

Accepted Manuscript

Crystallographic and CDprobing of ligand-induced conformational changes in a plant PR-10 protein

Joanna Śliwiak, Rafał Dolot, Karolina Michalska, Kamil Szpotkowski, Grzegorz Bujacz, Michał Sikorski, Mariusz Jaskolski

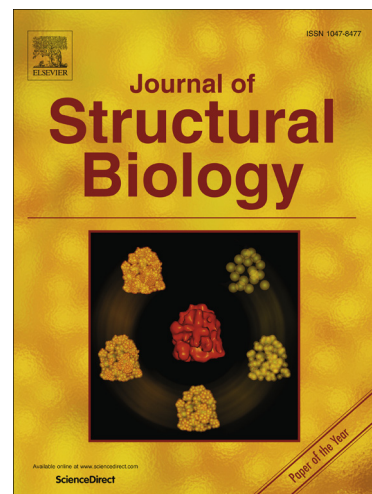
PII: S1047-8477(15)30105-2
DOI: <http://dx.doi.org/10.1016/j.jsb.2015.11.008>
Reference: YJSBI 6818

To appear in: *Journal of Structural Biology*

Received Date: 5 November 2015
Revised Date: 26 November 2015
Accepted Date: 27 November 2015

Please cite this article as: Śliwiak, J., Dolot, R., Michalska, K., Szpotkowski, K., Bujacz, G., Sikorski, M., Jaskolski, M., Crystallographic and CDprobing of ligand-induced conformational changes in a plant PR-10 protein, *Journal of Structural Biology* (2015), doi: <http://dx.doi.org/10.1016/j.jsb.2015.11.008>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Crystallographic and CD probing of ligand-induced conformational changes in a plant PR-10 protein

Joanna Śliwiak¹, Rafał Dolot², Karolina Michalska³ⁱ, Kamil Szpotkowski¹, Grzegorz Bujacz¹,
Michał Sikorski¹, Mariusz Jaskolski^{1,3*}

¹*Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland*

²*Centre of Molecular and Macromolecular Studies, Department of Bioorganic Chemistry, Polish Academy of Sciences, Lodz, Poland.*

³*Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland*

*** Corresponding author:**

Department of Crystallography

Faculty of Chemistry

A. Mickiewicz University

Grunwaldzka 6

60-780 Poznan, Poland

e-mail: mariuszj@amu.edu.pl

tel. +48618291274

fax +48618291505

For submission to: *J. Struct. Biol.*

PDB references: Crystal structure of yellow lupine LIPR-10.1A protein in complex with *trans*-zeatin, 4RYV; Crystal structure of yellow lupine LIPR-10.1A protein in ligand-free form, 4Y31; Crystal structure of yellow lupine LIPR-10.1A protein partially saturated with *trans*-zeatin, 5C9Y;

ⁱ Current address: Argonne National Laboratory, Biosciences Division, Argonne, IL, USA

Abstract

Plant pathogenesis-related class 10 (PR-10) proteins are a family of abundant proteins initially identified as elements of the plant defense system. The key structural feature suggesting PR-10 functionality is a huge hydrophobic cavity created in the protein interior by a scaffold composed of an extended β -sheet wrapped around a long and flexible C-terminal α -helix. Several crystallographic and NMR studies have shown that the cavity can accommodate a variety of small molecule ligands, including phytohormones. The article describes ~ 1.3 Å resolution crystal structures of a *Lupinus luteus* PR-10 isoform LIPR-10.1A, in its free form and in complex with *trans*-zeatin, a naturally occurring plant hormone belonging to the cytokinin group. Moreover we present the structure of the same protein where the saturation with zeatin is not complete. This set of three crystal structures allows us to track the structural adaptation of the protein upon *trans*-zeatin docking, as well as the order of the ligand-binding events, step-by-step. In addition, titration of LIPR-10.1A with *trans*-zeatin monitored in solution by CD spectra, confirmed the pattern of structural adaptations deduced from the crystallographic studies. The ligand-binding mode shows no similarity to other zeatin complexes of PR-10 proteins. The present work, which describes the first atomic models of the same PR-10 protein with and without a physiological ligand, reveals that the conformation of LIPR-10.1A undergoes a significant structural rearrangement upon *trans*-zeatin binding.

