Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

MAD phasing using the $(Ta_6Br_{12})^{2+}$ cluster: a retrospective study

Oliwia Pasternak,[a] Anna Bujacz,[b] Jacek Biesiadka,[a] Grzegorz Bujacz,[a,b] Michal Sikorski[a] and Mariusz Jaskolski[a,c]*

[a]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland, [b]Institute of Technical Biochemistry, Technical University of Lodz, Poland, and [c]Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland

Correspondence e-mail: mariuszj@amu.edu.pl

Received 12 February 2008
Accepted 21 March 2008

PDB Reference: cytokinin-specific binding protein, 3c0v, r3c0vsf.

As a birthday tribute, this paper is dedicated to Zbyszek Dauter, master of protein phasing by MAD, SAD and other methods, including use of the brain.

The crystal structure of cytokinin-specific binding protein (CSBP) containing four independent molecules with $4 \times 155 = 620$ residues in the asymmetric unit of the $P6_4$ unit cell has been solved by three-wavelength MAD using 1.8 Å resolution data recorded from a crystal derivatized with the dodecabromohexatantalum cation $(Ta_6Br_{12})^{2+}$. The diffraction data contained a very strong anomalous signal (allowing successful phasing even using peak SAD data alone) despite the fact that the five $(Ta_6Br_{12})^{2+}$ clusters found in the asymmetric unit have low occupancy (about 0.3). The derivative structure has been successfully refined to $R = 0.158$, providing interesting details on the geometry of the $(Ta_6Br_{12})^{2+}$ cluster, its interactions with the protein and on the backsoaking of a cytokinin ligand that was originally part of a CSBP–cytokinin complex in the native crystals used for $(Ta_6Br_{12})^{2+}$ derivatization. A simulation analysis of the phasing power of the $(Ta_6Br_{12})^{2+}$ ions at artificially imposed resolution limits shows that it is not possible to resolve the individual Ta atoms if the d_{min} limit of the data is higher than 2.9 Å. Additionally, for successful Ta identification the $(Ta_6Br_{12})^{2+}$ complex should be specifically bound and ordered. Good binding at the protein surface is facilitated by the presence of acidic groups, indicating higher pH buffer conditions to be preferable. In addition, the water channels in the crystal should be sufficiently wide (at least 11 Å) to allow free diffusion of the $(Ta_6Br_{12})^{2+}$ ions on soaking. A retrospective look at the initial molecular-replacement calculations provides interesting insights into how the peculiar packing mode and strong bias of the molecular-replacement-phased electron-density maps had hindered successful solution of the structure by this method.

1. Introduction

The multi-wavelength or single-wavelength anomalous diffraction (MAD or SAD) methods are the most widely used approaches for determination of the three-dimensional structure of new proteins. Of the many potential anomalously scattering atoms, selenium, which is introduced into the protein in the form of SeMet, has become the most popular. However, owing to various limitations, for example the absence of Met residues in the protein sequence, other anomalous scatterers must sometimes be used. This is the case with plant pathogenesis-related class 10 (PR-10) and cytokinin-specific binding proteins (CSBP), in the sequences of which (consisting of about 160 residues) sulfur-containing amino acids are very rare (Sikorski *et al.*, 1999; Pasternak *et al.*, 2006; Handschuh *et al.*, 2007). A very useful compound for chemical derivatization in such situations is the dodecabro-

mohexatantalum cation, $(Ta_6Br_{12})^{2+}$, which has the potential to introduce a huge anomalous signal. To date, the $(Ta_6Br_{12})^{2+}$ cluster has mostly been used for phasing protein structures at low resolution (Thygesen *et al.*, 1996; Ban *et al.*, 2000; Wahl *et al.*, 2000; Szczepanowski *et al.*, 2005), where it scatters the X-rays as a superatom, with the positions of the individual tantalum sites not being resolved. At this resolution, it is also not important whether the $(Ta_6Br_{12})^{2+}$ ion is ordered or not. However, it is possible to resolve the Ta positions if the resolution of the diffraction data is sufficiently high, provided that the cluster is well ordered, as demonstrated by Banumathi *et al.* (2003). In that work, four previously determined protein structures could be successfully phased by SAD at high resolution using the $(Ta_6Br_{12})^{2+}$ ion.

The crystal structure of the salt $(Ta_6Br_{12})Br_2.8H_2O$ recrystallized from water (Knäblein *et al.*, 1997) indicates a composition formula $[Ta_6Br_{12}(H_2O)_6]^{2+}Br^-,OH^-.5H_2O$, with a complex cation comprised of six octahedrally arranged metal–metal-bonded Ta atoms and 12 Br atoms forming bridges at the edges of the Ta octahedron. Six water molecules coordinated by the metal atoms extend from the apices of the Ta octahedron. The cluster is compact and of approximately spherical shape, with a van der Waals radius of about 5.5 Å and high symmetry. Two synthetic procedures have been described (Hay & Messerle, 2002; Koknat *et al.*, 1974), so the compound can be synthesized if necessary.

On derivatization, one $(Ta_6Br_{12})^{2+}$ ion adds 856 electrons to a protein molecule, a considerable contribution to the scattering power, and it is possible to use the anomalous signal of both the Ta and Br atoms. An advantage of using this compound is the green colour that is acquired by the crystal upon $(Ta_6Br_{12})^{2+}$ incorporation, which allows easy control of the soaking process. Moreover, $(Ta_6Br_{12})Br_2$ is soluble in a wide range of buffers even at high salt concentrations. The large size of the $(Ta_6Br_{12})^{2+}$ cluster makes its incorporation into tightly packed protein crystals difficult. Therefore, the solvent content of the crystal is important for successful derivatization by soaking.

In this work, we present the phasing strategy used to solve the unknown crystal structure of a cytokinin-specific binding protein (CSBP) from *Vigna radiata* (mung bean). The final structure of the protein in complex with *trans*-zeatin (a plant hormone from the cytokinin group), refined at atomic resolution using a native data set (1.2 Å), has been described in detail by Pasternak *et al.* (2006). The determination of the CSBP structure is the first example of MAD phasing at high resolution (1.8 Å) solely using the anomalous signal of tantalum introduced into the crystal in the form of $(Ta_6Br_{12})^{2+}$. At this resolution, 30 positions of individual Ta atoms, belonging to five Ta_6Br_{12} clusters, could be resolved.

Prior to the MAD experiment we had employed a variety of other phasing methods, all of which were unsuccessful. Initially, molecular replacement was tried using distant sequence homologues of CSBP as probes. Retrospectively, some interesting conclusions can be made about the failure of the molecular-replacement approach and these are also presented in this paper.

We also tried the method of multiple isomorphous replacement using a number of heavy atoms (Pt, Hg, Os, U, Ir and Au). All the MIR derivatization experiments failed to produce useful phasing information. The main problems were crystal disintegration upon soaking and lack of isomorphism. Two potential heavy-atom derivatives (Pt and Hg) which gave comparatively promising results have also been analyzed retrospectively using the final phase angles of the refined native structure. The peaks found in the difference Fourier maps have a low signal-to-noise ratio and a uniform height distribution, indicating that the derivatization rate was too low for phasing.

2. Materials and methods

2.1. Molecular-replacement calculations

The 17.8 kDa CSBP protein (155 residues) shows limited amino-acid sequence identity (about 20%) to a group of plant class 10 pathogenesis-related proteins (PR-10). The three-dimensional structures of several members of the PR-10 group have been determined previously and thus the molecular-replacement method was initially used to attempt to solve the structure of *V. radiata* CSBP. As search models, the coordinates of four PR-10 molecules were used: those of a birch pollen protein, Bet v1 (PDB code 1bv1; Gajhede *et al.*, 1996), and of three homologous yellow lupin proteins, LlPR-10.1A (1icx; Biesiadka *et al.*, 2002), LlPR-10.1B molecule *A* (1ifv; Biesiadka *et al.*, 2002) and LlPR-10.2A molecule *A* (1xdf; Pasternak *et al.*, 2005). In view of the low sequence identity, a second set of search models was prepared by truncating the side chains of all nonglycine residues to C^β and deleting most of the loop regions as well as five residues at the flexible C-terminus. The task of finding the proper molecular-replacement solution was additionally complicated by the wide range (1.90–4.76 $Å^3\,Da^{-1}$) of acceptable Matthews volumes (Matthews, 1968), indicating the presence of between five and two CSBP molecules in the asymmetric unit, with 35–75% solvent content, respectively. Moreover, the protein crystallized in an enantiomorphic ($P6_2/P6_4$) space group, doubling the number of possibilities. In summary, three problems had to be tackled with the molecular-replacement method: the low sequence similarity of the models, the fact that a single copy of the model molecule corresponded only to a small fraction of the unit-cell scattering matter and the space-group ambiguity.

