Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The landscape of cytokinin binding by a plant nodulin

M. Ruszkowski,[a] K. Szpotkowski,[a] M. Sikorski[a] and M. Jaskolski[a,b]*

[a]Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland, and [b]Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland

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

Received 7 July 2013
Accepted 6 August 2013

PDB References: MtN13, complex with *trans*-zeatin, 4jhg; complex with N^6-isopentenyladenine, 4gy9; complex with kinetin, 4jhh; complex with N^6-benzyladenine, 4jhi

Nodulation is an extraordinary symbiotic interaction between leguminous plants and nitrogen-fixing bacteria (rhizobia) that assimilate atmospheric nitrogen (in root nodules) and convert it into compounds suitable for the plant host. A class of plant hormones called cytokinins are involved in the nodulation process. In the model legume *Medicago truncatula*, nodulin 13 (MtN13), which belongs to the pathogenesis-related proteins of class 10 (PR-10), is expressed in the outer cortex of the nodules. In general, PR-10 proteins are small and monomeric and have a characteristic fold with an internal hydrophobic cavity formed between a seven-stranded antiparallel β-sheet and a C-terminal α-helix. Previously, some PR-10 proteins not related to nodulation were found to bind cytokinins such as *trans*-zeatin. Here, four crystal structures of the MtN13 protein are reported in complexes with several cytokinins, namely *trans*-zeatin, N^6-isopentenyladenine, kinetin and N^6-benzyladenine. All four phytohormones are bound in the hydrophobic cavity in the same manner and have excellent definition in the electron-density maps. The binding of the cytokinins appears to be strong and specific and is reinforced by several hydrogen bonds. Although the binding stoichiometry is 1:1, the complex is actually dimeric, with a cytokinin molecule bound in each subunit. The ligand-binding site in each cavity is formed with the participation of a loop element from the other subunit, which plugs the only entrance to the cavity. Interestingly, a homodimer of MtN13 is also formed in solution, as confirmed by small-angle X-ray scattering (SAXS).

1. Introduction

Nitrogen is one of the most important macroelements. Despite the high abundance of nitrogen in the atmosphere, only a very limited number of living organisms are capable of assimilating gaseous nitrogen. The leguminous plants (Fabaceae) have developed a solution to this drawback by encapsulating nitrogen-fixing bacteria from the *Rhizobium* genus in special root organs called nodules. As a result, legumes do not require or need only very little nitrogen fertilization. Legumes are thus the perfect crops and are important to humans. Peas, beans, soybeans, lentils, peanuts and chickpeas are only some examples of legumes that have been cultivated for millennia and whose nutritious values are also appreciated today. The legumes provide a roughly equal amount of fixed nitrogen to the chemical fertilizer industry, reducing the ecological load and cost of cropping. Legume plants release certain signalling flavonoids into the soil (Wasson *et al.*, 2006) which act as chemoattractants for soil-dwelling rhizobia. In response, the bacteria synthesize and secrete a variously modified lipochitin

oligosaccharide called the Nod factor (NF). NF is perceived in the root epidermis by the Nod-factor perception/LysM kinase receptor (NFP/LYK), which activates Nod-factor signalling pathways that eventually lead to nodule formation. One of the pathways, known as the regular mechanism of response to cytokinins (Hwang & Sheen, 2001; Kakimoto, 1996), involves increased cytokinin production, activation of a cytokinin receptor and a phosphorylation relay.

Cytokinins are plant hormones (phytohormones) that promote cell division and differentiation in various developmental processes (Hwang *et al.*, 2012). Naturally occurring cytokinins are N^6-substituted adenine derivatives (Fig. 1). Until recently, cytokinin perception was believed to be an extracellular process, implying that cytokinins do not cross the plasma membrane prior to the activation of the pathway. The recent detection of cytokinin receptors in the endoplasmatic reticulum (ER) has shed new light on cytokinin signalling (Caesar *et al.*, 2011; Lomin *et al.*, 2011; Wulfetange *et al.*, 2011). The diffusion of cytokinins within the plant cell is still elusive and there may be factors that mediate their translocation.

The proteins that are involved in nodulation are collectively termed nodulins. Gamas *et al.* (1998) reported that nodulin 13 from *Medicago truncatula* (MtN13), which is expressed only in the nodule cortex at the early stages of nodulation, belongs to the pathogenesis-related proteins of class 10 (PR-10). This class of proteins has been extensively characterized in a recent review (Fernandes *et al.*, 2013). PR-10 proteins are small (up to 19 kDa), monomeric, slightly acidic cytosolic plant-specific proteins that are expressed in response to various biotic and abiotic stress factors and are believed to be the key element of plant defence systems. Based on sequence homology, PR-10 proteins are divided into three subclasses. The first subclass includes classic PR-10 proteins that are induced by pathogens, and MtN13 belongs to this subclass. Members of the second subclass are found in latex and are called major latex proteins. Cytokinin-specific binding proteins (CSBPs) comprise the third subclass despite a very low level of sequence identity to the other groups. All PR-10 proteins have the same overall fold, which consists of a seven-stranded antiparallel β-sheet and three α-helices. A large hydrophobic cavity, which naturally suggests itself as a ligand-binding site, is formed between the β-sheet and the longest, C-terminal, α-helix α3. The hydrophobic cavity of PR-10 proteins has been shown to bind various ligands (Kofler *et al.*, 2012; Fernandes *et al.*, 2008, 2009; Pasternak *et al.*, 2006, 2008; Marković-Housley *et al.*, 2003). This cavity is believed to be the key element that determines the functional relevance of each PR-10 protein (Fernandes *et al.*, 2013).

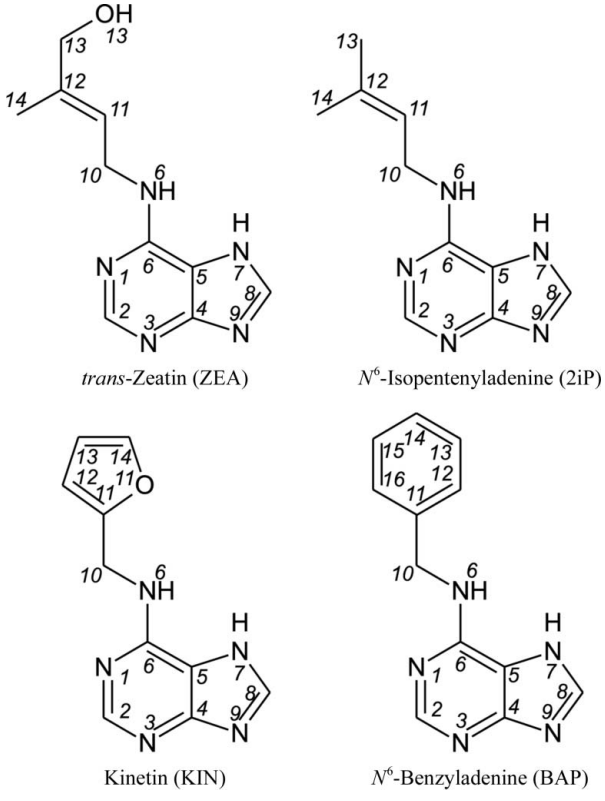

Figure 1
Chemical structures of four naturally occurring cytokinins. The correct atom-numbering scheme is shown. Please note that the PDB is enforcing a chemically incorrect numbering of zeatin atoms (Jaskolski, 2013). A number in superscript (N^6) indicates an exocyclic adenine atom.

Some PR-10 proteins have been found to bind cytokinin molecules (Kofler *et al.*, 2012; Fernandes *et al.*, 2008, 2009; Pasternak *et al.*, 2006). The most perplexing conclusion from these studies is that there is no unique mode of ligand binding or even constant stoichiometry and that the same protein can even bind the same ligand in several different but precisely defined ways (Pasternak *et al.*, 2006). The present work is focused on the question of whether the nodulation-specific MtN13 protein is also capable of binding cytokinins. Four adenine-type naturally occurring cytokinins, namely *trans*-zeatin (ZEA), N^6-isopentenyladenine (2iP), kinetin (KIN) and N^6-benzyladenine (BAP), were tested. Cocrystallization of MtN13 with these phytohormones yielded crystals that diffracted X-rays to high angles, allowing very accurate characterization of the complexes. In all cases, one ligand molecule was found in the hydrophobic cavity of MtN13 with excellent definition in the electron-density maps. Moreover, all four cytokinins were bound in the same manner. In addition to hydrophobic interactions, multiple hydrogen bonds are also formed between the protein and the hormone molecule. One of these interactions is formed with the participation of a loop element from another MtN13 molecule, which enters the binding cavity, sealing off the access route. In consequence, this interaction leads to the formation of a homodimeric MtN13–cytokinin complex. PR-10 dimerization with a proposed biological role has not previously been observed.

2. Materials and methods

2.1. Cloning, overexpression and purification of MtN13

The gene coding for MtN13 was found in the *M. truncatula* genome at transcript level by *Protein BLAST* (Altschul *et al.*, 1997). The MtN13 sequence has an 'Evidence at protein level' annotation in the UniProt Knowledgebase and is accessible

under code P93330. The MtN13 DNA-coding sequence was amplified by the polymerase chain reaction (PCR) using *M. truncatula* (ecotype J5) cDNA from root nodules as the template (the PCR primers used were forward, CACCATGGGTGTTATCACTTCAGAAAGC, and reverse, TTAGTTGCTGTCTTTGTTGTAATCAG). The reaction product was cloned into the pET-TOPO-151D vector (Invitrogen) and the correctness of the insert was confirmed by sequencing.

Overexpression was carried out in BL21 Magic *Escherichia coli* cells. The bacteria were cultured with shaking at 210 rev min^{-1} in LB medium supplemented with 120 µg ml^{-1} ampicillin and 25 µg ml^{-1} kanamycin at 310 K until the OD_{600} reached 1.0. The temperature was lowered to 291 K and isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 m*M*. The culture was grown for 18 h and then centrifuged at 4500 rev min^{-1} for 10 min at 277 K. The cell pellet from 1 l of culture was resuspended in 30 ml binding buffer [50 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl, 20 m*M* imidazole, 1 m*M* tris(2-carboxyethyl)phosphine (TCEP)], flash-frozen in liquid nitrogen and stored at 193 K. The samples were thawed and the cells were disrupted by sonication using bursts of total duration 4 min with appropriate intervals for cooling. Cell debris was removed by centrifugation at 15 000 rev min^{-1} for 30 min at 277 K. The supernatant was applied onto a column packed with 6 ml HisTrap HP resin (GE Healthcare) connected to a Vac-Man (Promega) and the chromatographic process was accelerated using a vacuum pump. After binding, the column was washed four times with 30 ml binding buffer and the purified protein was eluted with 15 ml elution buffer (50 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl, 300 m*M* imidazole, 1 m*M* TCEP). The His_6 tag was cleaved with TEV protease and excess imidazole was simultaneously removed by dialysis (4 h at 277 K). The solution was mixed with HisTrap HP resin to remove the His_6 tag and the His_6-tagged TEV protease. The flowthrough was