Keywords: PR-10 fold, phytohormone, cytokinin, *trans*-zeatin, ligand-binding,

Abbreviations used: 2iP, 6-(γ,γ -dimethylallylamino)purine; ABA, abscisic acid; ANS, 1-anilino-8-naphthalene sulfonate; CD, circular dichroism; CPPU, *N*-(3-chloropyridyl)-*N'*-phenylurea; CSBP, Cytokinin Specific Binding Protein; DPU, diphenylurea; GA3, gibberellic acid; IAA, 3-indoleacetic acid; LIPR, *Lupinus luteus* pathogenesis-related protein; MPD, 2-methyl-2,4-pentanediol; PDB, Protein Data Bank; PEG, polyethylene glycol; PhBP, Phytohormone Binding Protein; PR, pathogenesis-related; R.M.S.D., root mean square deviation; ZEA, *trans*-zeatin;

Introduction

Within the huge superfamily of plant pathogenesis-related (PR) proteins, the members of class 10 (PR-10) are defined as small, slightly acidic and mainly cytosolic proteins. In variance with other PR classes, PR-10 proteins do not have a clearly defined biological function and their classification is often made according to structural similarity, even though it is quite apparent now that not all plant proteins with PR-10 fold and physicochemical properties are linked to pathogenesis mechanisms (van Loon *et al.*, 2006). The key and common feature of the PR-10 fold is the presence of a hydrophobic cavity formed at the interface of the principal secondary structure elements, which are a long and flexible C-terminal α -helix α_3 , and a seven-stranded, antiparallel β -sheet. This folding scheme, resembling the thumb in a clenched fist, is completed by a number of loops (L1-L9) and by two additional short α -helices, α_1 and α_2 , which form a V-shaped support for the C-terminal end of helix α_3 . The cavity has two entrances: E1 formed by helix α_3 and loops L3, L5 and L7, and E2 situated between helix α_3 and strand β_1 (Fig. 1a).

Numerous structural and biophysical studies have demonstrated that the PR-10 cavity has the ability to bind diversified small molecules, leading to the suggestion that the PR-10 fold could be a generic solution to ligand binding utilized in different physiological processes in plants. Indeed, the PR-10 fold is exhibited by enzymes such as S-norcochlorogenic acid synthase (Berkner *et al.*, 2008) or TemN aromatase/cyclase (Ames *et al.*, 2008), hormone receptors such as Pyl1, which is a component of the abscisic acid (ABA) receptor (Miyazono *et al.*, 2009) or the nodulin MtN13 (Ruszkowski *et al.*, 2013), which takes part in cytokinin signaling in nodulating legumes. Other examples include proteins such as the birch pollen allergen Bet v 1 (Kofler *et al.*, 2012), peanut panallergen Ara h 8 (Hurlburt *et al.*, 2013), or yellow lupine isoform LI-PR-10.2B (Fernandes *et al.* 2008; Fernandes *et al.* 2009), with likely storage or transport functions, as they are capable of accommodating numerous copies of chemically diversified ligand molecules in their huge internal cavity. In an attempt to correlate the PR-10 fold with function, Sliwiak *et al.* (2013) proposed a taxonomy of PR-10 proteins according to their cavities. In this classification, PR-10 members that bind ligands in a highly specific manner and are often involved in signaling, have cavity of type 1, which is small and shallow, has only one entrance (E1) and usually accommodates only one ligand molecule. On the other hand, PR-10 proteins that bind chemically different ligands nonspecifically and thus tend to have storage function, possess a huge cavity of type 2, which spans the entire body of the protein interior from E1 to E2, and is capable of accommodating

two or more ligand molecules. Proteins from the latter group bind their ligands in an unusual manner, not only because of the variable number of the cargo molecules, but also because the number of strong, specific and anchoring interactions between the ligand and protein residues is very limited. However, despite this lack of specific interactions, atomic-resolution crystal structures have revealed in a number of cases a perfect order of the different ligand molecules, most often plant hormones.

Notably, plant hormones from the cytokinin group have been reported to be bound by numerous PR-10 proteins. In plant physiology, cytokinins stimulate cell division and additionally control the symbiotic root nodulation in legumes (Hwang *et al.*, 2012). The most common among these adenine N6-derivatives is *trans*-zeatin (Fig. 2). Interestingly, synthetic urea derivatives, e.g. diphenylurea (DPU) or *N*-(3-chloropyridyl)-*N'*-phenylurea (CPPU), show very strong physiological activity almost identical to that of natural cytokinins (Ricci & Bertolotti, 2009) despite no chemical similarity.

