature factor for these residues when compared with the rest of the protein (Fig. 2*c*).

In the triclinic and monoclinic crystal forms, contacts between monomeric subunits are primarily mediated by the surface metal ion. This comparison between crystal forms thus suggests that the presence of surface metal ions and the absence of a third salt bridge leads to a loop (residues 114–126) remaining flexible in the monoclinic

and triclinic crystal forms. This feature is more apparent from the observation that packing BacB structures solved in either the monoclinic or triclinic crystal forms into the tetragonal cell results in steric clashes involving this loop segment (Supplementary Fig. 2). We further note that this mode of crystal packing is also adopted by other cupin proteins (Supplementary Fig. 3 and Table 4).

## 4. Conclusions

The differences in the packing between the three crystal forms of BacB can be rationalized on the basis of metal ions bound on the protein surface that mediate lattice contacts. Along one dimension, the asymmetric unit contacts in all crystal forms are mediated by a surface metal ion. In crystals belonging to the tetragonal lattice, the surface metal ion and an additional salt bridge lead to a more compact packing arrangement which brings the interacting $\beta$-strands closer than in the other crystal forms. This comparative analysis of different crystal forms of BacB thus provides a template for the design of crystallization experiments using metal-ion cofactors or the engineering of metal-binding sites on the surface to enhance lattice contacts.

## References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). *Acta Cryst.* D**58**, 1948–1954.
Cohen, G. E. (1997). *J. Appl. Cryst.* **30**, 1160–1161.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Crowley, P. B., Matias, P. M., Mi, H., Firbank, S. J., Banfield, M. J. & Dennison, C. (2008). *Biochemistry*, **47**, 6583–6589.
Dasgupta, S., Iyer, G. H., Bryant, S. H., Lawrence, C. E. & Bell, J. A. (1997). *Proteins*, **28**, 494–514.
DeLano, W. L. (2002). *The PyMOL Molecular Viewer.* DeLano Scientific, San Carlos, California, USA. http://www.pymol.org.
Derewenda, Z. S. (2004). *Structure*, **12**, 529–535.
Dokmanić, I., Šikić, M. & Tomić, S. (2008). *Acta Cryst.* D**64**, 257–263.
Durbin, S. D. & Feher, G. (1996). *Annu. Rev. Phys. Chem.* **47**, 171–204.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.
Lawson, D. M., Artymiuk, P. J., Yewdall, S. J., Smith, J. M., Livingstone, J. C., Treffry, A., Luzzago, A., Levi, S., Arosio, P. & Cesareni, G. (1991). *Nature (London)*, **349**, 541–544.
Leslie, A. G. W. (1999). *Acta Cryst.* D**55**, 1696–1702.
Moroz, O. V., Burkitt, W., Wittkowski, H., He, W., Ianoul, A., Novitskaya, V., Xie, J., Polyakova, O., Lednev, I. K., Shekhtman, A., Derrick, P. J., Bjoerk, P., Foell, D. & Bronstein, I. B. (2009). *BMC Biochem.* **10**, 11.
Paik, S. R., Shin, H.-J., Lee, J.-H., Chang, C.-S. & Kim, J. (1999). *Biochem. J.* **340**, 821–828.
Qiu, X. & Janson, C. A. (2004). *Acta Cryst.* D**60**, 1545–1554.
Rajavel, M. & Gopal, B. (2006). *Acta Cryst.* F**62**, 1259–1262.
Rajavel, M., Mitra, A. & Gopal, B. (2009). *J. Biol. Chem.* **284**, 31882–31892.
Salgado, E. N., Lewis, R. A., Mossin, S., Rheingold, A. L. & Tezcan, F. A. (2009). *Inorg. Chem.* **48**, 2726–2728.
Schneider, T. R. (2002). *Acta Cryst.* D**58**, 195–208.