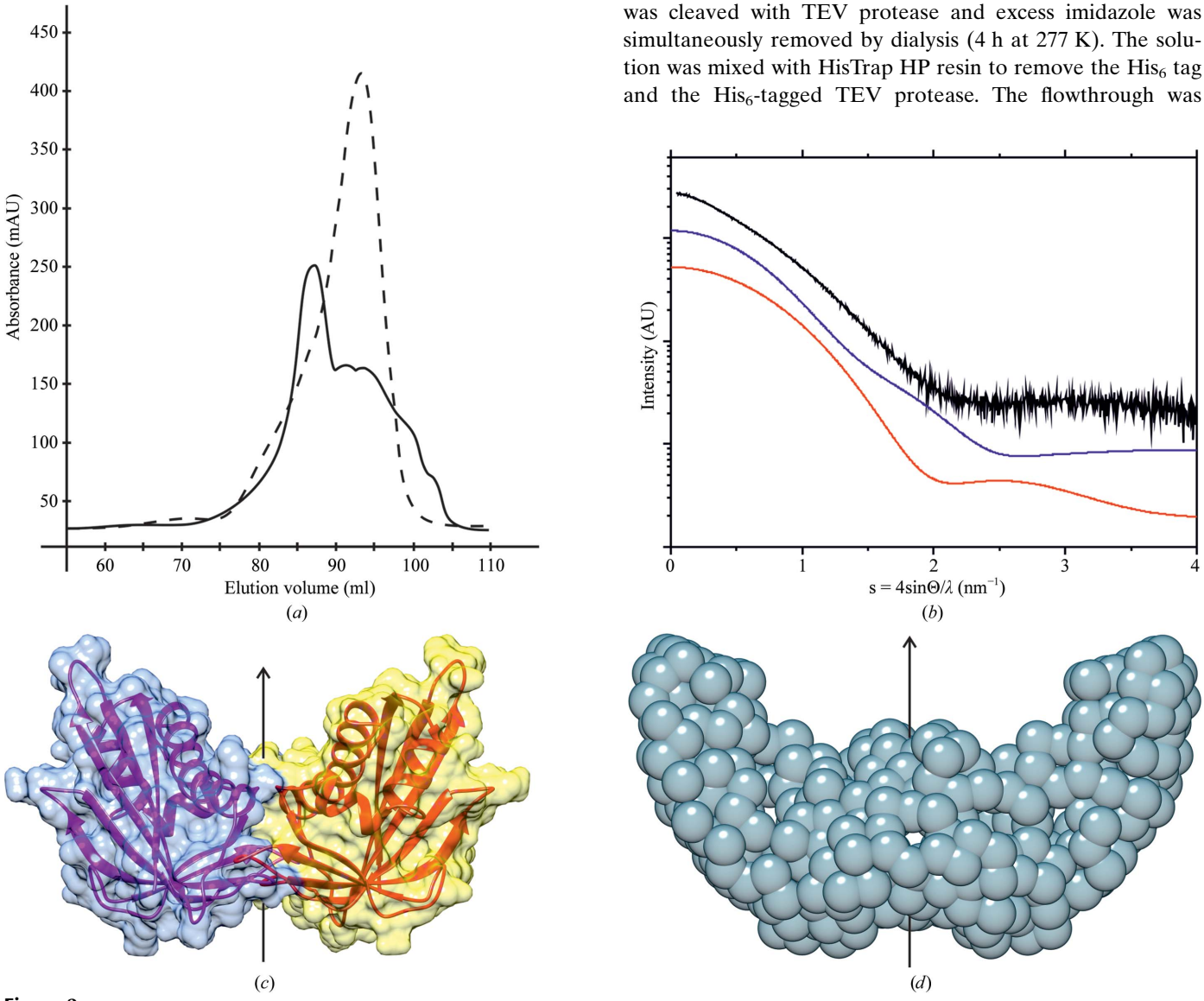

Figure 2
Dimerization of MtN13. (*a*) Size-exclusion chromatography (SEC) of the MtN13 protein. The curves show the elution profiles of protein samples applied to the SEC column immediately after the preceding chromatographic step (dashed line) or after a 24 h delay (continuous line). (*b*) Comparison of the experimental SAXS curve for MtN13 (black) with theoretical SAXS scattering curves calculated for the monomer (red) and dimer (blue) using the crystallographic coordinates from the present study. (*c*) The crystallographic twofold-symmetric dimer of MtN13 with a semitransparent surface. The twofold axis is indicated. (*d*) A model of MtN13 in solution generated by *ab initio* calculations using SAXS data. The scale is the same as in (*c*).

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Complex with	ZEA	2iP	KIN	BAP
Data collection				
Radiation source	APS, ANL Argonne	PETRA III, DESY Hamburg	BESSY, Berlin	MAX-lab, Lund
Beamline	22-BM	P14	BL14.1	I911-2
Wavelength (Å)	1.0000	1.22343	0.91801	1.04172
Temperature (K)	100	100	100	100
Space group	$P6_222$	$P6_222$	$P6_222$	$P6_222$
Unit-cell parameters (Å)	a = 96.1, c = 113.2	a = 96.6, c = 112.0	a = 96.4, c = 113.4	a = 96.0, c = 113.4
Oscillation range (°)	0.5	1	0.5	1
No. of images	180	120	100	90
Resolution (Å)	30.0–1.85 (1.92–1.85)	48.3–2.04 (2.16–2.04)	48.2–2.20 (2.33–2.20)	46.9–2.60 (2.76–2.60)
Reflections (collected/unique)	281527/26831	287179/20324	94999/16292	105014/10018
Completeness (%)	99.8 (99.4)	99.7 (98.2)	99.0 (94.4)	99.8 (99.6)
Multiplicity	10.5 (8.9)	14.1 (14.0)	5.8 (4.9)	10.5 (10.7)
R_{merge}† (%)	7.5 (98.1)	9.9 (70.6)	9.4 (64.7)	12.8 (68.4)
$\langle I/\sigma(I)\rangle$	9.5 (2.1)	24.3 (4.0)	14.5 (2.0)	17.0 (4.3)
Refinement				
Unique reflections (work + test)	26803	20321	16286	10017
Test reflections	1349	1004	1002	802
No. of atoms (non-H)				
Protein	1305	1303	1290	1274
Cytokinin/malonate	16/7	15/7	16/7	17
Metal (Na)	3	6	1	1
Solvent	121	146	94	25
R_{work}/R_{free} (%)	18.2/21.6	16.5/19.8	18.6/21.5	17.0/22.5
R.m.s.d. from ideal geometry				
Bond lengths (Å)	0.019	0.019	0.020	0.015
Bond angles (°)	1.6	1.4	1.7	1.6
Ramachandran statistics (%)				
Favoured	96.4	96.4	96.9	98.1
Outliers	0.6	0.0	0.6	0.0
PDB code	4jhg	4gy9	4jhh	4jhi

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of observation i of reflection hkl.

collected, concentrated to 4 ml and applied onto a HiLoad Superdex 200 16/60 column (GE Healthcare) equilibrated with a buffer composed of 50 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl, 1 m*M* TCEP. The size-exclusion chromatography step yielded a homogenous protein fraction of about 6 ml. The entire purification procedure (from thawing the cell pellet to ligand addition) had to be finished within 10 h, as even with a 24 h delay the protein exhibited a significant level of dimerization (Fig. 2*a*). The sample was concentrated to 7 mg ml^{-1} as determined by the method of Bradford (1976) and used immediately in crystallization experiments.

2.2. Crystallization and data collection

Owing to differences in their solubility, the ligands were added to the protein solution using two different protocols. In both cases the protein–ligand mixture was incubated overnight prior to crystallization. *Trans*-zeatin hydrochloride (20 mg ml^{-1}; Sigma–Aldrich catalogue No. Z2753) was added as an aqueous solution in a threefold molar excess. The remaining hormones (2iP, KIN and BAP) were added in pulverized form in submilligram quantities to 300 µl protein solution. It is of note that addition of these ligands as solutions in organic solvents invariably led to either precipitation of the protein or to failure of the subsequent crystallization experiments. Also, the ligands had to be added immediately after the final size-exclusion chromatography step or the protein would not crystallize. Following overnight incubation, the protein–ligand solutions were centrifuged at 14 000 rev min^{-1} for 5 min at room temperature to remove undissolved ligand. Initial screening for crystallization conditions was performed using a robotic sitting-drop vapour-diffusion setup (Mosquito) with Crystal Screen HT, Index HT and SaltRx reagents from Hampton Research. All four complexes crystallized from conditions containing sodium malonate at pH 7.0. Manual optimization of the crystallization hits in hanging drops gave the following malonate concentrations: 1.9 *M* for the complex with ZEA, 1.5 *M* for that with 2iP, 1.85 *M* for that with KIN and 1.7 *M* for that with BAP. The hanging drops consisted of a mixture of 2 µl reservoir solution and 2 µl protein–ligand solution. Hexagonal bipyramidal crystals grew at 292 K to typical dimensions of approximately 0.50 × 0.40 × 0.40 mm for the ZEA complex, 0.40 × 0.30 × 0.30 mm for the 2iP complex, 0.25 × 0.20 × 0.20 mm for the KIN complex and 0.15 × 0.10 × 0.1 mm for the BAP complex within a week. The crystals were transferred to reservoir solutions with an elevated concentration of sodium malonate for controlled dehydration. Dehydration was carried out stepwise (0.1 *M* per day) and the final sodium malonate concentration was 0.3 *M* higher in each case compared with the crystallization condition. After one month, the crystals were harvested with 0.1 mm nylon loops (Hampton Research) and vitrified in liquid nitrogen for synchrotron-radiation data collection. Additional cryosolutions were not used since sodium malonate at concentrations above 1.8 *M* is expected to provide sufficient cryoprotection (Bujacz *et al.*, 2010). X-ray diffraction data were collected at 100 K at four different synchrotron facilities (Table 1). The diffraction data for the MtN13–ZEA complex were processed using the *HKL*-2000 package (Otwinowski & Minor, 1997) while *XDS* (Kabsch, 2010) was used in the remaining cases. Statistics of data collection and processing are summarized in Table 1.

Attempts to crystallize MtN13 without a ligand or in complex with other plant hormones (such as synthetic cytokinins, auxins, gibberellic acid, abscisic acid, salicylic acid, jasmonic acid and methyl jasmonate) failed (no crystals were obtained).

2.3. Determination and refinement of the crystal structures

The crystal structure of MtN13 in complex with zeatin was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007). The crystal structure of a PR-10 protein from *Lupinus luteus* (PDB entry 2qim; Fernandes *et al.*, 2008) with 45% sequence identity was used as a search probe. Automatic model building was carried out with the online version of *ARP*/*wARP* (Langer *et al.*, 2008). The protein chain from the MtN13–ZEA complex was used as the initial model for refinement of the other complexes, which are all isomorphous. *Coot* (Emsley *et al.*, 2010) was used for manual fitting in electron-density maps between rounds of model refinement in *phenix.refine* (Adams *et al.*, 2010). Riding H atoms for the protein chain were included in the complexes with ZEA, 2iP and KIN. For these three models, TLS parameters (Winn *et al.*, 2001) were refined for six, five and six groups, respectively, as suggested by the refinement program. The models were validated with *MolProbity* (Chen *et al.*, 2010). The refinement statistics are listed in Table 1.

