Acta Crystallographica Section F
**Structural Biology and Crystallization Communications**
ISSN 1744-3091

# How to get the magic triangle and the MAD triangle into your protein crystal

**Tobias Beck,* Carlos Eduardo da Cunha and George M. Sheldrick**

Department of Structural Chemistry, Georg-August-Universität Göttingen, Tammannstrasse 4, 37077 Göttingen, Germany

Correspondence e-mail: tbeck@shelx.uni-ac.gwdg.de

Received 9 June 2009
Accepted 11 September 2009

**PDB Reference:** bovine trypsin crystallized with B3C, 3iti, r3itisf.

The magic triangle 5-amino-2,4,6-triiodoisophthalic acid (I3C) and the MAD triangle 5-amino-2,4,6-tribromoisophthalic acid (B3C) are two representatives of a novel class of compounds that combine heavy atoms for experimental phasing with functional groups for protein interactions. These compounds are readily available and provide easy access to experimental phasing. The preparation of stock solutions and the incorporation of the compounds into protein crystals are discussed. As an example of incorporation *via* cocrystallization, the incorporation of B3C into bovine trypsin, resulting in a single site with high occupancy, is described.

## 1. Introduction

Obtaining a suitable heavy-atom derivative is still a challenge for experimental phasing. We have developed a new class of compounds that combine heavy atoms for anomalous phasing with functional groups for interaction with macromolecules. The 'magic triangle' 5-amino-2,4,6-triiodoisophthalic acid (I3C; Beck & Sheldrick, 2008) may be used for experimental phase determination *via* single-wavelength anomalous dispersion (SAD) or single isomorphous replacement plus anomalous scattering (SIRAS). I3C is commercially available as a powder from different chemical suppliers, but can also be acquired as part of a ready-to-use kit. The 'MAD triangle' 5-amino-2,4,6-tribromoisophthalic acid (B3C; Beck *et al.*, 2009) may be utilized for multi-wavelength anomalous dispersion (MAD) experiments. Limited amounts of B3C may be obtained by contacting TB. The compounds I3C and B3C are two representatives of this class of novel phasing tools. Here, we describe some practical aspects of the incorporation of these small molecules into protein crystals.

## 2. Incorporation of the compounds into protein crystals

### 2.1. Stock solutions

A stock solution with a high concentration of I3C or B3C is desirable, especially for soaking experiments. To obtain a high concentration in the stock solution, the compounds cannot be dissolved directly in water. The two carboxylate groups are deprotonated with double equimolar amounts of aqueous base to obtain a salt solution (Fig. 1). In case of I3C, only the use of lithium hydroxide solution as a base produces an I3C salt with high solubility. If sodium or potassium hydroxide solution or bases such as triethylamine or ammonia are used, the resulting salt has limited solubility. Lithium cations as counterions have a high hydration energy, with the result that the

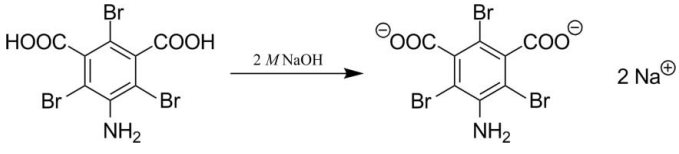
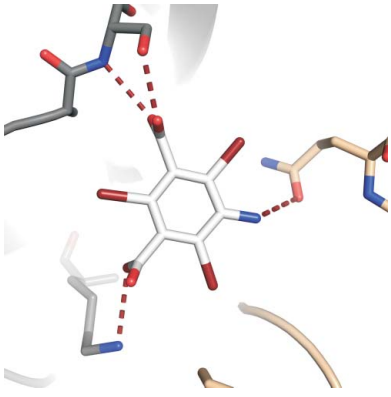



**Figure 1**
Dissolving B3C in 2 *M* NaOH solution. Double equimolar amounts of base are required to deprotonate both carboxyl groups. The resulting sodium B3C salt has a high solubility.



solvation energy exceeds the lattice energy of the salt formed with the doubly charged anion of I3C, so the salt is highly soluble. In the case of B3C, sodium hydroxide solution gives satisfactory results (Fig. 1), possibly because the smaller double anion B3C has a higher solvation energy. 1 *M* sodium B3C solution can be obtained by dissolving 0.043 g B3C in 0.1 ml 2 *M* NaOH solution. The lithium I3C or sodium B3C stock solutions can be used for protein-crystal derivatization as described below.