PR-10 proteins were first linked to cytokinin binding by Fujimoto *et al.* (1998), who isolated a cytosolic fraction from *Vigna radiata* with high cytokinin affinity. The protein responsible for this interaction, named Cytokinin Specific Binding Protein (CSBP), was later shown to bind cytokinins with lower affinity (Pasternak *et al.*, 2006) than originally reported and the binding mode (within a type 1 cavity) turned out to have a puzzling diversity. Moreover, a recent study revealed that the CSBP proteins bind gibberellin, which is an entirely different phytohormone, with higher affinity and specificity (Ruszkowski *et al.*, 2014). Accordingly, the term Phytohormone Binding Protein (PhBP) has been proposed to replace CSBP as more appropriate. Nevertheless, another PR-10 protein with type 1 cavity, *Medicago truncatula* nodulation protein MtN13, involved in cytokinin signaling (Ruszkowski *et al.*, 2013), was shown to bind cytokinins in a highly specific and reproducible manner.

The PR-10 members that have a large type 2 cavity and are promiscuous in ligand selection are also capable of cytokinin binding, as illustrated by several crystal structures of the birch pollen allergen Bet v 1 in complex with kinetin (PDB IDs 4a85, 4a86) and yellow lupine LIPR-10.2B protein in complex with *trans*-zeatin (2qim). However, no common cytokinin binding pattern can be derived from all these complexes, as the binding modes are characterized by different stoichiometries and protein-ligand interactions. Moreover, the structural studies have emphasized the amazing promiscuity in terms of the chemical nature of the ligands acceptable by these two PR-10 members, as in addition to cytokinins, such

molecules as steroids or flavonoids could form complexes as well (Kofler *et al.*, 2012; Sliwiak, unpublished results).

Despite the large number of crystal structures of PR-10 proteins deposited in the PDB, there is no single case where the same protein would be studied in its free and ligand-bound states. The closest situations one can find for such a pair are presented by Bet v 1 (4a88/4a85) and the Hyp-1 protein from St John's wort (3ie5/4n3e), but the “free” states in both these cases correspond in fact to structures in complex with organic molecules from the crystallization buffers, MPD and PEG, respectively, raising doubts if such structures can indeed serve as templates for the ligand-free conformations. On the other hand, there are several crystal structures of PR-10 proteins in their free form, such as the yellow lupine isoforms LIPR-10.1B (1ifv) and LIPR-10.2A (1xdf), for which no matching complexes have been studied.

In this paper, we describe high-resolution crystal structures of the LIPR-10.1A protein from yellow lupine, which represents the subgroup of PR-10 members with a large type 2 cavity, in complex with *trans*-zeatin (1.38 Å) and in free state (1.32 Å). The latter structure significantly extends the resolution of the model already available in the PDB (1icx, 1.95 Å; Biesiadka *et al.*, 2002). Interestingly, unlike the LIPR-10.2B isoform mentioned above, LIPR-10.1A seems to be much less promiscuous in its ligand “choice”, as despite numerous attempts, it was not possible to obtain crystalline complexes with other ligands. Upon co-crystallization with *trans*-zeatin, the protein internalized three hormone molecules in the binding cavity, and two additional ligand molecules were trapped at the protein surface, where they mediate protein-protein interactions in the crystal lattice. In addition, we also present the structure of an LIPR-10.1A complex unsaturated with *trans*-zeatin (1.50 Å), which can be treated as an “intermediate binding state”. This complement of three high-resolution crystal structures allows us to visualize the conformational rearrangements of the protein that are induced by the docking of the ligand molecules in the binding pocket. These structural transformations were additionally monitored and correlated with the situation in solution using circular dichroism (CD) measurements.