Initial attempts to solve the molecular-replacement problem, which were carried out using the native data set and the default settings of the programs *AMoRe* (Navaza, 1994) and *EPMR* (Kissinger *et al.*, 1999), did not result in any satisfactory solution. Further calculations with adjustments of the program settings did not provide a solution either, but revealed that the search was very sensitive to the resolution limits and to the model used. Therefore, searches were performed in a systematic manner by automated runs of the *EPMR* program in each of the two space groups and by applying all combinations of the following resolution limits:

20, 18, 16, 12 and 10 Å (low-resolution limit) and 4.5, 4.0, 3.5, 3.2 and 3.0 Å (high-resolution limit). Since the Bet v 1 protein shows the highest sequence identity (25%) among all available PR-10 structures, two search models were constructed based on the Bet v 1 coordinates. Combination of all the above variables resulted in 100 molecular-replacement runs that took about 10 d of calculation on a Linux PC. Only one of the trials, that with the truncated model in the $P6_4$ space group and using the resolution limits 16–3.2 Å, was characterized by a correlation coefficient (0.27) that was significantly higher than the average (about 0.19). Inspection of the molecular packing of the four protein molecules forming this solution strongly supported its correctness. The four CSBP molecules formed a dimer of similar dimers, all related by NCS twofold axes, and filled the unit cell without clashes.

Initial model building was conducted in *ARP/wARP* (Perrakis *et al.*, 1999; Morris *et al.*, 2002). Two strategies were utilized, model rebuilding and *ab initio* model building from a free-atom model, but they did not provide satisfactory results. The program was only able to build about 40 residues of the 620 expected. Also, rigid-body refinement and simulated annealing (Brünger *et al.*, 1998) failed, with R dropping from 0.57 to 0.50 but $R_{\rm free}$ remaining unchanged (0.58–0.57). Manual rebuilding of the model according to the initial $2F_{\rm o} - F_{\rm c}$ electron-density maps did not improve the situation, mostly because the maps had many breaks in the main-chain trace.

Table 1
MAD data-collection and processing statistics.

For comparison, statistics for the high-resolution native data set are also included. Values in parentheses are for the last resolution shell.

	Ta_6Br_{12} derivative			
	Peak	Edge	Remote	Native†
Space group	$P6_4$			$P6_4$
Unit-cell parameters (Å)	a = 113.2, c = 85.1			a = 113.6, c = 86.8
EMBL/DESY beamline	BW7A			BW7B
Detector type	MAR CCD 165 mm			MAR 345 mm IP
Temperature (K)	100			100
Resolution limits (Å)	30.0–1.8 (1.86–1.80)			30.0–1.20 (1.22–1.20)
Wavelength (Å)	1.2547	1.2580	1.2703	0.8441
No. of measured reflections	421616	421594	421917	1277993
No. of unique reflections	56894	56882	56873	189769
$R_{\rm int}$	0.053 (0.448)	0.036 (0.351)	0.033 (0.391)	0.070 (0.639)
$R_{\rm int}$‡	0.032 (0.410)	0.028 (0.323)	0.030 (0.361)	—
$\langle I/\sigma(I)\rangle$	31.5 (4.3)	48.6 (5.9)	55.1 (5.5)	19.7 (2.6)
Redundancy	7.4 (7.0)	7.4 (7.0)	7.4 (7.0)	6.7 (4.5)
Completeness (%)	99.0 (98.0)	99.0 (98.0)	99.0 (97.8)	95.7 (93.2)
Anomalous completeness‡ (%)	97.5 (97.3)	98.4 (97.3)	98.6 (97.2)	—

† See Pasternak *et al.* (2006) and Bujacz *et al.* (2003). ‡ 'Scale anomalous' option in *SCALEPACK*, Bijvoet pairs kept separate.

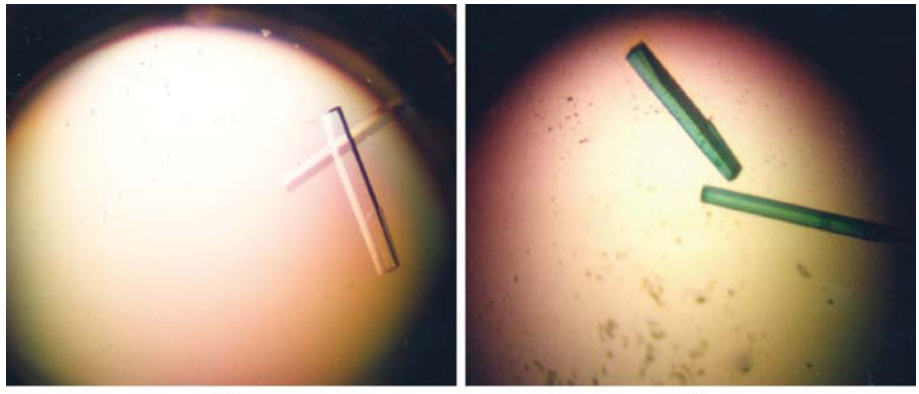

Figure 1
Single crystals of *V. radiata* CSBP (*a*) before and (*b*) after soaking in $(Ta_6Br_{12})Br_2$ solution. The figure shows the acquisition of a green colour by the crystals upon soaking.

2.2. $(Ta_6Br_{12})^{2+}$ derivatization

For derivatization, CSBP single crystals were soaked for 24 h in 5 mM $(Ta_6Br_{12})^{2+}$ solution. The solution was prepared by mixing 2 µl mother liquor with 2 µl 10 mM $(Ta_6Br_{12})^{2+}$ (dissolved in well solution). Using this method, the concentration of the salt and buffer do not change and only the concentration of the zeatin ligand is halved. During the soaking procedure, the crystals gradually acquired a deep green colour with simultaneous fading of the soaking solution (Fig. 1). The subsequent diffraction experiments confirmed that the relatively long soaking time did not have any adverse effects on the crystal quality. Since at the time of our structure-solution efforts having a zeatin complex was not a priority, the $(Ta_6Br_{12})^{2+}$-derivatization experiments were carried out in buffers with reduced zeatin concentration, thus raising the danger of ligand removal through backsoaking.

2.3. MAD data collection

Low-temperature X-ray diffraction data were collected to 1.8 Å resolution for the $(Ta_6Br_{12})^{2+}$ derivative on EMBL beamline BW7A of the DESY synchrotron in Hamburg. Since sodium citrate is a good cryoprotectant at concentrations higher then 1.2 M at pH 7.5 (Bujacz, unpublished results), the crystal was briefly transferred into a drop of the well solution to remove the $(Ta_6Br_{12})^{2+}$ ions from the mounting drop and flash-frozen in a cold nitrogen stream directly on the goniometer. The images were indexed and integrated using *DENZO* and scaled in *SCALEPACK* (Otwinowski & Minor, 1997).

To determine the precise wavelengths for the experiments, a fluorescence scan near the tantalum $L_{\rm III}$ absorption edge

Table 2
Statistics of the solution obtained in *SOLVE* v.2.03.

The criteria are explained in §2.4.

	Over all solutions		Final solution	
Criterion	⟨Value⟩	σ⟨Value⟩	Value	*Z* score
Patterson	1.49	0.545	2.07	1.06
Fourier cross-validation	9.28	7.72	169.0	20.7
Native Fourier CC × 100	8.93	2.09	20.9	5.72
Mean FOM × 100	0.0	5.0	40.7	8.15
Correction for *Z* score				−13.4
Overall *Z* score				22.3

(energy range 9800–9900 eV) was carried out prior to data collection. The values of f' and f'' were estimated from the fluorescence spectrum through the Kramers–Kronig transformation using the program *CHOOCH* (Evans & Pettifer, 2001; Fig. 2). The three wavelengths selected for the MAD experiment were as follows: (i) 1.2547 Å at the absorption peak corresponding to maximum f'' (15.7 e), (ii) 1.2580 Å at the inflection point of the absorption curve corresponding to minimum f' (−22.2 e) and (iii) a 1.2703 Å low-energy remote wavelength.

A full MAD data collection was performed using a single crystal with dimensions 0.15 × 0.15 × 0.35 mm. To ensure high redundancy, for each data set 120° total rotation was covered in 240 frames, with the crystallographic sixfold direction approximately parallel to the rotation spindle. No special techniques such as the inverse-beam method were applied. The statistics of the intensity data are shown in Table 1.

2.4. MAD structure solution

The positions of the tantalum sites were found using the program *SOLVE* v.2.03 (Terwilliger & Berendzen, 1999) with standard parameters. Since the molecular-replacement solution indicated the presence of four protein molecules in the asymmetric unit, the number of anomalous scatterers to look for was initially estimated as 24 (assuming one Ta_6Br_{12} cluster per CSBP monomer). Also, the space-group enantiomorph ($P6_4$) indicated by the molecular-replacement calculations was selected for the first *SOLVE* run. Since it turned out to be the correct choice, no comparative calculations in the alternative space group were made. The Bijvoet differences that were measured at the peak and inflection energies were strongly correlated up to a resolution of 2.38 Å (correlation of above 0.6 in the resolution range 3.6–2.38 Å, correlation of below 0.3 in the range 2.7–1.8 Å), suggesting that it would be possible to resolve the individual tantalum sites.