### 2.2. Cocrystallization

A way to incorporate the phasing tools into biological macromolecules is to add the small molecules to the crystallization drop. During crystal growth, the compounds are integrated into the crystal lattice. The compounds may be added to the protein solution before crystallization setup or directly into the already pipetted drops. It is generally advised to use at least a tenfold excess of I3C or B3C for incorporation *via* cocrystallization.

### 2.3. An example of incorporation *via* cocrystallization: B3C in bovine trypsin

**2.3.1. Crystallization and data collection**. Here, we report the incorporation of B3C into bovine trypsin by cocrystallization. Bovine trypsin (223 residues, 23.3 kDa; Sigma–Aldrich) was dissolved to a final concentration of 60 mg ml$^{-1}$ in a solution containing 10 mg ml$^{-1}$ benzamidine (inhibitor to prevent auto-cleavage), 3 m*M* $CaCl_2$, 0.02% $NaN_3$ and 30 m*M* sodium B3C (about a tenfold excess of B3C with respect to the protein concentration). Crystals were grown using the sitting-drop vapour-diffusion method. Wizard Screen I (Emerald BioSystems) was set up with a Tecan dispenser (100 µl reservoir volume). 0.1 µl protein solution was mixed with an equal volume of reservoir solution using a Mosquito robot. Crystals appeared within one week in several conditions. Data were collected in-house at 100 K from a crystal taken from a drop with the reservoir solution containing 30%(*w*/*v*) PEG 3000 and 0.1 *M* CHES pH 9.5. Prior to data collection, the crystal was flash-cooled in liquid nitrogen without any additional cryoprotection.

**2.3.2. Data analysis, structure solution and refinement**. Although the anomalous signal of Br is relatively weak at Cu *Kα* ($f'' = 1.25$ e; Fig. 2 shows the anomalous signal for this data set), the ratio of $R_{\rm anom}$ to $R_{\rm p.i.m.}$, a measure of the quality of the anomalous signal (Weiss, 2001), indicates that structure solution should be possible using in-house data (Table 1). The presence of a calcium ion that also contributes to the anomalous signal at this wavelength ($f''$ about the same as for Br) is probably also helpful for structure solution. The heavy-atom substructure solution was successfully solved with *SHELXD* (Schneider & Sheldrick, 2002) using the single-wavelength anomalous dispersion (SAD) method [CC = 18.31; CC(weak) = 9.97]. The first four peaks found with *SHELXD* are the Ca atom and three peaks forming an equilateral triangle with a side of about 5.6 Å, *i.e.* the Br atoms in B3C. *SHELXE* (Sheldrick, 2002) was used for phase extension and density modification. The new beta test version also includes a protein main-chain tracing algorithm (Sheldrick, 2010). 218 of 223 residues were traced, belonging to five different chains. The experimental electron-density map obtained from *SHELXE* shows a high correlation with the final refined map (CC = 0.892) and a low mean phase error (26.8°). Although experimental phasing with B3C at a single wavelength was successful in this example, the use of B3C in multi-wavelength anomalous dispersion (MAD) experiments is recommended in order to take advantage of the extra phase information available from a MAD experiment.