2.4. Stereochemical restraint libraries for the ligands

The structure-factor refinement of all of the models used maximum-likelihood (ML) targets and was carried out under the control of stereochemical restraints imposed with weights that led to root-mean-square deviations (r.m.s.d.s) from standard geometry for bonds of within 0.02 Å (Jaskolski *et al.*, 2007). For the protein part, the stereochemical library was as defined by Engh & Huber (1991). The geometrical restraints for the cytokinin molecules were generated in *phenix.elbow* (Moriarty *et al.*, 2009) using small-molecule targets from the Cambridge Structural Database (CSD; Allen, 2002). Specifically, CSD structures with the following accession codes were used: DUTZOH10 (Soriano-Garcia *et al.*, 1987) for ZEA, XACJIV for 2iP, KINTIN10 (Soriano-Garcia & Parthasarathy, 1977) for KIN and CATBII (Raghunathan *et al.*, 1983) for BAP. Restraints were only applied to covalent bonds, angles and planar groups, while torsion angles were left unrestrained and thus could refine freely.

2.5. Experiments in solution

Protein samples for experiments in solution were purified as described in §2.1 and stored at 277 K for at least 24 h prior to measurements. Dynamic light-scattering (DLS) experiments were performed using a Zetasizer μV (Malvern Instruments) and an MtN13 sample at 1 mg ml^{-1} concentration. Small-angle X-ray scattering (SAXS) data (Fig. 2*b*) were collected from protein solutions at 5 mg ml^{-1} concentration on the BioSAXS beamline BM29 of the storage ring at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). All SAXS measurements were performed at 288 K for scattering vectors $0.05 < s < 5.0$ nm^{-1} ($s = 4\pi\sin\theta/\lambda$, where θ is the scattering angle and λ is the X-ray wavelength). Integration, scaling and buffer subtraction were accomplished using the program *PRIMUS* (Petoukhov *et al.*, 2012). The program *CRYSOL* (Svergun *et al.*, 1995) was applied for the evaluation of the solution scattering patterns using the present MtN13 crystal structure. *Ab initio* modelling with twofold symmetry restraints was performed with *GASBOR* (Svergun *et al.*, 2001).

2.6. Other software used

Molecular figures were created with *UCSF Chimera* (Pettersen *et al.*, 2004). Assignment of secondary-structure elements was based on the *DSSP* algorithm (Kabsch & Sander, 1983). The *PISA* server (Krissinel & Henrick, 2007) and *Intersurf* (Ray *et al.*, 2005) were used to calculate the surface area buried upon homodimer formation. The surface of the protein internal cavity was generated with *SURFNET* (Laskowski, 1995). Structural alignments based on C$^{\alpha}$ atoms were performed using *UCSF Chimera* and *PDBeFold* v.2.55 (Krissinel & Henrick, 2004). Torsion angles were calculated with *Mercury* 3.1 (Macrae *et al.*, 2008).

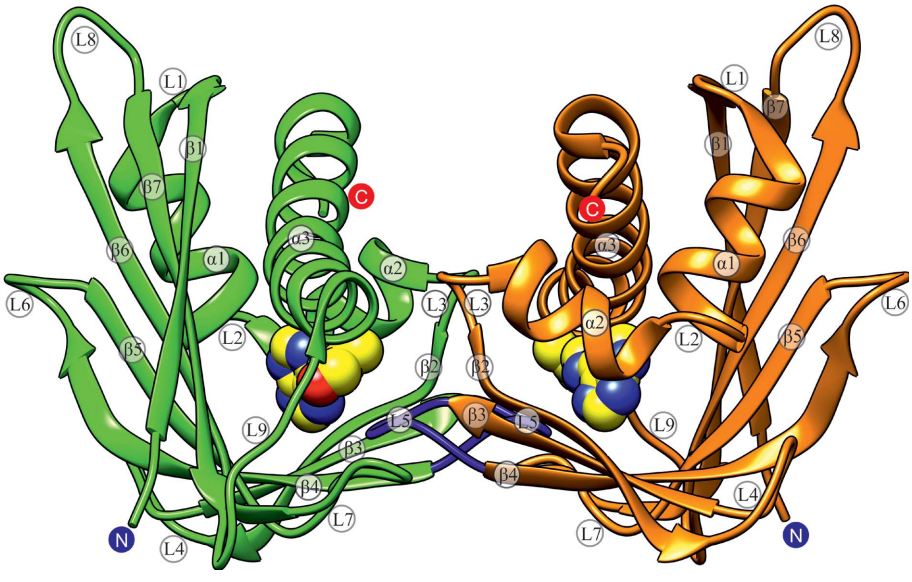

Figure 3
An overall structure of the crystallographic twofold-symmetric dimer of MtN13 (cartoon with subunits coloured green and orange) in complex with *trans*-zeatin (ZEA), which is shown in van der Waals representation in each binding cavity. The L5 loops from both subunits, seen as plugs in the cavities of the complementary subunits, are shown in blue. The N- and C-termini are labelled blue and red, respectively. Secondary-structure elements (α-helices, β-strands and loops) are marked as assigned by the *DSSP* algorithm (Kabsch & Sander, 1983).

3. Results and discussion

3.1. Overall features of the MtN13–cytokinin complexes

All four crystal structures of the MtN13 complexes with cytokinins were solved in space group $P6_222$. In each case there is one protein molecule in the asymmetric unit located close to the dyad in the diagonal direction, which

creates a tight crystallographic dimer. The hydrophobic cavity of each subunit is occupied by one ligand molecule (Fig. 3).

Despite obvious differences resulting from the different data resolutions (from 1.85 to 2.6 Å), the electron-density maps obtained were very clear and allowed the tracing of almost all of the residues of the MtN13 sequence. Only five residues at the C-terminus (NKDSN) were impossible to trace in each structure owing to disorder. On the other hand, five (IDPFT) and four (DPFT) additional residues preceding the genuine MtN13 protein sequence, introduced as cloning artifacts, were modelled at the N-terminus in the complexes with ZEA, 2iP and KIN and in the complex with BAP, respectively. The cytokinin ligands had excellent definition in difference electron-density maps and could be modelled without any ambiguity (Figs. 4*c*–4*f*). The refined atomic displacement parameters (ADPs) for each cytokinin molecule are similar to or lower than those for the protein side chains of the surrounding residues. In the complexes with ZEA, 2iP and KIN a malonate anion (from the crystallization buffer) was also modelled next to Gly87. Three, six, one and one sodium cations were modelled in the ZEA, 2iP, KIN and BAP complexes, respectively. One Na^+ cation [$Na^+(1)$] stabilizes the

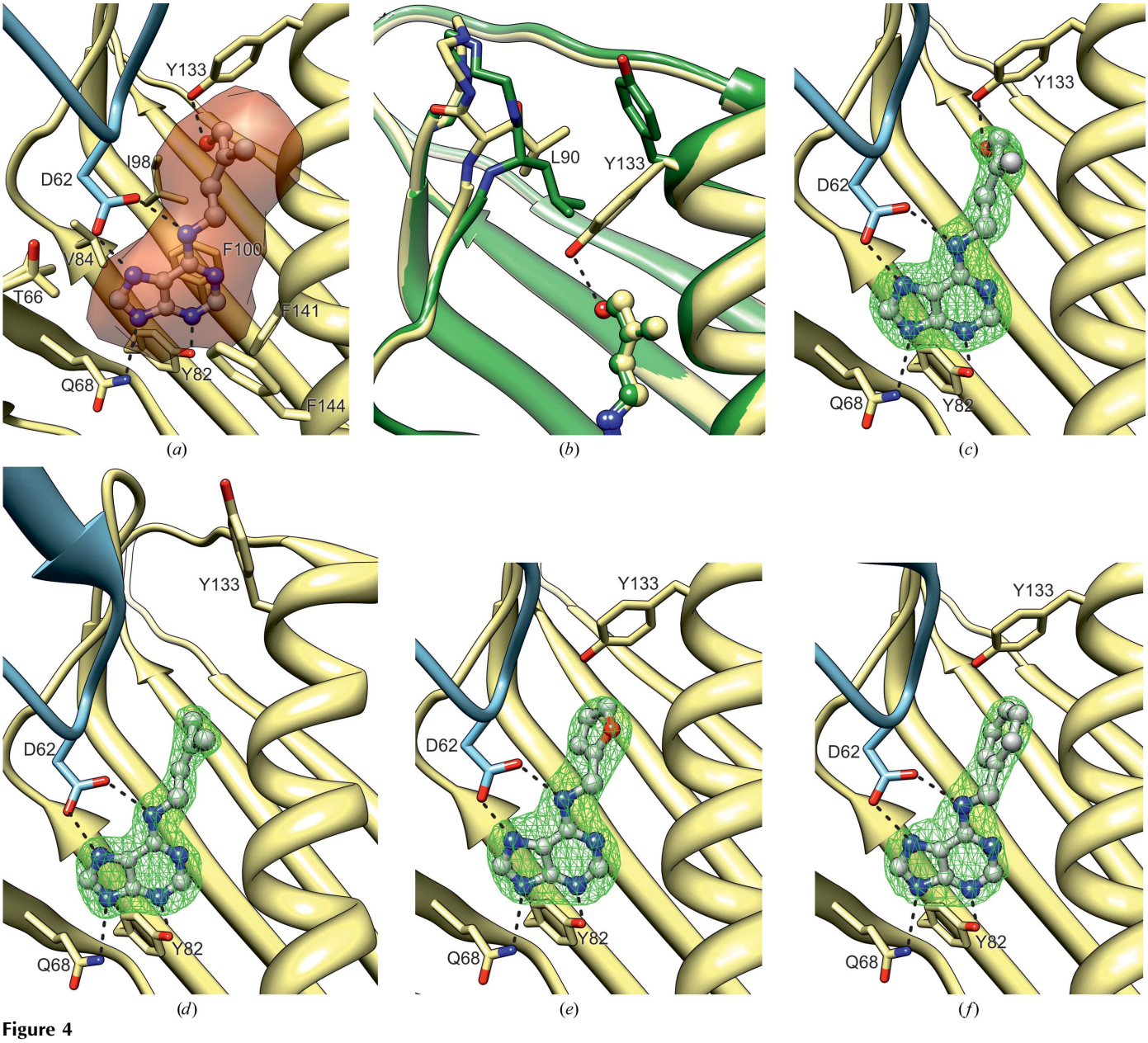

Figure 4
The cytokinin-binding site of the MtN13 protein. (*a*) The internal cavity (represented by a semitransparent surface) is formed by amino-acid residues creating either hydrogen bonds or hydrophobic interactions with the cytokinin molecule, here illustrated by *trans*-zeatin. Asp62 is contributed by the second protein molecule of the homodimer (blue). (*b*) Structural adaptation of the cavity to different cytokinin N^6-substituents (tails). Note the flip of Tyr133 and the rearrangement of the L7 loop (marked by Leu90) on the replacement of ZEA (yellow) by 2iP (green). (*c*)–(*f*) ZEA (*c*), 2iP (*d*), KIN (*e*) and BAP (*f*) bound in the internal cavity of MtN13. The ligands (in ball-and-stick representation) are shown with OMIT $F_o - F_c$ electron-density maps contoured at 4.5σ.

L5 loop and is in the same position in each structure. Two additional Na^+ cations found in the complexes with ZEA and 2iP occupy the same sites, next to Ser65 [$Na^+(2)$] and to Thr0, preceding the genuine Met1 [$Na^+(3)$]. Three additional Na^+ cations are found in the 2iP complex only. The metal-cation identification was based on the octahedral coordination and metal···O distances, which were shorter (∼2.2–2.5 Å) than typical hydrogen bonds. In addition, the presence of Na^+ cations was confirmed by the calcium bond-valence sum (CBVS) method (Müller *et al.*, 2003).