In *SOLVE*, the solutions are rated according to a *Z*-score parameter based on four criteria: (i) the agreement between the experimental and calculated Patterson maps, (ii) a cross-validation difference Fourier analysis of peak heights at each heavy-atom position excluded from phasing, (iii) the phasing figure of merit and (iv) the connectivity of electron density in the native maps. The statistics of the tentative structure solution (Table 2) were very promising. The *Z*-score value of the native Fourier of 5.7 was a very good indicator of a correct solution (Terwilliger & Berendzen, 1999). The peak heights of the tantalum sites in cross-validation difference Fourier maps were between 28σ and 7σ. The overall *Z* score of 22 was also quite high, but for so many anomalous scatterer sites it would be expected to be even higher. Also, the overall figure of merit of 0.41 seemed plausible, but the authors suggest that for good MAD data sets a figure of merit of over 0.65 is expected. These slightly lower than expected statistics could be explained by the low occupancies of the tantalum sites. Their values, shown in Table 3, are in agreement with the findings of Banumathi *et al.* (2003).

Since the dodecabromohexatantalum cluster contains six Ta atoms, it is expected that for general position $(Ta_6Br_{12})^{2+}$ ions with the individual tantalum positions resolved, the final anomalous-scatterer substructure should consist of a number of sites that is a multiple of six. Since the *SOLVE* procedure revealed the positions of 22 tantalum sites, it was obvious that the solution was not completely correct, *i.e.* that some tantalum sites were missing or/and there were some spurious sites. Examination of the anomalous difference Fourier map showed that the peaks were grouped into five clusters. Three tantalum positions within clusters 1 and 2 were located incorrectly, while 11 atoms were missing (Fig. 3). The anomalous difference map and the known geometry of the Ta_6Br_{12} cluster allowed unambiguous location of the missing atoms. The *ANALYSE* procedure was used to refine the positions of all 30 sites and to calculate the final phase angles.

Solvent flattening, performed using the program *DM* (Cowtan, 1994), improved the phase angles significantly, as illustrated by the increase in the mean figure of merit of phasing from 0.46 to 0.84. The resultant electron-density map was of excellent quality.

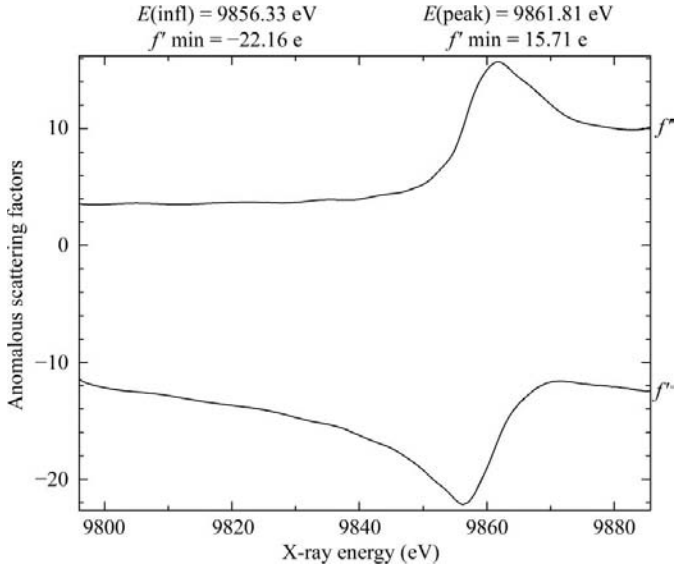

Figure 2
Plot of experimental f' and f'' values *versus* X-ray photon energy prepared using the program *CHOOCH* (Evans & Pettifer, 2001) from an X-ray florescence scan recorded for a $(Ta_6Br_{12})^{2+}$-derivatized CSBP crystal prior to diffraction measurements.

Table 3
Occupancy and *B*-factor statistics for the Ta atoms after structure solution and structure refinement.

Ta_6Br_{12} cluster	1	2	3	4	5
Occupancy					
SOLVE	0.12†	0.17†	0.07	0.08	0.08
	0.14†	0.14	0.07	0.14†	0.11†
	0.17†	0.15†	0.09	0.10†	0.06
	0.13	0.22†	0.09†	0.13†	0.11†
	0.16†	0.15†	0.08†	0.07	0.08†
	0.17†	0.14	0.09†	0.10	0.11†
Final structure‡	0.35	0.35	0.25	0.25	0.25
B factors (Å^2)					
SOLVE	20.2	35.8	28.1	35.2	49.9
	33.2	27.4	24.7	42.8	56.9
	35.2	26.6	30.4	38.0	36.8
	28.7	48.4	34.2	57.7	60.0
	31.6	30.6	31.2	23.7	29.6
	33.8	24.5	33.2	35.4	35.5
Final structure	36.9	27.8	38.3	32.8	34.8
	34.1	28.5	36.9	33.9	35.2
	34.9	28.6	36.0	32.1	32.1
	35.3	28.5	36.5	34.9	35.2
	36.4	30.1	37.9	33.7	36.6
	33.8	28.5	37.9	33.3	34.8

† Ta sites found in the initial *SOLVE* run. ‡ Occupancy factor for all atoms within each cluster.

2.5. Retrospective SAD calculations

Since the solution of the phase problem using a full three-wavelength MAD data set collected at the tantalum L_{III} absorption edge was very straightforward, we were interested whether only a SAD data set corresponding to the absorption-maximum wavelength would also suffice for this purpose. A retrospective SAD calculation in *SOLVE* v.2.12 (with properly set input parameters; see §2.7) identified 24 potential Ta peaks, all of which corresponded to correct Ta sites. This result is somewhat surprising, as the automatically selected Ta set from MAD calculations with a similar number of sites always contained some spurious peaks (regardless of the version of *SOLVE*). This observation illustrates the power of SAD for protein phasing. It has to be admitted, however, that while the SAD phases were perfectly sufficient for successful solution of the structure, the quality of the original MAD-phased maps was significantly better. This is also reflected by the initial SAD figure of merit (0.29), which was significantly lower than in the MAD case (0.41).

Since after density modification of the SAD-phased map it was obvious that the remaining steps of the phasing process would be nonproblematic, the SAD test was terminated at this stage. Like the MAD calculations, the SAD protocol was only performed in the *a priori* selected correct space-group enantiomorph.

2.6. $(Ta_6Br_{12})^{2+}$-derivative structure refinement

Model building of the CSBP–zeatin–$(Ta_6Br_{12})^{2+}$ complex structure was carried out in *ARP*/*wARP* (Perrakis *et al.*, 1999; Morris *et al.*, 2002) using structure-factor amplitudes from the remote-wavelength data set and phases derived from the final 1.2 Å resolution model of native CSBP (PDB code 2flh). Firstly, reflections for R_{free} testing were selected. Up to 1.8 Å resolution, the same reflections as in the native data set were chosen. The subset was then randomly extended to include more than 1000 test reflections. The missing *cis*-proline residues (not interpreted by *ARP*/*wARP*) and the ligand mole-