Materials and methods

Crystallization and data collection

LIPR-10.1A was expressed and purified as previously described (Biesiadka *et al.*, 2002), with one important exception. Specifically, the bacteria were grown overnight at 18°C (instead of 37°C), which resulted in the expression of the protein exclusively in the soluble fraction (preventing the formation of inclusion bodies and obviating the need for refolding). Crystallization was carried out using the hanging drop vapor diffusion method at 278 K. For the crystallization of the ligand-free form, protein concentration of 11 mg/ml was used. Single crystals of ligand-free LIPR-10.1A were obtained using as precipitant a solution containing 18% PEG 4K, 0.2 M sodium acetate and 0.1 M HEPES pH 7.0. The crystals were cryoprotected in a 1:1 mixture of the mother liquor and 50% PEG400, and vitrified at 100 K in a cold nitrogen gas stream. Diffraction data extending to 1.32 Å resolution (designated **F**) were collected at beamline BL14.1 of the BESSY synchrotron in Berlin.

Prior to crystallization of the zeatin-saturated complex, the LIPR-10.1A protein at 10 mg/ml was incubated with five-fold molar excess of *trans*-zeatin (ZEA). The incubation mixture was then used for setting up the crystallization drops. Crystals of the LIPR-10.1A/ZEA complex grew over a reservoir solution containing 1.8 M (NH₄)₂SO₄ and 0.1 M MES pH 6.5. Diffraction data extending to 1.38 Å resolution (**Z**) were collected using beamline I911-2 at MAX-lab in Lund. Crystals of a complex unsaturated with *trans*-zeatin were obtained using protein at 20 mg/ml concentration preincubated with a three-fold molar excess of the ligand, and 1.6 M (NH₄)₂SO₄ as the precipitating agent. Diffraction data extending to 1.50 Å resolution (**U**) were collected at beamline X11 of the DORIS storage ring at EMBL/DESY in Hamburg. Crystals of both complexes were cryoprotected using 25% glycerol in the mother liquor and vitrified at 100 K in a cold nitrogen gas stream.

The diffraction data for **F** and **Z** were indexed, integrated and scaled in *HKL2000* (Otwinowski *et al.*, 1997) and for the unsaturated complex **U** in *XDS* (Kabsh, 2010).

Structure solution and model refinement

The present structure **F** of ligand-free LIPR-10.1A is isomorphous with the orthorhombic structure (*P*2₁2₁2₁) reported previously at 1.95 Å resolution by Biesiadka *et al.* (2002) and the coordinates of the protein atoms of the PDB model 1icx could be used directly in the refinement against the new 1.32 Å data. Manual model rebuilding was carried out in *Coot* (Emsley *et al.*, 2010), while for the crystallographic refinement against maximum-likelihood targets the *Refmac5* (Murshudov *et al.*, 1997) program was used.

The crystals of the LIPR-10.1A/ZEA complex **Z** represent a new monoclinic form (C2) whereas the crystals of the unsaturated complex **U** have the same space group as the ligand-free structure. The phase problem was solved by molecular replacement in the *Phaser* program (McCoy *et al.*, 2007), using 1icx as the search model in both cases. Improvement of the initial solution was achieved by manual rebuilding in *Coot* and the final crystallographic refinement was carried out in *phenix.refine* (Afonine *et al.*, 2012). A model of *trans*-zeatin for stereochemical restraint targets was taken from the PDB library. Five molecules of ZEA were modeled in the electron density of the ligand-saturated structure **Z**, whereas the unsaturated structure **U** was refined and deposited in the Protein Data Bank (PDB) without ligand atoms because the corresponding electron density, while clearly indicating the presence of the ligand molecules in the binding cavity, was too fragmentary and ambiguous for reliable modeling of the hormone molecules.

All three structures were refined anisotropically with riding H atoms of the protein molecules included in F_c calculations. The final models were validated with *MolProbity* (Chen *et al.*, 2010). The refinement statistics are given in Table 1. Atomic coordinates and structure factors have been deposited in the PDB with the accession codes 4ryv (**F**), 4y31 (**Z**) and 5c9y (**U**).

Circular Dichroism measurements

Circular dichroism (CD) spectra were recorded on a JASCO J-815 CD spectrometer equipped with a Peltier thermostated cell holder. Protein was dialyzed against 10 mM sodium phosphate buffer pH 7.0, containing 75 mM NaF. Protein samples at the concentration of 0.01 mg/ml were incubated for 1 h with 1-, 2-, ... 10-fold molar excess of *trans*-zeatin before measurements. 600 μ l of protein sample were placed in a 2-mm quartz cuvette for the measurements. Each CD spectrum was the average of three scans at continuous scanning mode, corrected by subtracting the spectrum of the buffer solution at identical conditions. Each scan in the range of 185-340 nm was obtained with a scanning rate of 100 nm/min, 1 nm bandwidth, 0.5 nm data pitch and data integration time of 1 s. The spectra were analyzed using the *DICHROWEB* server (Whitmore & Wallace, 2004, <http://www.cryst.bbk.ac.uk/cdweb/html/> accessed 01.04.2015).