**Table 1**
Data statistics for bovine trypsin with B3C.

Values in parentheses are for the highest resolution shell.

| | |
|---|---|
| Unit-cell parameters (Å, °) | $a = 53.66$, $b = 56.88$, $c = 66.81$ |
| Space group | $P2_12_12_1$ |
| Wavelength (Å) | 1.54178 |
| Resolution (Å) | 41.8–1.55 (1.65–1.55) |
| $R_{\rm merge}$† | 0.0658 (0.288) |
| Completeness | 99.9 (99.7) |
| Multiplicity | 7.69 (3.61) |
| $\langle I/\sigma(I)\rangle$ | 18.9 (3.87) |
| $R_{\rm anom}$‡/$R_{\rm p.i.m.}$§ | 2.81 (2.20) |

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.
‡ $R_{\rm anom} = \sum_{hkl} 2|I(hkl) - I(\overline{hkl})| / \sum_{hkl}[I(hkl) + I(\overline{hkl})]$.
§ $R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

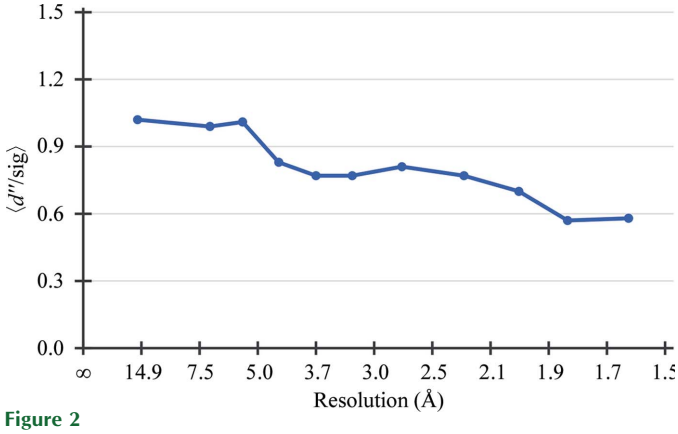



**Figure 2**
Anomalous data statistics for B3C in bovine trypsin. Three Br atoms (B3C) and one Ca atom contribute to the anomalous signal. Data were truncated to 1.9 Å for heavy-atom substructure solution. Although the anomalous signal is rather weak, heavy-atom substructure solution and phase extension could successfully be carried out. This figure was prepared using *HKL2MAP* (Pape & Schneider, 2004).

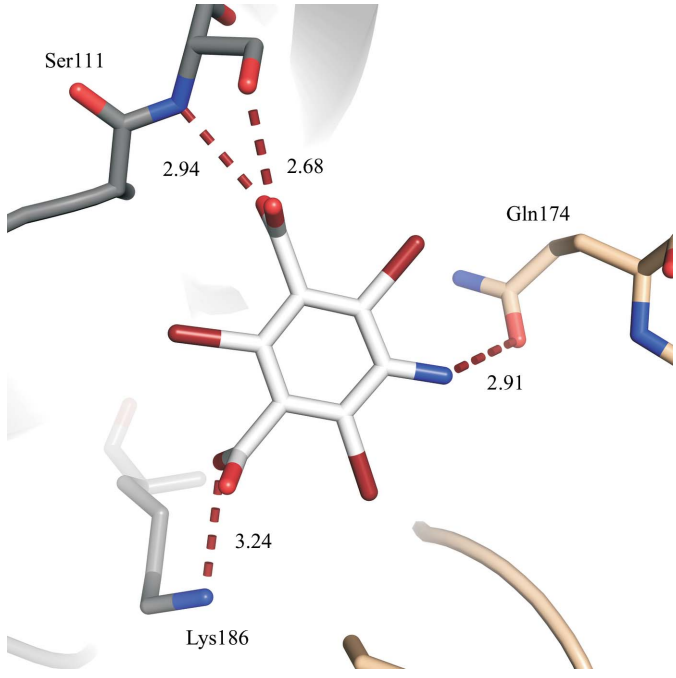



**Figure 3**
The B3C site in bovine trypsin. Hydrogen bonds to Lys186, Ser111 (including the interaction with the amide proton) and Gln174 are shown as dashed lines; distances are given in Å. This figure was prepared using *PyMOL* (DeLano, 2002).