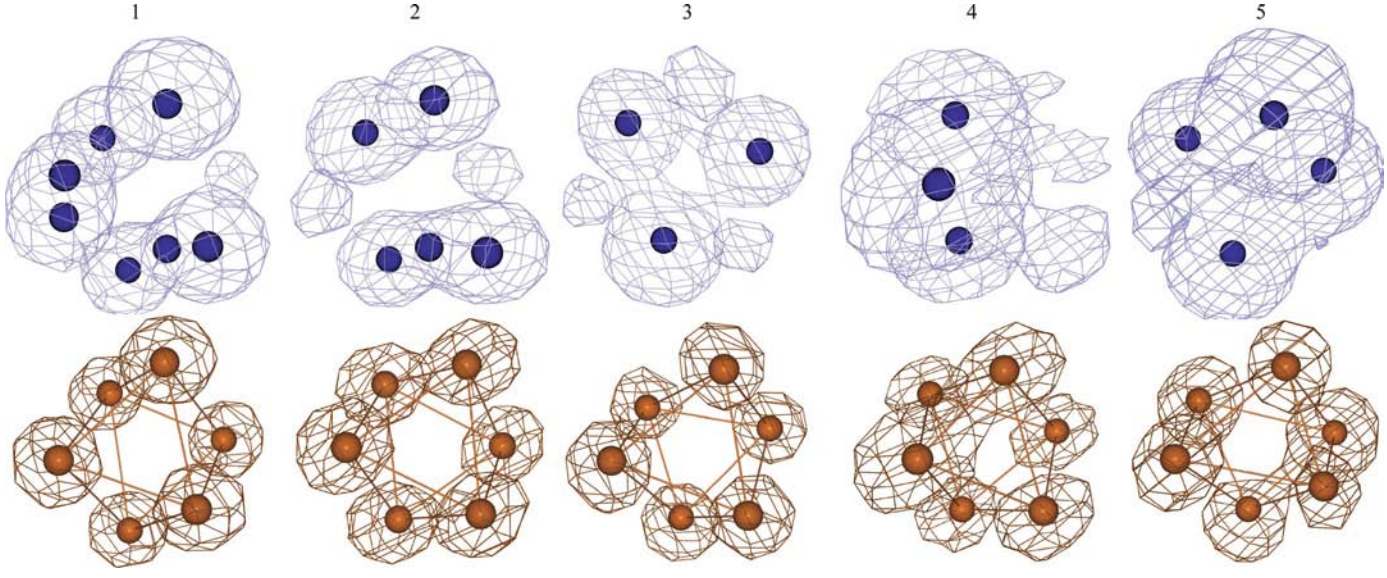

Figure 3
The Ta atoms belonging to five dodecabromohexatantalum clusters in the *V. radiata* CSBP structure numbered according to their heights in anomalous difference maps. The figure shows the sites and the anomalous difference maps obtained in the initial *SOLVE* run (upper row) and after the *ANALYSE_SOLVE* procedure (lower row) (Terwilliger & Berendzen, 1999). The anomalous difference maps calculated using phases obtained in *SOLVE* (blue) are contoured at the following levels: for clusters 1 and 2 at 9σ, for clusters 3 and 4 at 4σ and for cluster 5 at 3σ. The maps obtained in the *ANALYSE_SOLVE* procedure (red) are contoured at 18σ, 10σ and 8σ, respectively.

cules were manually added to the *ARP/wARP* model. The *REFMAC*5 (Murshudov *et al.*, 1997) structure-factor refinement used all 30–1.8 Å resolution reflections and included 13 TLS groups (Winn *et al.*, 2001): two for each CSBP monomer and one for each Ta_6Br_{12} cluster. The stereochemical restraint dictionaries for the zeatin ligand and for the Ta_6Br_{12} cluster were prepared in *Monomer Library Sketcher* from the *CCP*4*i* package (Collaborative Computational Project, Number 4, 1994). The restraints for the $(Ta_6Br_{12})^{2+}$ cation were derived from the atomic resolution model of the cluster (Knäblein *et al.*, 1997). Considering the high symmetry of the cation, only the Ta—Ta (2.92 Å) and Ta—Br (2.61 Å) distances and the Br—Ta—Br (88.0° and 158.4°) and Ta—Br—Ta (68.2°) angles were restrained. The latter restraint is geometrically redundant, but was required by the refinement (*REFMAC*5) and modelling (*Coot*) programs. The rounds of refinement alternated with manual rebuilding sessions, carried out initially in the *Xfit* program from the *XtalView* package (McRee, 1999) and then in *Coot* (Emsley & Cowtan, 2004) using $2F_o - F_c$ and $F_o - F_c$ Fourier maps. After the first round of refinement, water molecules were added automatically in *Xfit* and then verified manually. In all subsequent cycles, the solvent structure was only built by hand in the *Coot* program. The refinement converged with R and R_{free} factors of 0.158 and 0.207, respectively. Analysis with *PROCHECK* (Laskowski *et al.*, 1993) shows that 91.6% of the residues are in the most favoured regions of the Ramachandran plot (Ramachandran *et al.*, 1963), with 8.4% of the residues occupying additionally allowed regions of the plot. The refinement results for the derivative structure are summarized in Table 4.

Table 4
Refinement statistics.

Program used	*REFMAC*5
Resolution limits (Å)	30.0–1.8
Reflections	
Total	55378
R_{free}	1160
Rejection criteria	None
Atoms	
Protein	4948
Zeatin	96
Ta/Br	30/60
Other metals (Na)	2
HEPES	15
Solvent	578
R/R_{free}	0.158/0.207
Average B factors (Å^2)	
Protein atoms	28.02
Zeatin atoms	34.54
Inner	24.63
Outer	44.40
$(Ta_6Br_{12})^{2+}$	32.69
HEPES	32.04
Solvent	38.65
R.m.s. deviations from ideality	
Bonds (Å)	0.018
Angles (°)	1.8
Torsion angles (°)	5.7
Chiral volumes (Å^3)	0.123
Ramachandran plot statistics (%)	
Most favoured regions	91.6
Additionally allowed regions	8.4

2.7. Resolution of the Ta_6 metal cluster

Since the original MAD data are of comparatively high resolution, we were interested to investigate the effect of successive limitation of the resolution on the usefulness of the data for phasing. For this purpose, the *SOLVE* runs were repeated using the $(Ta_6Br_{12})^{2+}$ MAD data truncated at different high-resolution levels.

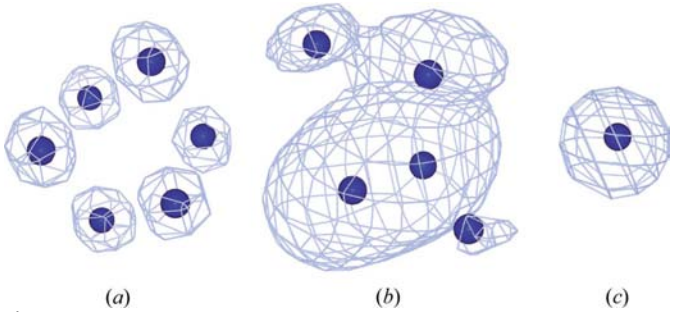

Figure 4
Anomalous difference map calculated for the best Ta_6Br_{12} cluster, TBR1, using phases generated by *SOLVE* and subsequent density modification for Ta-MAD data truncated at different d_{min} levels: 1.8 Å (*a*), 2.6 Å (*b*) and 2.9 Å (*c*). The maps were contoured at levels of 18σ for (*a*) and (*c*) or 6σ for (*b*). (*b*) must be contoured at a lower σ level to emphasize its features or it would appear to be spherical. In (*c*), lowering of the contour level does not reveal any real features. The dark balls indicate the Ta positions located automatically by *SOLVE*.

Two comments must be made here. (i) The data quality (and in consequence the phasing power) of a 1.8 Å set truncated to, for instance, 2.8 Å resolution is not the same as for a set measured to 2.8 Å resolution. A truncated high-resolution data set will be of high quality at 2.8 Å, while for a genuine low-resolution data set the data quality at 2.8 Å will be marginal. (ii) Because of the replacement of the old (v.2.03) version of *SOLVE* with a new one (v.2.12), we were unable to repeat the simulations in exactly the same way as the original calculations. For instance, the default values of the new version are different and an automatic run of *SOLVE* v.2.12 would find only one Ta atom in each cluster (regardless of resolution), while no such limitation was present in v.2.03. To overcome this drawback of *SOLVE* v.2.12, the user has to set the value of ntol_site (minimum distance, in grid units, for separate sites) to instruct the program to recognize the individual Ta positions as separate.

Our *SOLVE* simulations were calculated with the high-resolution limit set to 1.8, 2.0, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0 and 3.5 Å. For each run, the phases calculated by the program were used directly for phase improvement by density modification. (In the real run, the Ta_6 clusters were first analyzed manually and completed according to the electron-density maps and their octahedral geometry.) The results indicate that at a resolution of at least 2.5 Å, all the 'identifiable' electron-density peaks correspond to separate Ta atoms of the Ta_6Br_{12} clusters. At d_{min} = 2.6 Å, all the $(Ta_6Br_{12})^{2+}$ cations, except for the best ion, TBR1, are visible as single 'superatom' peaks. With worsening resolution the TBR1 cluster also becomes increasingly diffuse and at 2.9 Å resolution its electron density coalesces into a single peak (Fig. 4).

3. Results and discussion

3.1. Overall fold and crystal packing

Since the atomic resolution (1.2 Å) structure of *V. radiata* CSBP has been described previously (Pasternak *et al.*, 2006), only some general data concerning the 1.8 Å MAD structure are provided here. The general fold of the CSBP protein consists of a seven-stranded antiparallel β-sheet which forms a grip around a long C-terminal helix α3. Between these two main structural elements, an internal cavity is present which is closed at one end by two short helices α1 and α2, which join the β1 and β2 strands forming the opposite edges of the β-sheet (Fig. 5). There are nine loops in the polypeptide fold, five of which are β-hairpin loops (L4–L8) connecting the consecutive β-strands of the β-sheet. There are four CSBP molecules (*A*, *B*, *C* and *D*) in the asymmetric unit. Apart from the four protein monomers, the asymmetric unit contains six zeatin molecules, five Ta_6Br_{12} clusters, two Na^+ cations, one HEPES molecule and 578 modelled water molecules. The zeatin-binding sites are found within the cavity formed inside the protein molecule. In *V. radiata* CSBP, the binding cavity is capable of incorporating two zeatin ligands. One of them (inner) is bound deep in the protein core, while the second (outer) ligand is more accessible to solvent. In the 1.2 Å structure each protein molecule binds two zeatin molecules, with the exception of polypeptide chain *C*, in which only the inner ligand molecule was found. In addition, two zeatin ligands were found outside of the cavity between the protein chains. Generally, the zeatin-binding pattern is similar in the native and MAD structures. The most striking difference is the absence of the single inner ligand molecule in the binding cavity of molecule *C*. The three remaining binding cavities (in molecules *A*, *B* and *D*) are occupied by both the inner and the outer ligands. However, the outer ligand in monomer *D* has fractional occupancy. Also, in contrast to the 1.2 Å structure, no interstitial zeatin ligands were found between the protein chains. These differences indicate backsoaking of the zeatin molecules during $(Ta_6Br_{12})^{2+}$ derivatization, since no zeatin supplement was present in the soaking solution.