Software

The volumes of the protein internal cavities were calculated using the *SPACEBALL* (Chwastyk *et al.*, 2014) server (<http://www.ifpan.edu.pl/~chwastyk/spaceball/> accessed 20.10.2015). UCSF *Chimera* (Pettersen *et al.*, 2004) and *PyMol* (The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC) were used for structural alignments and for the preparation of the figures. The *ALIGN* program (Cohen, 1997) was used for C α superpositions and R.M.S.D. calculations.

Results and Discussion

Crystal formation

The final crystallization conditions for the LIPR-10.1A/ZEA complex were established after several trials with the commercial crystal screens (Peg/Ion I and II, Crystal Screen I and II; Hampton Res.). Interestingly, these screens were used for testing co-crystallization of three yellow lupine PR-10 isoforms, LIPR-10.1A, LIPR-10.1B and LIPR-10.2B, with several phytohormones and signal molecules including cytokinins (*trans*-zeatin, kinetin, isopentenyladenine (2iP), DPU, CPPU), gibberellin (GA3), auxin (IAA), abscisic acid (ABA), salicylic acid and flavonoids (naringenin, quercetin). While the LIPR-10.2B protein formed crystalline complexes with *trans*-zeatin (2qim), DPU (3e85), CPPU, 2iP, ANS (8-anilinoanthracene-1 sulfonate, a fluorescent dye) and both flavonoids (Sliwiak, unpublished results), LIPR-10.1A formed only a crystalline complex with *trans*-zeatin, and its co-crystallization with other cytokinins invariably resulted in the ligand-free form only. LIPR-10.1B on the other hand, gave only crystals of the free form in all these tests and it was never possible to detect any complex formation.

Overall features of the three crystal structures

LIPR-10.1A is monomeric in solution as confirmed by size-exclusion chromatography and native gel electrophoresis (not shown). Moreover, no stable quaternary structure could be predicted by *PISA* (Krissinel & Henrick, 2007) analysis of both the protein-ligand and ligand-free crystal structures. The crystals of LIPR-10.1A in complex with *trans*-zeatin were created upon crystallization of protein pre-incubated with five- or three-fold molar excess of *trans*-zeatin. The level of ligand saturation also affects the space group of the crystal structure, as the presence of five-fold molar excess of the ligand resulted in the formation of C2 crystals **Z**, whereas three-fold molar excess of the ligand resulted in *P*2₁2₁2₁ crystals of the same

symmetry as the free form but with slightly bigger unit cell (Table 1). The different space group and crystal packing is also reflected in different solvent content, as in the case of **Z** it is 52.5% (with Matthews volume $V_M = 2.59 \text{ \AA}^3/\text{Da}$), whereas in the case of the unsaturated complex **U** it is 35.9% ($V_M = 1.92 \text{ \AA}^3/\text{Da}$), slightly more than for the ligand-free structure **F**, 32.9% ($V_M = 1.83 \text{ \AA}^3/\text{Da}$). In all three forms there is one protein molecule in the asymmetric unit.

All crystals diffracted X-rays to high angles and it was possible to refine the structures at high resolutions of 1.38 Å (**Z**), 1.50 Å (**U**) and 1.32 Å (**F**). Atomic displacement parameters of the protein, ligand and solvent non-H atoms were refined anisotropically for the **Z** and **F** structures, and the same refinement strategy was also used for **U** (with anisotropic ADPs for the protein component only) after confirming that it significantly reduced both R_{work} and R_{free} and as the TLS option was unsatisfactory. Also, anisotropy analysis carried out for **U** using the *PARVATI* server (Merritt, 2012) (<http://skuld.bmsc.washington.edu/parvati/>) did not indicate any abnormalities.