**Table 2**
Refinement details for bovine trypsin with B3C.

| | |
|---|---|
| PDB code | 3iti |
| No. of reflections | 30344 |
| $R_{cryst}/R_{free}$ (%) | 16.7/21.3 |
| No. of protein atoms | 1615 |
| No. of ligand/ion atoms | 27 |
| No. of water atoms | 243 |
| *B* factors (Å$^2$) | |
| Protein | 13.51 |
| Ligands | 14.30 |
| Waters | 28.30 |
| R.m.s. deviations | |
| Bond length (Å) | 0.009 |
| Angle distance (Å) | 0.024 |
| Ramachandran plot | |
| Favoured (%) | 98.64 |
| Disallowed (%) | 0 |

The initial backbone model from *SHELXE* was first manually completed with *Coot* (Emsley & Cowtan, 2004). The side chains were docked within *Coot* using the 'dock sequence' extension. Refinement was carried out with *SHELXL* (Sheldrick, 2008). Refinement statistics are given in Table 2. Stereochemical analysis of the refined structures was performed with the *MolProbity* server (Davis *et al.*, 2007). There is a single site for B3C in the crystal structure. Its occupancy was refined to 0.88 for the scaffold atoms of the phenyl ring (carbon), the functional groups (oxygen, nitrogen and carbon) and two Br atoms. The third Br atom only has an occupancy of 0.73 (the occupancies of the Br atoms were refined independently). This may be rationalized by specific radiation damage at this bromine site (even with in-house data). B3C interacts with two molecules of trypsin and shows similar hydrogen-bond interactions to those observed in other protein structures (Fig. 3).

### 2.4. Soaking

The introduction of I3C or B3C into protein crystals by soaking has been described elsewhere (Beck *et al.*, 2008, 2010; Sippel *et al.*, 2008). For soaking experiments, the use of a high concentration of B3C or I3C (*e.g.* 500 or 200 m*M*) in conjunction with short soaking times (*e.g.* 10 s) is generally recommended if this is tolerated by the crystal. If crystal degradation occurs, lower concentrations together with longer soaking times are advised. A gradient soak in which the concentration of the ligand is increased stepwise (*i.e.* decreasing the disturbance at each step for the crystal) can also be a valuable option.

## 3. Conclusion

Here, we have shown that cocrystallization is another practical method, in addition to soaking of native protein crystals, for the incorporation of a B3C molecule for phasing. Cocrystallization of B3C with bovine trypsin results in a single site with high occupancy. In-house SAD phasing was carried out based on the anomalous signal of one Ca atom and three Br atoms. However, the recommended usage for B3C is to carry out a multi-wavelength anomalous dispersion (MAD) experiment. It has been shown that MAD phasing with B3C yields improved experimental phases (Beck *et al.*, 2010). Since some crystals do not tolerate soaking with high concentrations of I3C or B3C owing to disturbance of the crystal lattice, cocrystallization with these compounds might prove to be a valuable option for incorporating the phasing tools, especially if no suitable native crystals could be obtained in the first place.

Financial support of the International Centre for Diffraction Data (Ludo Frevel Scholarship Award 2009 to TB) and the German Research Foundation (DFG International Research Training Group 1422 Biometals) is greatly appreciated. Stefan Becker is acknowledged for the use of the crystallization robot facilities. Christian Grosse is thanked for help regarding the crystallization setup. Andrea Thorn is thanked for assistance with the data collection.

## References

Beck, T., Gruene, T. & Sheldrick, G. M. (2010). Submitted.
Beck, T., Herbst-Irmer, R. & Sheldrick, G. M. (2009). *Acta Cryst.* C**65**, o237–o239.
Beck, T., Krasauskas, A., Gruene, T. & Sheldrick, G. M. (2008). *Acta Cryst.* D**64**, 1179–1182.
Beck, T. & Sheldrick, G. M. (2008). *Acta Cryst.* E**64**, o1286.
Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B. III, Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007). *Nucleic Acids Res.* **35**, W375–W383.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific LLC, Palo Alto, California, USA.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, doi:10.1107/S0907444909038360.
Sippel, K. H., Robbins, A. H., Reutzel, R., Domsic, J., Boehlein, S. K., Govindasamy, L., Agbandje-McKenna, M., Rosser, C. J. & McKenna, R. (2008). *Acta Cryst.* D**64**, 1172–1178.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.