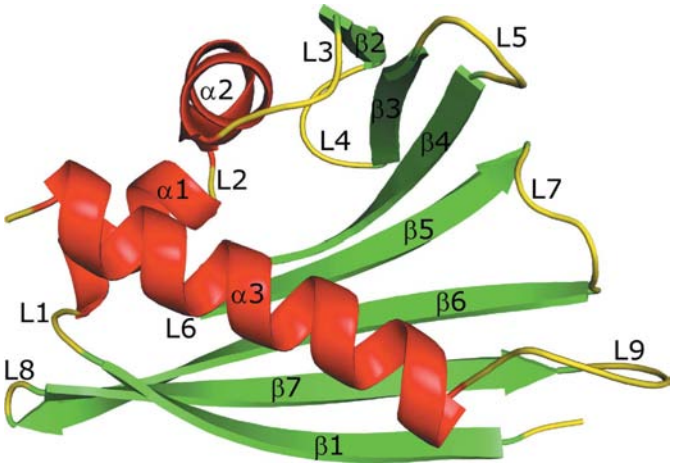

Figure 5
Protein fold of *V. radiata* CSBP. The annotated secondary-structure elements are also colour-coded using red for α-helices, green for β-strands and yellow for loop regions. Structural figures were prepared in *DINO* (Philippsen, 2003) and *PyMOL* (DeLano, 2002).

In the derivative structure, two Na^+ cations have been identified within loop L9 of molecules *B* and *C*, with similar coordination spheres as in the native 1.2 Å structure (two main-chain carbonyl O atoms and four water molecules in octahedral arrangement). One of the Na^+ sites (in molecule *C*) has identical coordination in both structures, involving Ser124 and Ile126 residues. In molecule *B*, the coordination is provided by Asp123 and Ile126. The cations were interpreted as sodium without ambiguity using the same procedure as for the 1.2 Å structure. Briefly, the calculated metal valences assuming sodium (Brese & O'Keeffe, 1991) are 0.98 and 1.05, which are very close to the theoretical value of 1.00, while the CBVS (Müller *et al.*, 2003) parameters are 1.54 and 1.65, which are again very close to the value expected for a sodium cation (1.57).

The four CSBP molecules in the asymmetric unit lie in one layer parallel to (001). The molecular layers are stacked in the *c* direction, with the shortest distances between the centres of protein molecules in consecutive layers being about 20 Å. The most prominent molecular contacts occur within the layers, organizing the monomers into different hexagons centred on the 6_4, 3_1 and 2 axes. The crystal packing consists of closely connected molecular dimers (*BC*), to which less tight but pseudosymmetric connections are formed by molecules *A* and *D*. In consequence, the four independent molecules in the crystal can be described as forming an intimate dimer of loose dimers (Figs. 6*a* and 7*b*).

The *c* dimension of the hexagonal unit cell is the only lattice parameter that undergoes a significant (over 2%) change on $(Ta_6Br_{12})^{2+}$ derivatization. Curiously, the unit cell shrinks on $(Ta_6Br_{12})^{2+}$ binding, indicating either crystal dehydration during soaking or, more likely, an attractive electrostatic force exerted by the dodecabromohexatantalum cation on the protein molecules in the adjacent (001) layers.

3.2. $(Ta_6Br_{12})^{2+}$ binding

There are five $(Ta_6Br_{12})^{2+}$ clusters (TBR1–TBR5) in the asymmetric unit. $2F_o - F_c$ and $F_o - F_c$ electron-density maps clearly indicated not only the positions of the Ta atoms but also of most of the Br atoms. The complete clusters were modelled with the aid of their known highly symmetrical geometry. At the beginning of the refinement, the occupancies of the Ta and Br atoms were set to 0.5. Inspection of the Ta/Br temperature factors and of $F_o - F_c$ electron-density maps was the basis for occupancy adjustments in subsequent refinement cycles. In the final model, all the TBR clusters are characterized by mean temperature factors of about 30 $Å^2$ and have fractional occupancies of 0.35 for TBR1 and TBR2 or 0.25 for the remaining clusters.

Generally, the $(Ta_6Br_{12})^{2+}$-binding sites are formed near the surface of the protein molecules, in close proximity to negatively charged glutamate or aspartate residues. The interactions of acidic side chains with the tantalum core are

supported by the interactions of basic and hydroxyl groups with the Br shell. Clusters TBR1 and TBR2 are anchored at NCS-related sites located near molecules *A* and *C*, respectively (Fig. 6). Similarly, TBR4 and TBR5 are bound to NCS-related sites located near molecules *D* and *C*, respectively. Cluster TBR3 does not have an NCS-related counterpart.

In more detail, the TBR1 binding site is formed between strand β2 of molecule *A*, helix α1 of molecule *B* and loop L9 of a symmetry-related molecule *D*. The TBR1 cluster forms closest interactions with the side chain of Asp40 located in strand β2 of molecule *A* and with Glu127 (loop L9) of a symmetry-related molecule *B*, with the Ta$\cdots$O distances being 3.2 and 2.3 Å, respectively. The amide group of Gln42 from molecule *A* makes a hydrogen-bonding contact with one of the Br atoms. A number of other Br atoms from this cluster form similar hydrogen bonds to N–H groups from molecule *B*. The TBR2 cluster is bound in a similar manner (Fig. 6). Here, the secondary-structure elements forming the binding site come from molecule *C* (strand β2), molecule *D* (helix α1) and molecule *B* (loop L9) and the interactions are formed by analogous side chains. The N—H$\cdots$Br hydrogen bonds also have a similar character. TBR4 is located between molecule *D* (strands β4 and β5) and a crystallographic copy of molecule *A* (helix α3 and loops L5 and L7; Fig. 6). This cluster only makes one close interaction with an acidic side chain (Ta1$\cdots$O$^{\varepsilon 2}$ Glu80*D*, 2.7 Å), but its Br centres form several interactions with $-NH_3^+$ and —OH donors, including the hydroxyl group of the zeatin ligand Zea2. The binding site of the TBR5 cluster is similarly formed by chains β4 and β5 from molecule *C* and helix α3, loop L5 and loop L7 from a symmetry-related molecule *C*. Obviously, there is no ZeaO—H$\cdots$Br hydrogen bond, as the binding pocket of molecule *C* is empty. TBR3 is located in the neighbourhood of the TBR1 ion. Its binding site is formed between loop L5 of molecule *A*, loop L3 of molecule *B* and loop L9 from a symmetry-related molecule *D*. The TBR3 cluster is located further from the protein molecules than the other $(Ta_6Br_{12})^{2+}$ clusters. It does not have contacts with acidic groups but is attached to the surface of molecule *A* *via* two (Lys)$NH_3^+\cdots$Br hydrogen bonds. The TBR1 and TBR3 binding sites are so close that in view of their charge and fractional occupancy it is obvious that they do not populate this area at the same time, but are two mutually