Overall, the backbone of the MtN13 protein has the canonical PR-10 fold (Fig. 3) consisting of a seven-stranded antiparallel β-sheet wrapped around the C-terminal helix α3. The β-strands are connected by β-hairpins, except for the β1–β2 crossover joining the edges of the β-sheet, which is formed by helices α1 and α2. The highly curved baseball-glove shape of the β-sheet is the result of six β-bulges. The β-hairpins and loops form finger-like structures at both ends of the β-sheet. In particular, the odd-numbered loops L3, L5, L7 and L9 resemble the 'fingers' of the glove. The two short helices (α1 and α2) create a V-shaped support for the C-terminal part of the long α3 helix, which is connected to the β7 strand of the β-sheet by loop L9. In its middle, helix α3 is slightly kinked towards the protein core.

A large hydrophobic cavity is formed between the α-helices and the β-sheet. All of the cytokinin molecules studied in this project are found inside the cavity. In the topology of the fold of the MtN13 molecule, the only prominent entrance to the cavity leads through a narrowing tapered tunnel. The entrance is bordered by loops L5 and L7 and the N-terminal part of helix α3. When unplugged, the entrance has a diameter of about 8 Å.

3.2. Dimerization of MtN13

In general, PR-10 proteins are monomeric in solution. Even though some exceptions have been reported (Ma *et al.*, 2006; Schöll *et al.*, 2005), they were neither examined from a structural point of view nor was the functional importance of the oligomerization indicated. It is therefore very exciting to observe tight MtN13 dimers in all of the presented crystal structures. The molecular surface area buried on dimerization is very large at ∼4000–5000 Å^2 per dimer, which is very impressive, especially in comparison with the surface area of the dimer itself (∼16 000 Å^2). The crystallographic dimer formation involves a mutual insertion of the tip of loop L5 into the internal cavity of the partner molecule (see below).

To verify the crystallographically observed dimerization of MtN13, two additional experiments were performed in solution, namely dynamic light scattering (DLS) and small-angle X-ray scattering (SAXS). The DLS measurements for both the unliganded and liganded form of MtN13 confirmed a particle mass of 37.6 ± 11.8 kDa, which corresponds to a homodimer. The hydrodynamic radius R_h obtained from the DLS experiment is 25.4 Å. Subsequently, SAXS experiments were performed on the unliganded MtN13 protein solution. The radius of gyration R_g of unliganded MtN13 calculated from the SAXS experimental curves (Fig. 2*b*) and extrapolated to zero concentration is 24.3 Å. Theoretical calculations based on the crystal structure of MtN13 (in complex with ZEA) give values

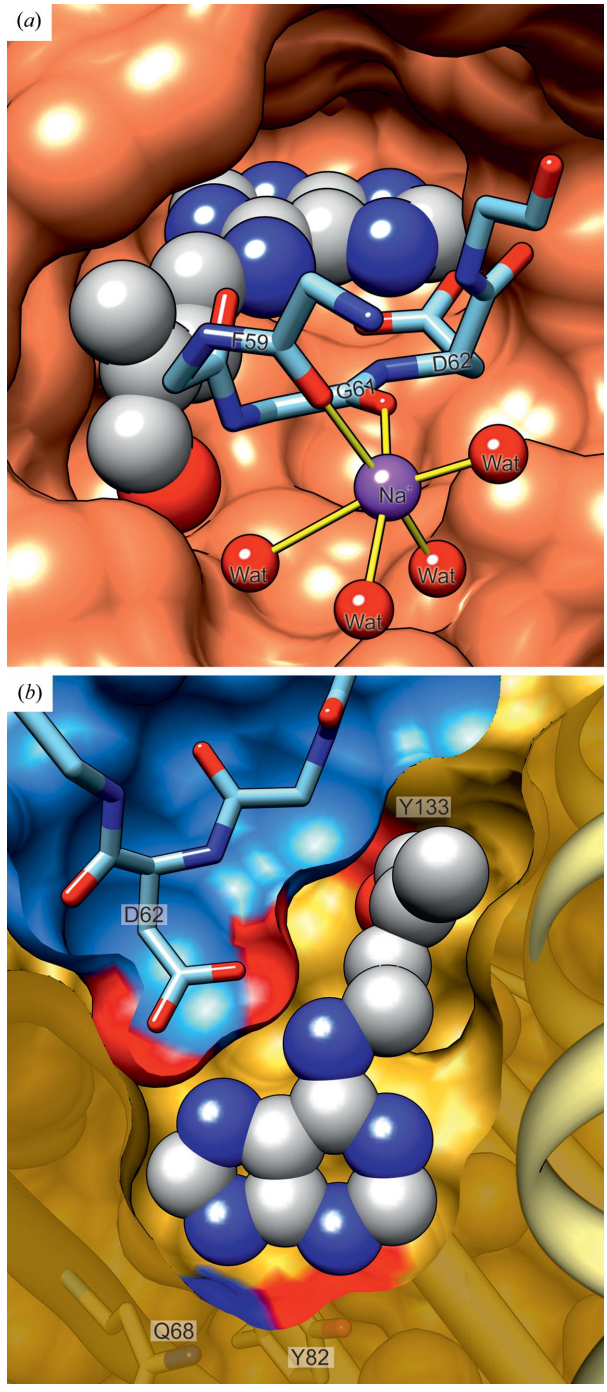

Figure 5
Loop L5 (sticks) in the crystal structures of MtN13 (surface representation). The tracing of the L5 loop is represented by its backbone atoms. The *trans*-zeatin molecule from a ZEA complex is shown in van der Waals representation, with spheres coloured grey for C atoms, blue for N atoms and red for O atoms. (*a*) The internal cavity of an MtN13 subunit viewed from the perspective of the penetrating L5 loop of the other subunit. Coordination bonds of a sodium cation [$Na^+(1)$] are shown as yellow sticks. (*b*) The L5 loop (with the Asp62 side chain shown in sticks) penetrating the internal cavity of the other subunit of the MtN13 homodimer. The blue surface corresponds to the L5 loop.

for the radius of gyration of 16.5 Å for the monomer and 21.8 Å for the homodimer. The theoretical value estimated for the homodimer is very close to the R_g obtained experimentally and the small discrepancy (2.5 Å) can be explained by the shell of hydration in solution and differences in molecular compactness induced by crystal packing and/or the presence of a bound ligand. The approximate relation between R_h and R_g for elongated particles (Berry, 2010), given by the equation $R_g = 0.8R_h$, is obeyed very well for the case of the MtN13 protein in solution.

The SAXS experiments also provided information about the shape of the MtN13 protein in solution. The obtained $p(r)$ function (Svergun & Koch, 2003) suggests that in solution the protein is an elongated ellipsoid with a maximum diameter of 75 Å. A comparison between an *ab initio* model calculated from the SAXS scattering curves and the homodimeric crystal structure is shown in Figs. 2(*c*) and 2(*d*). The general agreement is satisfactory and the more obtuse angle between the two halves of the SAXS model (compared with the crystallographic model) can be attributed to crystal-packing forces and to the effect of ligand binding.

In conclusion, MtN13 is also a dimer in solution. In the presence of the cytokinin molecules the dimer is formed by two protein molecules and two cytokinin ligands and thus the complex stoichiometry is best described as 2:2. However, dimer formation seems to be independent of cytokinin binding. Further experimental work is needed to examine the oligomeric state of MtN13 in the cytoplasm. At the very least, the physiological concentrations of both MtN13 and cytokinins during the nodulation phase in *M. truncatula* should be determined.

The L5 loop in the MtN13 topology is of particular interest (Fig. 5). It is five residues long ($_{60}$VGDFG$_{64}$) and connects strands β3 and β4. It forms an inverse γ-turn between Gly61 and Phe63, meaning that Asp62 is the central residue of the loop apex. The conformation of loop L5 is stabilized by a sodium cation with octahedral coordination. Almost all of the bonds between the metal cation and the electron donors are shorter than 2.5 Å. The main-chain O atoms of Phe59 and Gly61 together with water molecules are the ligands in the $Na^+(1)$ coordination sphere. The coordination octahedron (six O ligands) is complete in the two highest resolution structures: those with ZEA (1.85 Å resolution) and 2iP (2.0 Å). In the complexes with KIN (2.2 Å) and BAP (2.6 Å), respectively, one and three water molecules are defined by only weak peaks of positive electron density and are missing from the coordinate set. It is of note that in the 2.6 Å resolution structure of the BAP complex only a very limited number of water molecules (25 in total) were modelled. However, the $Na^+(1)$ cation was also modelled in the BAP complex by analogy to the better resolved structures of the other complexes.

The L5 loops are mutually exchanged by the two MtN13 molecules that form the homodimer. Topologically, L5 penetrates the binding cavity of the second subunit through the only prominent entrance, located in its own neighbourhood in the MtN13 topology. The two L5 loops that are swapped from one subunit to the other therefore pass each other in an antiparallel fashion, as required by the twofold symmetry of the dimer. The loops interact with each other *via* water-mediated hydrogen bonds between the main-chain O atom of Asp62 and the backbone N atom of Ser65 and the carbonyl O atom of Phe63 from the second protein molecule. The loop penetrates the second subunit so deeply that the Asp62 side chain at its apex creates a fork of two hydrogen bonds with the N^6 and N7 atoms (a number in superscript indicates an exocyclic adenine atom) of a cytokinin molecule docked at the bottom of the cavity. In this way, loop L5 acts as an intermolecular plug that seals off the only entrance to the cavity and nails the cytokinin ligand down in its binding site.

The L5 'plug' is secured in its 'socket' *via* numerous interactions. Asp62 is the only residue with a polar side chain in the L5 loop. Apart from binding the cytokinin ligand, the side chain of Asp62 also forms a hydrogen bond to Thr66 from the other protein subunit. There are no more direct hydrogen bonds formed by the L5 loop; however, a few water-mediated hydrogen bonds are formed between the L5 loop and the other protein molecule. The data resolution limits the total number of water molecules that are visible in the electron-density maps. For this reason, the water/sodium-mediated interactions at the interface between the 'plug' and the 'socket' are described below for the MtN13–ZEA and MtN13–2iP complexes, which have the highest resolutions. The carbonyl O atom of Val60 forms a water-mediated hydrogen bond to both N^η atoms of Arg140. Other interactions of this type are formed between two water molecules from the coordination sphere of the loop L5 $Na^+(1)$ cation (coordinated by the backbone O atoms of Phe59 and Gly61) and the carbonyl O atoms of Gly87 and Ala89. One more docking interaction is mediated by another Na^+ cation adjacent to the L5 loop. This $Na^+(2)$ cation is coordinated by the carbonyl O atom of Asp62 from the 'plug' and the carbonyl O atom of Ser65 and the hydroxyl O atom of Thr66 from the 'socket'. Hydrophobic contacts also play a role in the interaction between the L5 loop and the other protein molecule. Specifically, the side chain of Val60 interacts with Ala36 and the aromatic ring of Phe63 interacts with the aromatic rings of Phe38 and Phe59.