Inspection of the electron density maps of the ligand-saturated complex **Z** phased by the protein component only revealed the presence of five ZEA molecules, three of which (1-3) are bound in the protein interior, while the remaining two (4, 5) are located at the protein surface. Additional elongated patches of positive $F_o - F_c$ electron density were observed in the vicinity of the $\beta 3$, L3, $\beta 4$ and L5 elements of the protein fold but their poor quality was insufficient to allow reliable modeling of any chemical molecules present in the crystallization buffer. The final **Z** model includes one LIPR-10.1A molecule, five *trans*-zeatin molecules, 222 water molecules and one sulfate anion. The electron density of the protein molecule is of a very high quality, allowing unambiguous tracing of the entire polypeptide main chain (residues Gly1-Tyr155) and all side chains (Fig. 3a) without any breaks. The absence of the N-terminal Met0 residue is the effect of the *E. coli* methionylaminopeptidase (MAP) activity in the expression system used to produce the recombinant protein. The electron density of the internal ZEA molecules, especially of ZEA2 and ZEA3, is of superb quality, allowing easy modeling and determination of the ZEA conformation (Fig. 3b). The average B_{eq} values for the ZEA molecules after final refinement are 31, 17 and 19 Å² for ZEA1, ZEA2 and ZEA3, respectively, confirming their very good definition in electron density maps. The electron density of the superficial ZEA molecules is less perfect and indicates a degree of rotation of these molecules and a less fixed position. The model refinement converged with R_{work} and R_{free} of 14.3 and 17.0%, respectively. The high stereochemical quality of the model is

confirmed by *MolProbity* statistics and the distribution of main-chain torsion angles in the Ramachandran plot (Table 1).

In the structure of the unsaturated complex **U**, flat patches of positive $F_o - F_c$ electron density could be observed in the positions corresponding to ZEA1 and ZEA2, with the general shape indicating the presence of the zeatin purine ring and tail (Fig. 3c). However, the quality of these fragmented maps was too poor to allow responsible modeling of the ligands, which very likely populate the ZEA1 and ZEA2 sites at low occupancy. The $F_o - F_c$ maps also revealed residual electron density at the ZEA3 site (Fig. 3c) and in a new superficial position. However in these cases it was not even possible to figure out the ligand orientation. At the conclusion of the refinement of this unsaturated complex **U** it was decided to exclude the ZEA ligands from the coordinate file in spite of the evidence that the ZEA1 and ZEA2 ligands were present in the crystal structure at low occupancy. The corresponding coordinate file 5c9y was deposited in the PDB with the annotation that the structure might contain unmodeled ZEA ligands. The electron density of the protein model **U** allowed uninterrupted tracing of the entire main chain and most of the side chains (Fig. 3a). There is, however, lack of electron density for the following side chains in the $\alpha 3$ helix: Gln133, Lys135 and Phe136, indicating an increased mobility of this region. The final **U** model consists of one protein molecule and 146 water molecules. The model was refined to R_{work} and R_{free} of 18.1 and 24.3%, respectively.

The ligand-free model of LIPR-10.1A presented here (**F**) has significantly improved resolution (1.32 Å) compared to the previous model (1.95 Å) deposited in the PDB (1icx), thereby allowing better analysis of the ordered and disordered regions of the protein in the absence of stabilizing ligands. Indeed, analyzing the electron density one can observe that almost the entire main chain has excellent electron density with significant breaks corresponding to His59-Gly61 in loop L5 and to Ala134-Lys135 in helix $\alpha 3$, the latter break indicating a higher mobility of the middle part of the helix. Electron density is also missing for side chain atoms close to the above-mentioned mobile regions, namely Gln133, Phe136, Lys137 and His62. The final **F** model was refined to R_{work}/R_{free} of 15.3/21.2%. It consists of one protein molecule, 180 water molecules and one acetate anion.

Overall fold and cavities of the three LIPR-10.1A forms

As all other PR-10 proteins, the LIPR-10.1A molecule consists of a seven-stranded (strands $\beta 1$ - $\beta 7$) antiparallel β -sheet and three α -helices ($\alpha 1$ - $\alpha 3$). The long C-terminal helix $\alpha 3$ (residues Leu126-Ala151), connected to the rest of the protein by loop L9, is a centerpiece of the structure, as it is "gripped" by the β -sheet and supported at its C-terminus by the two short accessory helices $\alpha 1$ and $\alpha 2$. The $\alpha 3$ helix appears to be the key element responsible for the adaptation of the structure to the presence of the *trans*-zeatin ligands. In the ligand-free form **F**, it is unwound at the section of Arg131-Leu139 (Fig. 1a, b). In addition, there is a break in the main-chain electron density in the middle of this segment, at residues Ala134-Lys135, and no electron density for the side chains of Gln133-Lys137. The appearance of the electron density maps indicates that the helix is not only unwound but also very mobile or even disordered in this region in the absence of *trans*-zeatin. In the cases of the two ZEA complexes (**Z** and **U**), the $\alpha 3$ helix is straight and with properly preserved secondary structure.