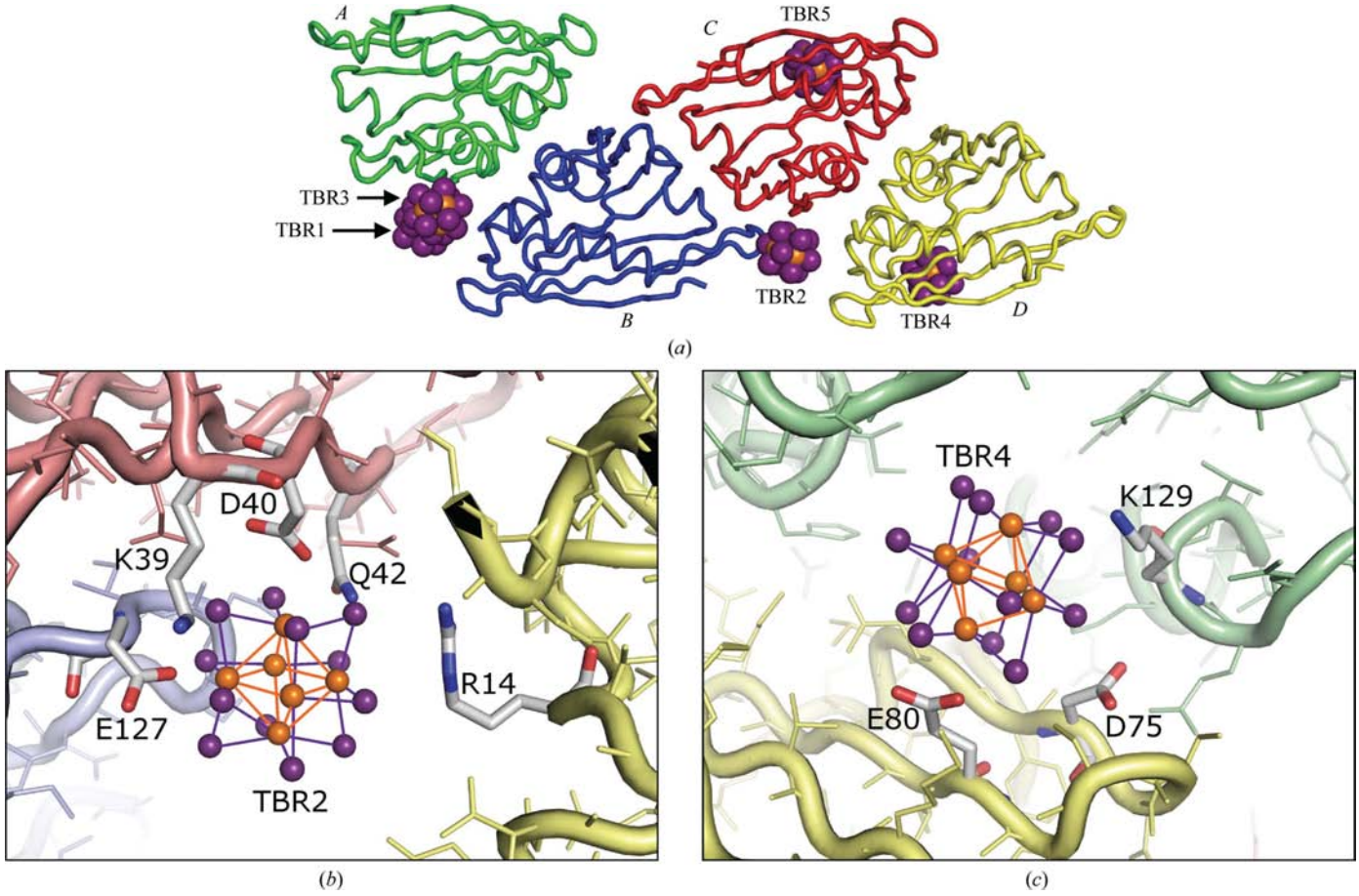

Figure 6
$(Ta_6Br_{12})^{2+}$-binding sites. The four CSBP molecules arranged as in the crystallographic asymmetric unit are coloured green (*A*), blue (*B*), red (*C*) and yellow (*D*). The $(Ta_6Br_{12})^{2+}$ clusters are marked using orange Ta spheres and violet Br spheres. (*a*) An overview of the relation between the four CSBP molecules and the five $(Ta_6Br_{12})^{2+}$ ions. The symmetry-related protein molecules that complete the $(Ta_6Br_{12})^{2+}$-binding sites have been omitted for clarity. (*b*) Close-up view of the binding details of two representative $(Ta_6Br_{12})^{2+}$ ions. The binding site of the TBR2 ion is shown on the left. The same interactions are observed for the TBR1 binding site. The binding site of the TBR4 ion is shown on the right. A similar environment is observed for TBR5. The amino-acid residues located in the vicinity of the $(Ta_6Br_{12})^{2+}$ ions are shown in stick representation.

exclusive alternatives. If both of them coexisted simultaneously, the closest Br···Br approaches would be 3.6 Å.

It has been shown by Knäblein *et al.* (1997) that in the atomic resolution model of the dodecabromohexatantalum cluster the six Ta atoms are coordinated by O atoms from water molecules at distances of about 2.25 Å. In the CSBP structure, several O atoms located at distances between 2.0 and 3.6 Å from the Ta atoms might serve as potential apical ligands. The discrepancy between the ideal coordination distance and some of the observed distances can be explained by the low occupancies of the $(Ta_6Br_{12})^{2+}$ clusters. It is possible that upon $(Ta_6Br_{12})^{2+}$ binding the nearby side chains assume dual conformations and water molecules are displaced with fractional occupancies. However, if the occupancies of the binding residues matched the low occupancy of the $(Ta_6Br_{12})^{2+}$ clusters, their alternate conformations would be impossible to model in 1.8 Å electron-density maps. A similar justification can be given for several N—H···Br hydrogen bonds (2.4–2.6 Å) that are shorter than the acceptable limits.

Superposition of the C^{α} atoms of the native and $(Ta_6Br_{12})^{2+}$-derivative structures shows that the protein conformation did not change upon derivatization. The r.m.s. deviations calculated using *ALIGN* (Cohen, 1997) were between 0.13 Å for molecule *A* and 0.25 Å for molecule *C*. This is significantly less than for pairwise superpositions of the CSBP molecules in the asymmetric unit of the native crystal (0.5 Å) for similar numbers of matched C^{α} atoms. The local influence of $(Ta_6Br_{12})^{2+}$ binding on the conformation of individual directly interacting side chains cannot be excluded, but the low level of incorporation of the clusters and the resolution limit of 1.8 Å are not sufficient for definite conclusions.

3.3. Molecular dimensions of the $(Ta_6Br_{12})^{2+}$ cation

In general, the $(Ta_6Br_{12})^{2+}$ cations have the expected dimensions and symmetry. Table 5 summarizes the geometrical parameters of the $(Ta_6Br_{12})^{2+}$ ions, including the Ta—Ta and Ta—Br bond lengths and some valence angles. There is no point in detailed analysis of these data as the geometry of the $(Ta_6Br_{12})^{2+}$ ions was restrained during refinement. The standard deviations and in consequence the deviations from ideal symmetry are generally somewhat higher for the cations with higher occupancy. This may reflect the fact that the $(Ta_6Br_{12})^{2+}$ cations with tighter binding are also more distorted by the interactions with their protein environment. Additionally, the geometry of the lower occupancy cations may be more strongly influenced by the idealizing restraints.

3.4. Molecular-replacement calculations

The ultimate determination of the CSBP structure by MAD phasing has enabled us to retrospectively analyze the failure of