3.3. The internal cavity of MtN13 as a cytokinin-binding site

The internal cavity is the most fascinating element of the PR-10 structure. The PR-10 cavities are formed mainly by hydrophobic residues in the core of the protein. A few polar side chains invariably point into the cavity, allowing the creation of selective hydrogen bonds to suitable heteroatom partners of the ligand molecules. This is also the case for the MtN13 complexes, in which several hydrophobic residues, namely Val84, Ile98, Phe100, Phe141 and Phe144, form the walls of the cavity (Fig. 4*a*). Gln68, Tyr82 and (in the complexes with ZEA, KIN and BAP) Tyr133 interact with the cytokinin molecules *via* hydrogen bonds. The volume of the cavity of the MtN13 dimer is reduced by the presence of the L5 loop from the second protein molecule. As a result, in the MtN13–ZEA complex the available volume of the cavity is

only 270 Å^3, which is close to the volume of 186 Å^3 (190 Å^3 including H atoms) calculated for an isolated *trans*-zeatin molecule. The volumes of the cavities of the other MtN13–cytokinin complexes are (values in parentheses refer to the volume of an isolated ligand molecule without/with H-atom contributions) 2iP, 300 Å^3 (179/181 Å^3); KIN, 250 Å^3 (174/177 Å^3); BAP, 280 Å^3 (188/194 Å^3). The cavity is the largest in the complex with 2iP owing to a different orientation of the Tyr133 side chain (see below). Another small cavity of about 80 Å^3 is found deep in the protein core. This small chamber is not connected to the main cavity and is occupied by two water molecules.

Two of the cytokinins used in this study, ZEA and 2iP, have substituents at the adenine N^6 atom (referred to as tails) that despite apparent similarity are in fact quite different in their chemical character. Specifically, the 2iP tail lacks the terminal hydroxyl group and is therefore highly hydrophobic. Structural comparison of the MtN13 complexes with ZEA and 2iP reveals important differences in the organization of the upper part of the cavity (Fig. 4*b*). With *trans*-zeatin bound in the cavity, the Tyr133 side chain is flipped towards the hormone molecule and forms a hydrogen bond to its terminal O atom. On the contrary, in the complex with 2iP Tyr133 points away from the cytokinin ligand, whereas Leu90 from the adjacent L7 loop is much closer (C$^\alpha$ shift of ∼2 Å) to the tail of 2iP compared with the ZEA complex. This means that the cavity can rearrange its functional elements in order to adapt their configuration for optimal binding of cytokinin ligands with different tails.

The four complexes of MtN13 with cytokinin ligands are shown in Figs. 4(*c*)–4(*f*). Except for the rearrangement of Tyr133 and loop L7 on 2iP binding, as described above, only minor tuning of the molecular environment is observed when the other cytokinins are bound. In all four complexes, Gln68 and Tyr82 form hydrogen bonds to the adenine N9 and N3 atoms, respectively. The hydrogen bonds to the side chain of Asp62 from the complementary protein molecule of the homodimer are strictly conserved. Owing to these interactions, the adenine framework of all the cytokinin molecules is fixed in the same orientation in all four complexes. The Tyr133 side chain points towards the cytokinins that have an aromatic group in the tail, as observed for KIN and BAP (Figs. 4*e* and 4*f*). This Tyr133 conformation is stabilized by a weak C—H···O hydrogen bond involving a C—H donor from the furanyl ring (KIN) or by a C—H···π interaction between the benzyl ring of BAP and the aromatic electrons of Tyr133.

The tautomeric form of the cytokinin ligands bound in the cavity of MtN13 can be determined by reference to the hydrogen-bond networks in which they participate. Usually, adenine derivatives are considered to bear an H atom attached to the N9 atom of the purine ring. However, this is not the case in the MtN13 complexes. Theoretically, the fork of hydrogen bonds to the adenine N^6 and N7 atoms formed by the side chain of Asp62 (from the complementary protein molecule) could use either a carboxylate (with N7 protonated) or a carboxylic (with N7 unprotonated) form. However, the latter case is highly unlikely considering the pH (7.0) of the crystallization buffers. The N9 atom cannot be protonated as it is the acceptor of an N$^\varepsilon$—H···N9 hydrogen bond from the side-chain amide of Gln68 in all of the complexes. The orientation of the Gln68 side chain was confirmed by careful analysis (and test calculations with the amide group flipped) of the atomic *B* factors (ADPs). Thus, the system of hydrogen-bond interactions is logically consistent with the assumption that N7 is protonated and N9 is not. Another interaction is formed by the adenine N^3 atom, which is obviously the acceptor of a hydrogen bond from the hydroxyl group of Tyr82. It is quite intriguing that the adenine N1 atom is not involved in any hydrogen-bond interactions in any of the complexes. This may be explained by the orientation of the N^6 substituent (tail), which in all cases is *cis*-oriented with respect to N1 and therefore shields this atom from intermolecular interactions with bulkier groups. It is noted that *trans*-zeatin was added to the incubation solution in a cationic (protonated) form as a salt. However, there is no evidence to suggest that the ligand in the MtN13–ZEA complex is protonated.

Considering the dimerization of the MtN13 protein and the shape of the internal cavity, one can postulate a mechanism of cytokinin binding. As suggested by the light-scattering and X-ray scattering studies of ligand-free solutions of MtN13, the presence of a cytokinin molecule is not a condition *sine qua non* for the homodimer to form. However, once it has been formed there is no entrance to or exit from the internal cavity. The MtN13 molecules that have dimerized without a cytokinin ligand most probably lose their binding properties because it seems unlikely that the dimers can dissociate back to the monomers, at least when the protein concentration is high, as in the light/X-ray scattering experiments. It is thus evident that cytokinin binding must precede or occur simultaneously with the dimerization event. In the former scenario, the ligand would be bound in the cavity of a monomeric protein first and two monomeric complexes would then dimerize, sealing each other's cavity with their L5 loops. In the second scenario, a cytokinin ligand would be recognized by the N^6,N7 synthon of hydrogen bonds with Asp62, and such a transient assembly would be inserted into the ligand-binding cavity of a dimerization partner in a mutual fashion. There is of course one more possible scenario, namely that the dissociation constant of the $(MtN13)_2$ dimer is high in the absence of a ligand and becomes drastically reduced on the presence of cytokinin ligands in the binding cavities. In this scenario, the monomer–dimer equilibrium could be shifted to the left (especially at low MtN13 concentration) and would shift to the right only in the presence of cytokinin ligands. However, this is merely speculation and its verification would require precise measurements of *in planta* physiological concentrations of both the MtN13 protein and its cytokinin ligands.

3.4. Comparison of the MtN13 structure with other PR-10 homologues

The four crystal structures of MtN13 in complexes with cytokinins described in this paper are isomorphous. The topology of the α-helices and the β-sheet is strictly conserved.

Table 2
Structural superpositions of selected PR-10 proteins with the present MtN13 structure, represented by the protein chain from the MtN13–ZEA complex.

The superpositions were calculated with *PDBeFold* v.2.55 (Krissinel & Henrick, 2004).

Protein structure	PDB code (chain ID)†	R.m.s.d. of C^{α} atoms (Å)	No. of aligned residues	Sequence identity (%)	*Q* score‡
VrCSBP–ZEA	2flh (*A*)	1.45	135	22	0.63
LlPR-10.1B	1ifv (*A*)	1.59	155	39	0.69
LlPR-10.2B–ZEA	2qim	1.77	149	44	0.64
LlPR-10.2B–DPU	3e85	1.80	146	44	0.62
Bet v 1a	1bv1	1.80	152	44	0.65
LlPR-10.1A	1icx	1.88	155	38	0.64

† If more than one protein chain was present in the crystal structure, only the chain with the lowest r.m.s.d. was used in the analysis. ‡ The *Q* score represents the quality function of C^{α} alignment. It reduces the effect of the r.m.s.d./$N_{\rm algn}$ (number of aligned residues) balance on the estimation of alignments: $Q = (N_{\rm algn}\cdot N_{\rm algn})/\{[1+(\text{r.m.s.d.}/R_0)^2]\cdot N_{\rm res1}\cdot N_{\rm res2}\}$, where $N_{\rm res1}$ and $N_{\rm res2}$ represent the number of residues in the aligned proteins and the empirical parameter R_0 is set to 3 Å.

The protein chains in all of the MtN13 complexes are very similar, as illustrated by the low root-mean-square deviations of their C^{α} superpositions (~0.2–0.3 Å). Unless specified otherwise, in the following structural comparisons the four MtN13 models are represented by the MtN13–ZEA structure, which has the highest resolution.

Table 2 shows structural alignments of representative PR-10 proteins with the present MtN13–ZEA structure characterized by r.m.s.d. values of their C^{α} superpositions. The r.m.s.d. values are quite high but similar (1.45–1.88 Å). A detailed analysis of selected structural elements of these comparisons is presented below. The β-sheet of classic PR-10 proteins is known to be a very conserved structural element. This is also the case for MtN13 when compared with the structures of other PR-10 proteins. For instance, in the crystal structures of the yellow lupin protein LlPR-10.2B in complexes with ZEA (Fernandes *et al.*, 2008) and *N*,*N*′-diphenylurea (DPU; Fernandes *et al.*, 2009), a water molecule was found to disrupt the canonical hydrogen-bonding pattern between strands β5 and β6 close to loop L7, mediating the backbone interactions between the O and N atoms of Val85 and Lys96, respectively. In the present MtN13 complexes, there is a similar motif with a water molecule sealing a rupture in the hydrogen-bonding pattern between the separating strands β5 and β6 with a hydrogen bond between the main-chain O atoms of Tyr85 and Glu96. This interaction certainly has a bearing on the conformation of loop L7, which as the nearest neighbour of the dimerization loop L5 must make room for the subunit–subunit interaction through loops L5.

Among the numerous loops in the PR-10 topology, the glycine-rich loop L4 has been most comprehensively analyzed (Fernandes *et al.*, 2008, 2009, 2013; Pasternak *et al.*, 2006). The sequence signature of loop L4, also conserved in MtN13, is EG*x*GG*x*GT, which is similar to the phosphate-binding loop of nucleotide-binding proteins (Saraste *et al.*, 1990). Nucleotides, however, are not among the molecules for which the PR-10 proteins have high affinity (Koistinen *et al.*, 2005). The L4 loop is highly conserved in the PR-10 proteins not only in sequence but also from the structural point of view, which is unusual considering the high content of Gly residues in its structure. Typically, the rigid L4 conformation is stabilized by hydrogen bonds between the side chain of a Thr residue in its sequence and the main-chain atoms. The rigidity and structure of loop L4 is also preserved in the MtN13 models. The loop is very well defined in electron density and its atoms have low ADP values. To date, no exact function has been proposed for the L4 loop of PR-10 proteins. The β-sheet itself and loop L4 are the most conserved structural elements of the PR-10 fold.

The odd-numbered loops L3–L9 of the PR-10 structure surround the main entrance to the internal cavity. In the PR-10 structures studied to date only loops L3 and L9 were reported to be involved in metal coordination. In most cases, a single sodium cation interacted with either of these loops. There are, however, two exceptions: (i) a calcium cation coordinated by the L3 loop in the crystal structure of the LlPR-10.2B–ZEA complex (Fernandes *et al.*, 2008) and (ii) two sodium ions coordinated by both loops in the LlPR-10.2B–DPU complex (Fernandes *et al.*, 2009). MtN13 is the first example of a PR-10 protein with loop L5 involved in metal-cation coordination. The presence of the $Na^{+}(1)$ cation causes deformation of loop L5, forcing it to bend away from the protein core. It is thus possible that the binding of the sodium cation is correlated with the dimer-formation process, in which loop L5 plays a prominent role. It is interesting to note that one of the additional sodium cations [$Na^{+}(3)$], found near Thr0 in the ZEA and 2iP complexes, is also coordinated by a main-chain carbonyl O atom (Gly124) from the L9 loop; however, this site is different from the L9 coordination site in the previous structures.