The secondary structure elements are connected by the nine loops L1-L9. Here again in the case of the ligand-free form **F** the position of loop L5 could not be modeled in the electron density at the segment of His59-Gly61. No such problem was encountered in case of the **Z** and **U** complexes.

As mentioned above, the spatial organization of the secondary structure elements leads to the formation of a huge cavity between the $\alpha 3$ helix and the concave face of the β -sheet. The volume and shape of this cavity seem to be dependent on the cargo content as the volume of the largest hilum is the smallest in the free form **F** (754 Å³), medium in the unsaturated complex **U** (1080 Å³) and the biggest in the saturated complex **Z** (1979 Å³). The large difference between the calculated volumes of the unsaturated and saturated complexes arises from the way the cavity volume is estimated by the *SPACEBALL* algorithm (Chwastyk *et al.*, 2014), which probes the void spaces with a fixed-radius sphere. In this method, several small, narrowly connected chambers may be assessed very differently than one large, sphere-like cavity of the same geometrical volume. A comparison of the shapes of the three LIPR-10.1A cavities (Fig. 4) indicates that the void of the unsaturated complex **U** is not fully opened and is separated into smaller sub-volumes by a narrowing at the center. Moreover, the pocket for ZEA3 binding, which in the free form **F** is absent altogether, in **U** is not fully formed and separated from the main cavity. Fig. 4 illustrates that the small cavity volume in **F** is the result of not only the absence of a ZEA3 pocket, but also of the doughnut shape of the cavity, which has a significant central constriction.

The side chains pointing to the lumen of the cavity are mostly non-polar, thus creating an environment suitable for hydrophobic ligands. The cavity has two entrances. One of them, E1, is located between the N-terminus of helix $\alpha 3$ and loops L3, L5 and L7, and is partially gated by the side chain of Lys137 in the saturated complex. The second opening, E2, is formed between the N-terminal fragment of helix $\alpha 3$ and strand $\beta 1$.

***trans*-Zeatin binding and conformation**

The internal trans-zeatin ligands

The hydrophobic cavity of the saturated complex **Z** is occupied by three *trans*-zeatin molecules (ZEA1, ZEA2, ZEA3) which were modeled with full occupancy. The molecules of ZEA2 and ZEA3 are perfectly defined in the electron density (Fig. 3b) and their stability is confirmed by the low average *B*-factors (17 and 19 Å², respectively). ZEA1 is somewhat less well ordered, with an average *B*-factor of 31 Å². It is possible that this ligand molecule could be bound in several slightly different orientations/conformations, but we were unable to model such a disorder, as introduction of any alternative positions of the ZEA1 molecule invariably led to clashes with the protein atoms. Moreover, refinement with ZEA1 in multiple conformations always resulted in a rise of *R*_{free}. For these reasons, the final **Z** model comprises only one set of coordinates for ZEA1.

ZEA1 is located near helices $\alpha 2$, $\alpha 3$ and strands $\beta 3$, $\beta 4$ (Fig. 1a) and has a vista of the E1 entrance (Fig. 5a). It forms three direct hydrogen bonds with the protein molecule (Fig. 5b). Two of them are formed by the hydroxyl group of the aliphatic tail of the ligand molecule with the side chains of Asp27 and Lys53. The adenine moiety of ZEA1 is also anchored to the protein with a hydrogen bond between the N3 atom and the hydroxyl group of one of the two alternative conformations of Tyr82. Additional stabilization of ZEA1 is achieved by van der Waals contacts involving Ile55, His68 and Leu141.