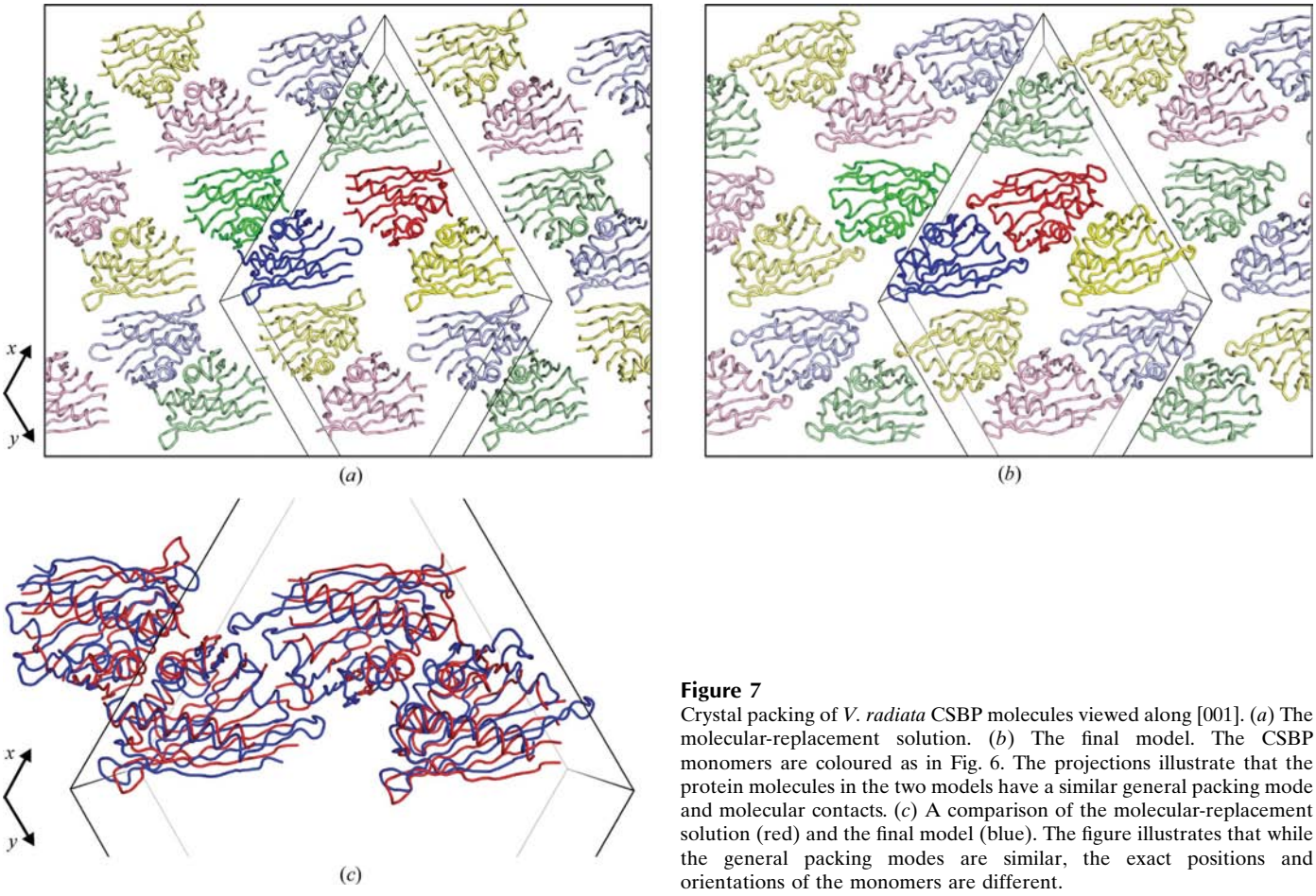

Figure 7
Crystal packing of *V. radiata* CSBP molecules viewed along [001]. (*a*) The molecular-replacement solution. (*b*) The final model. The CSBP monomers are coloured as in Fig. 6. The projections illustrate that the protein molecules in the two models have a similar general packing mode and molecular contacts. (*c*) A comparison of the molecular-replacement solution (red) and the final model (blue). The figure illustrates that while the general packing modes are similar, the exact positions and orientations of the monomers are different.

Table 5
Geometry of the $(Ta_6Br_{12})^{2+}$ clusters.

The numbers are mean values (with standard deviations) calculated for all instances of a given parameter. Where appropriate, the stereochemical restraint target is given in parentheses below the corresponding parameter.

Distances (Å)				Angles (°)				
Ta—Ta (2.92)	Ta—Br (2.61)	Br—Br (nearest)	Br—Br (farthest)	Ta—Ta—Ta	Ta—Ta—Ta	Ta—Br—Ta† (68.2)	Br—Ta—Br (88.0)	Br—Ta—Br (158.4)
2.92 (1)	2.60 (1)	3.61 (8)	7.23 (2)	60.0 (4)	90.0 (5)	68.1 (5)	87.9 (25)	158.0 (9)

† This triangulation-redundant restraint was required to avoid warning/error messages in *REFMAC*5 and in *Coot*.

Table 6
The operations (translation and rotation) required for superposition of the molecular-replacement model (either the whole tetramer or individual monomers) on the corresponding molecule(s) from the final model.

The r.m.s. deviation statistics for the specified number of atom pairs (in parentheses) are also listed. The calculations were made in *ALIGN* (Cohen, 1997) for all C^{α} atoms (auto mode).

Molecule	Translation along *a*, *b*, *c*† (Å)	Rotation (°)	R.m.s. deviation (Å) (No. of C^{α} pairs)
ABCD	−2.5, −1.9, −13.4	1.9	2.71 (500)
A	0.4, −1.7, −14.1	15.2	1.35 (126)
B	−2.2, −0.4, −13.8	4.6	1.30 (130)
C	−4.0, −2.8, −13.1	4.1	1.27 (124)
D	0.4, −2.4, −12.7	8.6	1.31 (126)

† Vector between the centroids of the original and superimposed molecules.

the molecular-replacement approach. The first conclusion is that CSBP indeed belongs to the PR-10 structural class even though PR-10 models failed to solve the structure. The r.m.s. deviation between the C^{α} atoms of CSBP and the closest PR-10 relatives (LlPR-10.1B, Bet v 1) is of the order of 1.5–1.6 Å. On the other hand, the individual CSBP molecules can be superposed with an average r.m.s. deviation of 0.5 Å.

The analysis of the molecular-replacement results was only performed using the model obtained with *EPMR*, as described in §2.1. The molecular-replacement solution consists of four CSBP-like PR-10 molecules placed in the asymmetric unit of the $P6_4$ unit cell, forming a dimer of dimers (Fig. 7). Through the operation of the space-group symmetry, the four molecules are propagated to form layers of different hexagons centred on the 6_4, 3_1 and 2 axes. Such a regular arrangement does not seem to be accidental and is strongly suggestive of a correct solution. However, numerous attempts to refine this model through exhaustive rigid-body optimization (*EPMR*, Kissinger *et al.*, 1999; *CNS*, Brünger *et al.*, 1998; *REFMAC*5, Murshudov *et al.*, 1997) followed by atomic refinement did not improve the electron-density maps and a better model could not be obtained by gradual rebuilding. A comparison of the molecular-replacement solution with the final CSBP model shows that the general packing scheme of the monomers within the unit cell was correct and that the intermolecular contacts did indeed resemble the real situation.

However, a closer inspection reveals that there are some significant crystal-packing differences. In space group $P6_4$, where the origin is not defined along the *c* axis, it would be expected that a superposition of the molecular-replacement solution on the final model should only correspond to a translation along the *z* direction. When the molecular-replacement model was translated manually (about 13 Å) along *c*, an approximate coincidence with the final model derived from MAD phasing could be obtained. The result of this operation, shown in Fig. 7(*c*), clearly indicates that while the two solutions are generally similar, there are still important discrepancies between them.

Superposition of the C^{α} atoms of the four independent molecules of the molecular-replacement solution on the corresponding tetramer of the final model reveals the detailed operations required to achieve the best fit (Table 6). Specifically, optimal superposition requires translations of 2.5 and 1.9 Å along the *a* and *b* directions, respectively, combined with a rotation of almost 2°. Even after these fine-tuning operations, the r.m.s. deviation between the corresponding C^{α} atoms (500 of 544 pairs) is still 2.7 Å. To check whether this overall discrepancy is not a result of incorrect orientation of, for instance, just one monomer, each molecule of the molecular-replacement tetramer has been separately superimposed on the corresponding molecule from the final structure. The results show (Table 6) that all the monomers need to be significantly rotated to match the correct solution. Relatively high rotations of 15° and 7° are required for molecules *A* and *D*, respectively, while the remaining two molecules only need to be slightly rotated by approximately 4°. Comparison of the overall rotation for the tetrameric set of molecules with the adjustments necessary for individual molecules indicates that the individual corrections are not correlated. It is also apparent that with orientation and translation errors as high as 15° and 2 Å, respectively, the molecular-replacement models had little chance of being expanded/corrected by the refinement procedure.

Inspection of the $2F_o - F_c$ maps calculated with the phase angles obtained after rigid-body refinement of the molecular-replacement solution shows that the secondary-structure elements are clearly visible (Fig. 8). However, owing to strong model bias, it was impossible to assign the amino-acid sequence or build the missing parts of the model. These observations emphasize the notion that model bias of molecular-replacement solutions with errors exceeding the convergence radius of structure-factor refinement can be a serious obstacle to successful completion of the model.

4. Conclusions

The successful phasing of a large and complicated protein crystal structure by MAD, or even by absorption-peak SAD, using the $(Ta_6Br_{12})^{2+}$ cluster underlines several interesting facts. It was shown by Banumathi *et al.* (2003) that the crystal

packing can influence the incorporation of the $(Ta_6Br_{12})^{2+}$ ions by soaking. In the crystal of *V. radiata* CSBP, the packing of the protein molecules results in the formation of solvent channels with a diameter of about 10–20 Å running parallel to *c*. These channels are wide enough for the $(Ta_6Br_{12})^{2+}$ ions, which have a diameter of about 11 Å, to penetrate the crystal and react with the protein molecules.

Theoretically, in order to resolve the individual Ta positions the diffraction data must have a d_{min} of at least 2.9 Å. However, our simulations with data truncated at different resolutions indicate that to determine the individual positions of low-occupancy Ta atoms, good-quality 2.6 Å data would be required.

Knowledge of the geometry of the $(Ta_6Br_{12})^{2+}$ cluster is certainly helpful in the identification of additional Ta sites and in pruning the spurious sites. It will more quickly lead to a complete model of the anomalous-scatterer substructure and, in consequence, to much more reliable experimental phasing.

With the use of $(Ta_6Br_{12})^{2+}$ scattering, very high-quality phase angles can be obtained despite the typically low occupancy of the Ta atoms. The occupancies of the Ta atoms obtained during the anomalous-scatterer substructure refinement (about 0.2–0.1) are in general agreement with the final occupancies obtained from structure-factor refinement of the complete structure (0.35–0.25), although it is noted that the site-occupancy fractions assigned jointly to all atoms of a given $(Ta_6Br_{12})^{2+}$ ion during structure-factor refinement are systematically higher than the values returned by the anomalous-scatterer refinement procedure. A possible source of this systematic error could be connected, for instance, to overestimated anomalous corrections assigned to Ta scattering.