The long C-terminal helix α3 is the most variable PR-10 element because of both very low sequence conservation and a very high degree of structural variability. At the same time, the α3 helix is a crucial component of the PR-10 fold and is responsible for shaping the internal cavity. The α3 helix can be either straight or kinked towards the protein core, thus changing the volume of the internal cavity. The helix can also be rotated or/and shifted along its axis and can have a different angle of approach at the L9 loop. All these structural factors as well as the sequence of the helix residues lead to a modulation of the size and chemical character of the cavity. Fernandes *et al.* (2008) suggested that it is ligand binding that causes the straightening of helix α3. The authors compared a PR-10 protein from yellow lupin, LlPR-10.2B, in complex with ZEA with a highly identical (91%) homologue from the same organism, LlPR-10.2A, crystallized without a ligand (Pasternak *et al.*, 2005). The strong kink towards the protein core in the middle of the α3 helix was not observed in the LlPR-10.2B–ZEA complex, making the cavity significantly larger compared with the unliganded structure of LlPR-10.2A. Unfortunately, there is no crystallographic study of the same PR-10 protein in forms both with and without a ligand. The recently reported crystal structures of the most studied PR-10 protein, the birch pollen allergen Bet v 1 (Kofler *et al.*, 2012), do not solve the problem either. Although the authors were

able to determine the structure without the ligand of their interest, ingredients of the crystallization cocktail were still present inside the cavity, making a reliable analysis of the helix α3 conformation impossible. Fig. 6(*a*) shows the structural

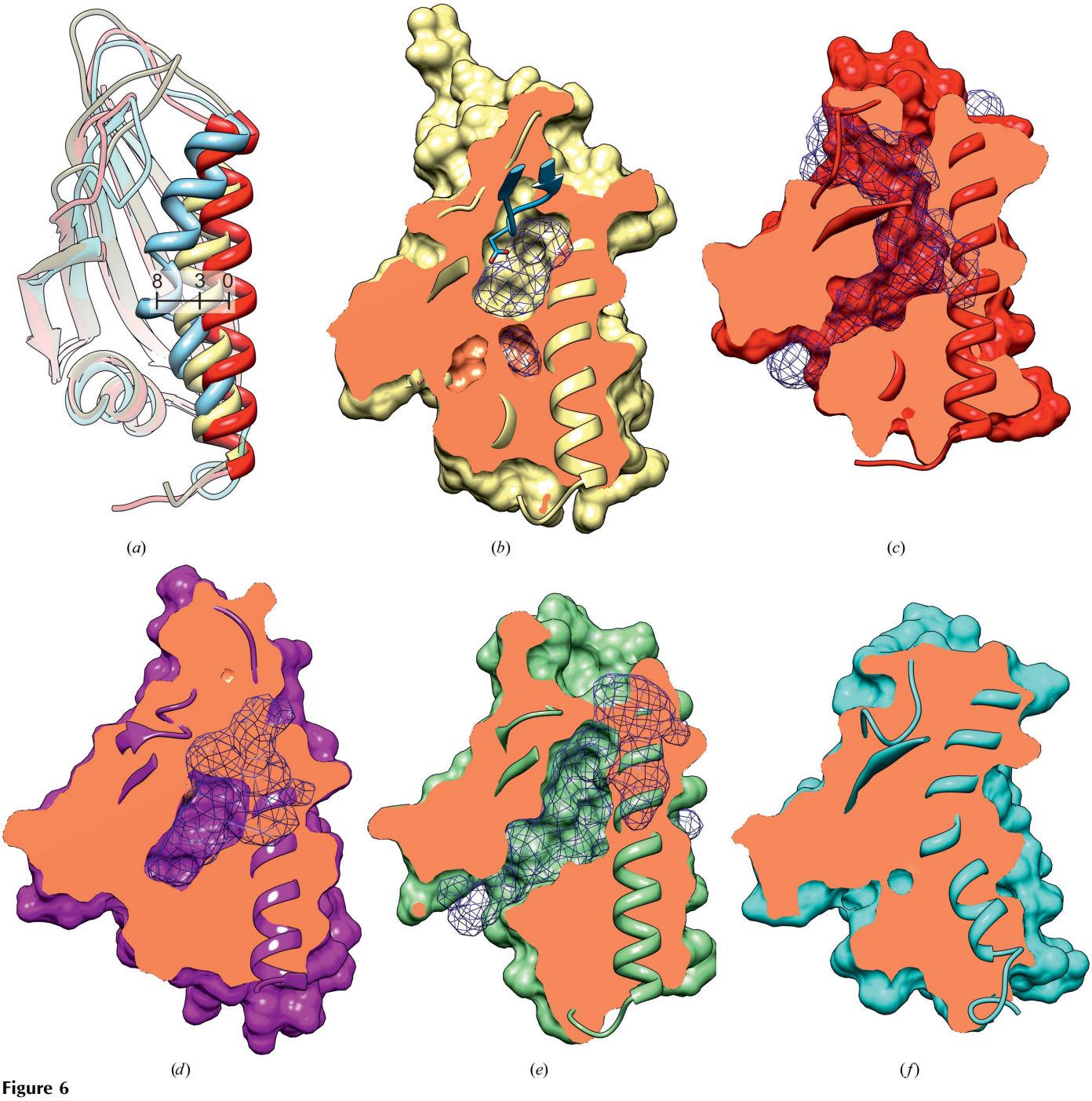

Figure 6
Structural analysis of MtN13 and other PR-10 proteins in the context of their internal cavities. Helix α2 as well as strands β2 and β3 have been omitted for clarity. (*a*) Comparison of the inward kink of helix α3. The three cases are illustrated by the crystal structures of (i) MtN13 in complex with ZEA (this work; yellow), (ii) LlPR-10.2B in complex with ZEA (PDB entry 2qim; red) and (iii) LlPR-10.2A without any ligand in the protein cavity (PDB entry 1xdf; blue). All secondary-structure elements except for helix α3 are semitransparent. Note that in the LlPR-10.2A structure the α3 helix is bent so strongly that the hydrogen-bond pattern is distorted and in fact it should be regarded as two L-shape-forming helices. The approximate indentations of the helices toward the lumen of the cavity are given in Å. (*b*)–(*f*) The internal cavities of five representative PR-10 proteins: (*b*) MtN13–ZEA complex (this work) with the cavity surface including elements from the second protein subunit (shown in blue), cavity volume 270 + 80 Å^3; (*c*) LlPR-10.2B in complex with ZEA (PDB entry 2qim), cavity volume 2440 Å^3; (*d*) VrCSBP in complex with ZEA (PDB entry 2flh, chain *A*), cavity volume 950 Å^3; (*e*) Bet v 1 in complex with KIN (PDB entry 4a85), cavity volume 1900 Å^3; (*f*) the case of LlPR-10.2A without any ligand (PDB entry 1xdf), no solvent-accessible cavity. The cavity volumes were calculated with *SURFNET* (Laskowski, 1995) and displayed as a three-dimensional blue mesh. A 1.5 Å gap sphere radius was used for the calculations to represent solvent-accessible areas. Solvent atoms, ions, ligands and H atoms were disregarded in the calculations. The protein surfaces have been clipped in a cut-away mode to show a maximum vista of each cavity.

alignment of three PR-10 proteins in the context of helix $\alpha3$ deformation. The $\alpha3$ helix of the MtN13 protein has a moderate kink (an ~3 Å deviation from a straight line in the middle) compared with the nearly straight helix reported for LlPR-10.2B in complex with ZEA. The inward kink of the helix in MtN13 is ~5 Å less pronounced than in the LlPR-10.2A structure without a ligand (~8 Å). As a result of the bending of helix $\alpha3$, the internal cavity of the MtN13 protein is small compared with other complexes of PR-10 proteins. The volume of the cavity of the MtN13 dimer is additionally restricted by the L5 loop from the other subunit. The internal cavities of five PR-10 proteins, including MtN13, are presented in Fig. 6. In the case of MtN13 there is only one entrance to the internal cavity, which, as mentioned above, is plugged upon MtN13 dimer formation. In contrast, the PR-10 proteins with large cavities have up to three entrances, as observed, for example, in the LlPR-10.2B–ZEA complex (Fernandes *et al.*, 2008).

Table 3
Comparison of the torsion angles (°) defining the conformation of the tails (N^6 substituents) of the cytokinin ligands in the MtN13 (top value; this work) and AHK4 (bottom value; Hothorn *et al.*, 2011) complexes.

The calculations were carried out in *Mercury* 3.1 (Macrae *et al.*, 2008). For cytokinin atom numbering, see Fig. 1.

Torsion angle	Atoms	ZEA	2iP	KIN	BAP
τ_1	N1—C6—N6—C10	7.6	−1.7	−1.6	−21.3
		7.3	8.8	−2.4	9.6
τ_2	C6—N6—C10—C11	124.9	123.4	126.8	131.6
		172.7	164.7	145.8	133.5
τ_3	N6—C10—C11—C12/O11†	−132.3	−141.9	157.3	157.6
		136.2	146.5	65.0	67.3
τ_4	C10—C11—C12—C13	174.1	179.1		
		−172.5	−174.1		
τ_5	C11—C12—C13—O13	−0.7			
		4.3			

† C12 in ZEA, 2iP and BAP or O11 in KIN.

3.5. Comparison of the cytokinin-binding sites in PR-10 proteins

Cytokinins are among the most frequently studied ligands in the crystal structures of PR-10 proteins. To date, three proteins from this family have been shown to bind to these phytohormones. Interestingly, in each case the cytokinin molecules are bound in a totally different manner and even the complex stoichiometry for the same PR-10–cytokinin pair is not conserved (Fernandes *et al.*, 2013). This variability is coupled with an astonishingly excellent definition of the ligand molecules in electron density. To sum up the previous results, LlPR-10.2B can bind either three ZEA or four DPU molecules inside the cavity. Bet v 1a was shown to bind a single KIN molecule in two orientations or both KIN and 8-anilino-1-naphthalene sulfonate (ANS) in the same cavity. In ZEA complexes of *Vigna radiata* CSBP (VrCSBP), one or two cytokinin molecules are found in the internal cavity and the tandem binding can be realised in two ways. The only uniformity is observed for the VrCSBP binding mode of the common inner ligand. In general, however, there is no common pattern of cytokinin binding, as the ligands can occupy different positions and orientations in the cavity.

On this background, the unique and conserved cytokinin-binding site of the MtN13 protein is a distinct novelty. Analysis of the amino-acid residues forming hydrogen bonds to the cytokinin molecules in other crystal structures of PR-10 complexes reveals that the binding mode is highly variable. There is no residue involved in cytokinin binding that is universally conserved in the PR-10 family. Involvement of amino-acid residues at the positions corresponding to Gln68 and Tyr82 of MtN13 is conserved among several PR-10 members; however, in each case a different cytokinin atom is recognized by these residues. In MtN13, Gln68 forms a hydrogen bond to the N9 atom of the cytokinin molecule, while in the VrCSBP–ZEA complex Glu69 (at the same structural position) binds to the N7 atom of the ligand. Tyr82 forms a hydrogen bond to either the N3 atom of the adenine ring (MtN13–ZEA) or to the terminal hydroxyl group from the cytokinin tail (LlPR10.2B–ZEA). Thus, neither the orientation nor the position of the cytokinin ligands is preserved in the internal cavity of the PR-10 proteins studied previously and here.

Despite all of the differences and variability, at least one common feature of cytokinin recognition by the PR-10 proteins can be gleaned, namely the recurrent fork-like interaction synthon between the N^6 and N7 atoms of the cytokinin scaffold and an acid or amide side chain. Such interactions occur in the LlPR-10.2B–ZEA complex, in which Asp7 is involved, and in the VrCSBP–ZEA complex, in which Glu69 (in all protein chains *A*–*D*) or Gln67 (chains *A* and *D*) participate in this synthon. Interestingly, in the VrCSBP cases in which Gln67 is involved the interaction implies that the ligand molecule must have the alternate tautomeric form with N9 protonated. One must, however, bear in mind that the fork-like hydrogen-bonding synthon can be realised in the internal cavity in very different ways. In MtN13, which is the most spectacular example, the synthon is created with the participation of loop L5 from the complementary subunit of this unusual PR-10 homodimer.

3.6. Common features of MtN13 and cytokinin receptor

Recently, Hothorn *et al.* (2011) reported a set of crystal structures of the sensor domain of *Arabidopsis thaliana* histidine kinase 4, AHK4, in complexes with different cytokinins. This protein had previously been known as a cytokinin receptor (Kakimoto, 2003). The overall fold of the (monomeric) hormone-binding domain of AHK4 is different from that of MtN13 (or from the PR-10 canon, for that matter); however, the hormone-binding cavity is created in a similar fashion between an antiparallel β-sheet and an α-helix. The binding mode of different natural cytokinins is conserved among all of the studied complexes with the AHK4 sensor domain. This makes MtN13 the second protein with known three-dimensional structure that is able to bind various cyto-

kinins in the same manner. Despite the differences in the location of the cytokinin ligands in the internal cavity, some important similarities between MtN13 and AHK4 can also be noted (Fig. 7). In all complexes of MtN13 and AHK4 the N^6 and N7 atoms of the cytokinin framework are recognized by the fork-like synthon with the participation of Asp residues (Asp62 in MtN13 and Asp262 in AHK4). One must keep in mind, however, that in the complexes of MtN13 the Asp62 residue is provided by the second protein subunit of the homodimeric assembly. The tautomeric form of the bound cytokinins is the same in both structures. In all cases, the N7 atom of the cytokinin must be protonated as this is the only form that logically explains the existing networks of hydrogen bonds.

Another interesting aspect concerns the portion of the cavity that is responsible for binding the cytokinin tail. In the MtN13 complexes Tyr133 is flipped towards the cytokinin tail whenever a binding interaction is possible (in the complexes with ZEA, KIN and BAP; see above). In AHK4 there is also a Tyr residue at a similar position (Tyr318). Surprisingly, in the AHK4 complexes Tyr318 is always flipped away from the ligand molecule, as in the MtN13–2iP complex studied here (Fig. 7*b*). In the AHK4 complexes hydrogen bonds to the cytokinin tails are formed by Thr294. More precisely, the hydroxyl group of Thr294 forms a direct hydrogen bond to the terminal O atom of the *trans*-zeatin tail in the AHK4–ZEA complex or a water-mediated hydrogen bond to the O atom from the furanyl ring of kinetin in the AHK4–KIN complex.

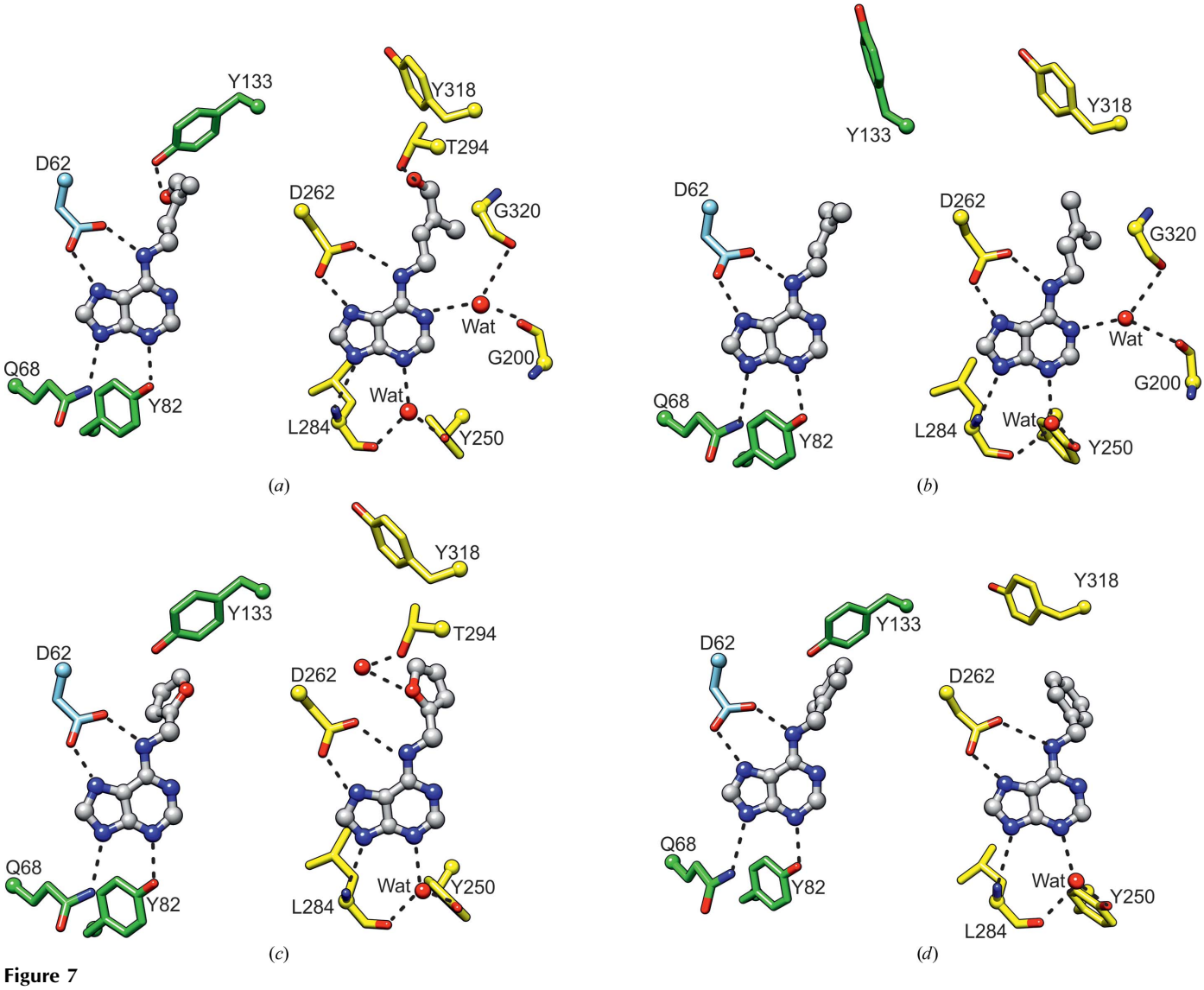

Figure 7
Comparison of the cytokinin-binding mode of MtN13 (present complexes) and of the cytokinin-binding domain of the cytokinin receptor AHK4 studied in complexes with ZEA (PDB entry 3t4l), 2iP (PDB entry 3t4j), KIN (PDB entry 3t4s) and BAP (PDB entry 3t4k) (Hothorn *et al.*, 2011). Each panel (*a*–*d*) compares the structures of MtN13 (left) and AHK4 (right) with the same ligand as follows: (*a*) ZEA, (*b*) 2iP, (*c*) KIN and (*d*) BAP. The side chains of the MtN13 molecule that is the docking site for the ligand are shown in green, while the second molecule of the homodimer is shown in blue. The side chains of the AHK4 protein are shown in yellow. The views are perpendicular to the purine ring of the cytokinins, disregarding any protein alignment. C^{α} atoms are highlighted as balls and water molecules are shown as red spheres. Unless necessary (Gly200, Leu284 and Gly320 of AHK4), the main-chain atoms were omitted for clarity.

3.7. Conformation of the ligand molecules

The geometry of the bound cytokinin ligands, together with the tautomeric/protonation state, is an important aspect of the binding mode of these phytohormones by their protein targets. The best characterization of the cytokinin ligand geometry is obtained by analysis of the torsion angles around all of the covalent bonds of the N^6 tails. It must be emphasized that the present cytokinin models have excellent definition in electron density in all the complexes and that they were refined without restraining any torsion angles. The values for the MtN13 complexes in Table 3 therefore provide a reliable and bias-free picture. A side-by-side comparison of the same ligands in the MtN13 and AHK4 complexes shows that all of the N^6 substituents (tails) are *cis*-oriented with respect to the adenine N1 atom, as defined by the torsion angle N1—C6—N6—C10 (τ_1), which is close to 0°. Even the most divergent pair, noted for BAP (−21.3 and 9.6°, respectively), are still within the *cis* range. No analogous difference is observed for kinetin. There is, however, a systematic ∼90° difference in the rotation τ_3 of the aromatic rings (of BAP and KIN) when the MtN13 and AHK4 complexes are compared. This indicates that the specific molecular environments in the two proteins are slightly different. The torsion angle τ_2 (C6—N6—C10—C11) is remarkably similar (123.4–131.6°) in all MtN13-bound ligands. In the AHK4 series it covers a much wider range, from a *trans* conformation for ZEA (172.7°) to a *gauche* conformation resembling the MtN13 situation for BAP (133.5°). It should be noted that because of the different hybridization states of the pivot atoms (N^6 is sp^2-type because of partial conjugation with the purine system, whereas C10 is sp^3), both conformers have hydrogen-eclipsed steric hindrance in the Newman projection in the *gauche* conformation at C6 and in the *trans* conformation at C11. The τ_3 torsion angle (N6—C10—C11—C12), which in KIN and BAP defines the orientation of the aromatic ring of the tail (see above), has similar absolute values (∼139°) in all complexes with ZEA and 2iP. Here again one of the pivot atoms is sp^3 (C10) and the other is sp^2 (C11). Remarkably, the angles in the MtN13 and AHK4 series have opposite signs. Because the total rotation is less than ∼90°, the ligands still look similar although their conformations are quite distinct (Fig. 8). The configuration around the C11═C12 double bond described by the τ_4 torsion angle is *trans*, as expected, in both ZEA complexes. Despite the small conformational variations within the MtN13 and AHK4 series (Table 3), an overall view shows that the ligands in a given protein place their tails in a very consistent way (Fig. 8), which strongly corroborates the notion that they are docked in very specific binding sites, although the binding sites of the two proteins have their individual characteristics. One might quote the phrase coined by Fernandes *et al.* (2009) that cytokinins as binding partners are also adaptable molecules that can adjust themselves (*via* protonation, tautomerism or conformation) to the requirements of their protein targets.