As a summary, several practical conclusions are possible. (i) Despite the low level of incorporation into the protein crystal lattice, $(Ta_6Br_{12})^{2+}$ clusters can still lead to successful MAD or even SAD phasing because of the large number of electrons that are present in this complex ion. (ii) Since the positively charged $(Ta_6Br_{12})^{2+}$ clusters are primarily associated with acidic groups on the protein surface, buffers of higher pH that guarantee the anionic form of carboxylic acid side chains are preferred. The carboxylate O atoms coordinate the Ta atoms at the vertices of the Ta_6 octahedron. Auxiliary binding is provided by N/O—H···Br hydrogen bonds to the bromine 'shell' of the $(Ta_6Br_{12})^{2+}$ cation, which can also include positively charged $-NH_3^+$ donors. (iii) The solvent channels in the crystal should be at least 11 Å wide to

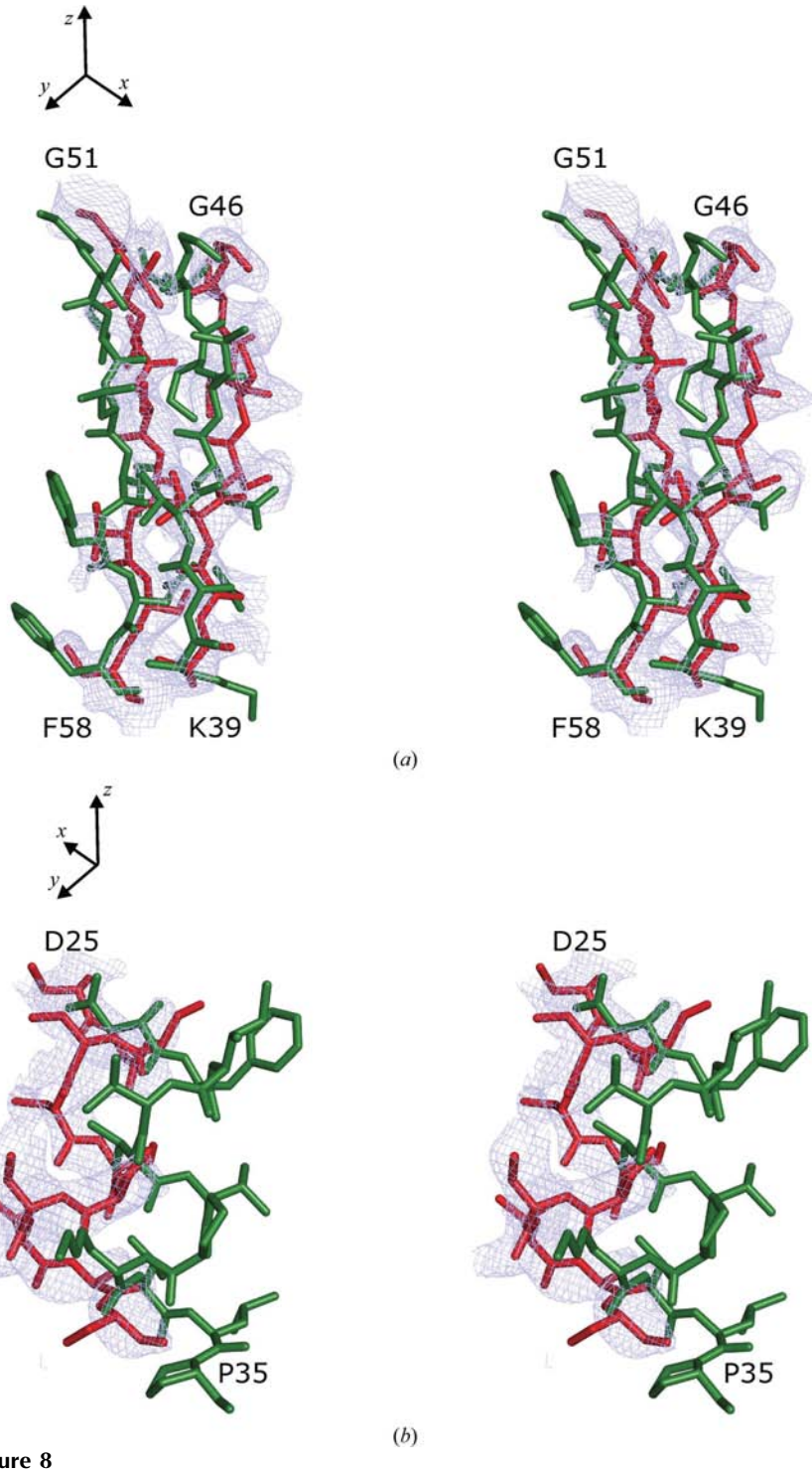

Figure 8
Detailed comparison of the models obtained by molecular replacement (red) and by MAD phasing (green). A fragment of the β-sheet of molecule *B* (*a*) and of helix α2 of molecule *C* (*b*) are shown in stick representation (stereoview). The $2F_o - F_c$ electron-density map was calculated with phases obtained by molecular replacement after rigid-body refinement. The figure illustrates that a translation along *c* (vertical) is not sufficient to superpose the two models and that the strong model bias of the molecular-replacement-phased map makes model rebuilding impossible.

allow free diffusion of the $(Ta_6Br_{12})^{2+}$ ions to their binding sites. (iv) With d_{min} = 2.6 Å diffraction data it is possible to separate the Ta atoms even in low-occupancy (about 0.2) Ta_6Br_{12} clusters. No Ta separation is possible with d_{min} > 2.9 Å.

5. Protein Data Bank depositions

The atomic coordinates of *V. radiata* CSBP in complex with zeatin and $(Ta_6Br_{12})^{2+}$ have been deposited in the Protein Data Bank together with the complete three-wavelength MAD X-ray diffraction data set (1.8 Å resolution) under accession code 3c0v. The isomorphous native structure of *V. radiata* CSBP in complex with zeatin is available from the PDB together with atomic resolution (1.2 Å) diffraction data under accession code 2flh.

This work was supported in part by grants from the State Committee for Scientific Research to MS (grants 6 P04B 004 21 and 2 P04A 053 27) and by a subsidy from the Foundation for Polish Science to MJ. Some of the calculations were performed in the Poznan Metropolitan Supercomputing and Networking Center.

References

Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000). *Science*, **289**, 905–920.
Banumathi, S., Dauter, M. & Dauter, Z. (2003). *Acta Cryst.* D**59**, 492–498.
Biesiadka, J., Bujacz, G., Sikorski, M. M. & Jaskolski, M. (2002). *J. Mol. Biol.* **319**, 1223–1234.
Brese, N. E. & O'Keeffe, M. (1991). *Acta Cryst.* B**47**, 192–197.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Bujacz, G., Pasternak, O., Fujimoto, Y., Hashimoto, Y., Sikorski, M. M. & Jaskolski, M. (2003). *Acta Cryst.* D**59**, 522–525.
Cohen, G. E. (1997). *J. Appl. Cryst.* **30**, 1160–1161.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, G. & Pettifer, R. F. (2001). *J. Appl. Cryst.* **34**, 82–86.
Gajhede, M., Osmark, P., Poulsen, F. M., Ipsen, H., Larsen, J. N., van Neerven, R. J. J., Schou, C., Lowenstein, H. & Spangfort, M. D. (1996). *Nature Struct. Biol.* **3**, 1040–1045.
Handschuh, L., Femiak, I., Kasperska, A., Figlerowicz, M. & Sikorski, M. M. (2007). *Acta Biochim. Pol.* **54**, 783–796.
Hay, D. N. T. & Messerle, L. (2002). *J. Struct. Biol.* **139**, 147–151.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.
Knäblein, J., Neuefeind, T., Schneider, F., Bergner, A., Messerschmidt, A., Löwe, J., Steipe, B. & Huber, R. (1997). *J. Mol. Biol.* **270**, 1–7.
Koknat, F. W., Parson, J. A. & Vongvusharintra, A. (1974). *Inorg. Chem.* **13**, 1699–1702.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283-291.
McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2002). *Acta Cryst.* D**58**, 968–975.
Müller, P., Köpke, S. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 32–37.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pasternak, O., Biesiadka, J., Dolot, R., Handschuh, L., Bujacz, G., Sikorski, M. M. & Jaskolski, M. (2005). *Acta Cryst.* D**61**, 99–107.
Pasternak, O., Bujacz, G. D., Fujimoto, Y., Hashimoto, Y., Jelen, F., Otlewski, J., Sikorski, M. M. & Jaskolski, M. (2006). *Plant Cell*, **18**, 2622–2634.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Philippsen, A. (2003). *DINO*. http://www.dino3d.org.
Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 95–99.
Sikorski, M. M., Biesiadka, J., Kasperska, A. E., Kopcińska, J., Łotocka, B., Golinowski, W. & Legocki, A. B. (1999). *Plant Sci.* **149**, 125–137.
Szczepanowski, R. H., Filipek, R. & Bochtler, M. (2005). *J. Biol. Chem.* **280**, 22006–220011.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Thygesen, J., Weinstein, S., Franceschi, F. & Yonath, A. (1996). *Structure*, **4**, 513–518.
Wahl, M. C., Bourenkov, G. P., Bartunik, H. D. & Huber, R. (2000). *EMBO J.* **19**, 174–186.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.