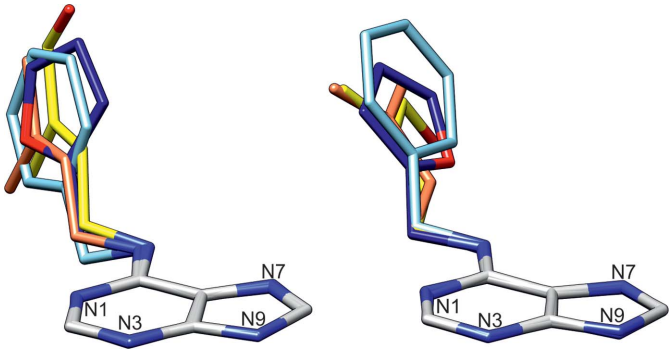

Figure 8
Structural superpositions (based on the adenine ring only) of the four cytokinin molecules from the complexes with MtN13 (left panel, this work) and with the cytokinin-binding domain of the cytokinin receptor AHK4 (right panel; Hothorn *et al.*, 2011). The four cytokinin ligands are ZEA (yellow), 2iP (coral), KIN (dark blue) and BAP (light blue). The adenine rings have grey C atoms. Note that the termini of the cytokinin tails are nearly coplanar within the MtN13 and the AHK4 clusters. However, the principal orientation of the terminal groups of the tails in the two clusters is roughly perpendicular.

There is an interesting observation concerning the placement of the O atom in the tails of ZEA and KIN in the MtN13 and AHK4 complexes. Although the furanyl rings in the two KIN complexes are nearly perpendicular (when the adenine fragments of the two molecules are aligned) and it is not easy to match their atoms, in general the furanyl O atom points towards the N1 fragment of the adenine system in MtN13 and towards N7 in AHK4 (Fig. 8). This is reflected in the different interactions that this atom makes in the two binding sites (see above). The orientation of the furanyl ring in the MtN13–KIN structure is consistent with the atomic *B* factors and patterns of intermolecular interactions. The hydroxyl group of the *trans*-zeatin tail is in both cases *cis* in terms of the τ_5 angle, which places it in the preferred *trans* orientation with respect to the bulky C14 methyl group. It is remarkable that the hydroxyl atom of the ZEA tail in AHK4 is placed in the general area of the furanyl O atom of KIN, highlighting the importance of this point of attachment in the cytokinin-binding site of AHK4. In the MtN13–ZEA complex, on the other hand, the hydroxyl group of ZEA occupies a well defined but different position, indicating that the cytokinin tail-recognition mechanisms of these two cytokinin-binding proteins are not the same. Interestingly, the ZEA molecules found in the binding cavity of VrCSBP had the conformational parameters $\tau_1 \simeq 0°$, $\tau_2 \simeq -65°$, $\tau_3 \simeq -152/-75°$, $\tau_4 \simeq 180°$ and $\tau_5 \simeq 120/0°$, which makes these *trans*-zeatin conformers significantly different from those observed for MtN13 and AHK4, especially with regard to τ_2, which has a totally unique range, as well as to τ_3 and τ_5, which in addition to conformations found in the MtN13–ZEA complex can also assume otherwise unobserved values. These observations reinforce the notion of cytokinin plasticity and (at least with reference to *trans*-zeatin) adaptability to different cytokinin-binding protein targets.

4. Conclusions and outlook

The presented crystal structures of the plant nodulation protein MtN13 in complexes with four natural cytokinins

reveal that this protein is a versatile binder of different cytokinins. More importantly, in each of the complexes the single cytokinin molecule occupies the same binding site and interacts with the same amino-acid residues of the protein. Thus, MtN13 is the first example of a PR-10 protein that has a conserved ligand-binding site. The ligand-binding cavity is compatible, with respect to shape and volume, with a cytokinin ligand. Multiple hydrophobic interactions together with a number of conserved hydrogen bonds suggest strong and specific binding of the cytokinin ligands by the MtN13 protein. The most intriguing aspect of cytokinin binding by MtN13 is the finding that the complexes exist as homodimeric species with a protein:ligand stoichiometry that should be described as 2:2. The dimerization is connected with mutual plugging of the only entrance to the binding cavity by loop L5 from the other subunit. At the same time, Asp62 from the apex of this loop provides the most conspicuous synthon of cytokinin recognition, consisting of tandem hydrogen bonds from the forked Asp carboxylate group to the (obviously protonated) N^6 and N7 atoms of the cytokinin scaffold. This hydrogen-bonding pattern also uniquely defines the tautomeric form of the adenine moiety. Because the dimerization of MtN13 results in plugging of the only exit route from the protein cavity, cytokinin release to the solvent is unlikely without dimer disruption. It will be extremely interesting to search for cellular factors that may influence the MtN13 dimerization equilibrium, thus controlling the capture and release of the cytokinin cargo. One possible factor could be pH change, which, by modifying the protonation pattern of the cytokinin molecule, could disrupt its specific hydrogen-bonding interactions with the protein and, in consequence, eject it from the binding site.

There is a striking analogy between the binding modes of different cytokinins by MtN13 and by the cytokinin receptor protein AHK4. However, there are also important differences, the main difference being that while AHK4 is also a dimer its dimerization is not related to cytokinin binding, in contrast to the MtN13 protein, which must dimerize for productive cytokinin complexation. In addition, the two series of cytokinin complexes emphasize the earlier notion that cytokinins are plastic ligands that are capable of adaptation (*via* mechanisms that involve protonation, tautomerism and conformational flexibility) to their protein targets.

Cytokinins are important signalling molecules that play a key role during the nodulation phase of legume plants. This work unambiguously demonstrates that *M. truncatula* nodulin 13 is a specific and strong binder of cytokinin phytohormones. Since MtN13 is only expressed during nodulation, we may speculate that it regulates nodulation by controlling the levels of freely available cytokinins. Since MtN13 is only detected in the nodules, it is not expected to interfere with other cytokinin-related processes in the plant physiology.

Financial support for this project was provided by the European Union within the European Regional Developmental Fund and by the Polish Ministry of Science and Higher Education (grant Nos. NN 301 003739 and NN 301 204233). X-ray diffraction data for MtN13 in complex with *trans*-zeatin were collected on the Southeast Regional Collaborative Access Team (SER-CAT) beamline of the APS/ANL. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. The *Chimera* package from the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco is supported by NIH grant P41 RR001081.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.
Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Berry, G. C. (2010). *Protein Sci.* **19**, 94–98.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Bujacz, G., Wrzesniewska, B. & Bujacz, A. (2010). *Acta Cryst.* D**66**, 789–796.
Caesar, K., Thamm, A. M., Witthöft, J., Elgass, K., Huppenberger, P., Grefen, C., Horak, J. & Harter, K. (2011). *J. Exp. Bot.* **62**, 5571–5580.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Fernandes, H., Bujacz, A., Bujacz, G., Jelen, F., Jasinski, M., Kachlicki, P., Otlewski, J., Sikorski, M. M. & Jaskolski, M. (2009). *FEBS J.* **276**, 1596–1609.
Fernandes, H., Michalska, K., Sikorski, M. & Jaskolski, M. (2013). *FEBS J.* **280**, 1169–1199.
Fernandes, H., Pasternak, O., Bujacz, G., Bujacz, A., Sikorski, M. M. & Jaskolski, M. (2008). *J. Mol. Biol.* **378**, 1040–1051.
Gamas, P., de Billy, F. & Truchet, G. (1998). *Mol. Plant Microbe Interact.* **11**, 393–403.
Hothorn, M., Dabi, T. & Chory, J. (2011). *Nature Chem. Biol.* **7**, 766–768.
Hwang, I. & Sheen, J. (2001). *Nature (London)*, **413**, 383–389.
Hwang, I., Sheen, J. & Müller, B. (2012). *Annu. Rev. Plant Biol.* **63**, 353–380.
Jaskolski, M. (2013). *Acta Cryst.* D**69**, 1865–1866.
Jaskolski, M., Gilski, M., Dauter, Z. & Wlodawer, A. (2007). *Acta Cryst.* D**63**, 611–620.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Kakimoto, T. (1996). *Science*, **274**, 982–985.
Kakimoto, T. (2003). *Annu. Rev. Plant Biol.* **54**, 605–627.
Kofler, S., Asam, C., Eckhard, U., Wallner, M., Ferreira, F. & Brandstetter, H. (2012). *J. Mol. Biol.* **422**, 109–123.
Koistinen, K. M., Soininen, P., Venäläinen, T. A., Häyrinen, J., Laatikainen, R., Peräkylä, M., Tervahauta, A. I. & Kärenlampi, S. O. (2005). *Phytochemistry*, **66**, 2524–2533.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Laskowski, R. A. (1995). *J. Mol. Graph.* **13**, 323–330.
Lomin, S. N., Yonekura-Sakakibara, K., Romanov, G. A. & Sakakibara, H. (2011). *J. Exp. Bot.* **62**, 5149–5159.
Ma, Y., Gadermaier, G., Bohle, B., Bolhaar, S., Knulst, A., Markovic-Housley, Z., Breiteneder, H., Briza, P., Hoffmann-Sommergruber, K. & Ferreira, F. (2006). *Int. Arch. Allergy Immunol.* **139**, 53–62.

Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466–470.
Marković-Housley, Z., Degano, M., Lamba, D., von Roepenack-Lahaye, E., Clemens, S., Susani, M., Ferreira, F., Scheiner, O. & Breiteneder, H. (2003). *J. Mol. Biol.* **325**, 123–133.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. (2009). *Acta Cryst.* D**65**, 1074–1080.
Müller, P., Köpke, S. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 32–37.
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, A., Biesiadka, J., Bujacz, G., Sikorski, M. & Jaskolski, M. (2008). *Acta Cryst.* D**64**, 595–606.
Pasternak, O., Bujacz, G. D., Fujimoto, Y., Hashimoto, Y., Jelen, F., Otlewski, J., Sikorski, M. M. & Jaskolski, M. (2006). *Plant Cell*, **18**, 2622–2634.
Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. (2012). *J. Appl. Cryst.* **45**, 342–350.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Raghunathan, S., Sinha, B. K., Pattabhi, V. & Gabe, E. J. (1983). *Acta Cryst.* C**39**, 1545–1547.
Ray, N., Cavin, X., Paul, J.-C. & Maigret, B. (2005). *J. Mol. Graph. Model.* **23**, 347–354.
Saraste, M., Sibbald, P. R. & Wittinghofer, A. (1990). *Trends Biochem. Sci.* **15**, 430–434.
Schöll, I., Kalkura, N., Shedziankova, Y., Bergmann, A., Verdino, P., Knittelfelder, R., Kopp, T., Hantusch, B., Betzel, C., Dierks, K., Scheiner, O., Boltz-Nitulescu, G., Keller, W. & Jensen-Jarolim, E. (2005). *J. Immunol.* **175**, 6645–6650.
Soriano-Garcia, M. & Parthasarathy, R. (1977). *Acta Cryst.* B**33**, 2674–2677.
Soriano-Garcia, M., Toscano, R. A. & Arroyo-Reyna, J. A. (1987). *J. Crystallogr. Spectrosc. Res.* **17**, 221–230.
Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
Svergun, D. I. & Koch, M. H. J. (2003). *Rep. Prog. Phys.* **66**, 1735–1782.
Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. (2001). *Biophys. J.* **80**, 2946–2953.
Wasson, A. P., Pellerone, F. I. & Mathesius, U. (2006). *Plant Cell*, **18**, 1617–1629.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.
Wulfetange, K., Lomin, S. N., Romanov, G. A., Stolz, A., Heyl, A. & Schmülling, T. (2011). *Plant Physiol.* **156**, 1808–1818.