ZEA2 sits very deep in the cavity, at the interface between helix $\alpha 3$ and strands $\beta 6$ and $\beta 7$ (Fig. 1a) where it is completely buried and has no access to any of the entrances (Fig. 5a). It has only one direct hydrogen bond with the protein molecule but in addition forms four contacts mediated by structural water molecules (Fig. 5c). The adenine N9 atom interacts with the O δ 1 atom of Asn7. This interaction indicates that the ZEA2 molecule has a proton attached to the N9 atom, either as a result of protonation or N7/N9 tautomerism. The remaining hydrogen bonds, involving the N1, N6 and N7 nitrogen atoms and the terminal

hydroxyl group of ZEA2, are established via three water molecules one of which is shared by the N6 and N7 atoms and anchors them to the O atom of Lys102. The water molecule interacting with the N1 atom relays this interaction to the O atom of Gly138, whereas the other water molecule, interacting with the ZEA2 hydroxyl group, is ultimately bonded with Leu78 and Ser101. Additionally, Gln9, Leu22, Tyr80, Ser101, Ile115 and Phe142 participate in van der Waals interactions with the ZEA2 ligand.

ZEA3 partially protrudes out of the cavity through the E2 entrance (Fig. 5a). This molecule is directly attached to the protein by only one hydrogen bond, which links the terminal hydroxyl group of ZEA3 and the side chain of Asp132 (Fig. 5d). Additional hydrogen bonds exist between the N3, N6, N7 and N9 atoms and three water molecules. In particular, N3 interacts with a water molecule that relays the interaction to Tyr99. The N6 and N7 atoms again share one hydrogen-bonded water molecule, through which they are connected with the N82 atom of Asn7. The van der Waals contacts of ZEA3 are established with residues Phe5, Asn7, Val117, Phe119, Arg131 and Lys135.

All three *trans*-zeatin molecules bound within the protein cavity have limited numbers of direct hydrogen bonds with the protein molecule. However, the stability of the ZEA2 and ZEA3 ligands seems to be attributed to the dense network of hydrophobic interactions as well as to the water-mediated hydrogen bonds with the protein atoms. In the case of ZEA1, these two types of interactions are less extensive. The weaker docking of the ZEA1 molecule in the binding pocket is related to the fact that the space available for this ligand is much bigger than the molecular volume of a *trans*-zeatin molecule and it is difficult to detect any shape complementarity between the interacting partners. In fact, ZEA1 seems to be accidentally trapped in the cavity and then blocked by the Lys137 gating of the E1 opening. ZEA2 and ZEA3, on the other hand, are found in much better fitting spaces, resulting in stronger binding and improved stabilization.

The superficial trans-zeatin ligands

Apart from the zeatin molecules bound in the protein interior of **Z**, two ligand molecules, ZEA4 and ZEA5, were also identified at the protein surface. ZEA4 is located with 0.5 occupancy at a twofold axis and bridges two symmetry-related protein molecules via van der Waals contacts involving residues Gly140, Lys143, Ala144 in the central part of the solvent-exposed face of helix α 3. ZEA5 (0.7 occupancy), found near loop L5 where its adenine ring

makes stacking interactions with the imidazole ring of His62, is also located near a twofold axis, but at a distance (~ 1.5 Å) that allows it to form adenine-adenine stacking interactions across the dyad, thus extending the aromatic stack to His62-ZEA5-ZEA5'-His62'. From their protein interactions and location at special sites in the crystal lattice, it is evident that ZEA4 and ZEA5 have an important crystal packing role in **Z**. This observation explains why the unsaturated form **U**, in which the protein molecules are evidently charged (at least to some degree) with the internal *trans*-zeatin molecules, but where there are no interstitial ligands, crystallized isomorphously with the free form **F** and not with the saturated complex **Z**.

In the unsaturated crystal structure **U**, there are patches of electron density that coincide with the positions of ZEA1 and ZEA2 (Fig. 3b). At the location corresponding to ZEA3, however, the patch of residual electron density is too small to unambiguously correlate it with *trans*-zeatin (Fig. 3c). There is also residual electron density near loop L5 but at a different position than ZEA5. In general, there is no electron density to indicate any traces of the ZEA4 and ZEA5 binding sites. The absence of external ZEA molecules in the unsaturated complex **U** may be responsible for retaining the *C2* space group symmetry of the free form **F**.

Conformation of the trans-zeatin ligands

All the zeatin molecules have evident *trans* configuration, as defined by the C10-C11-C12-C13 torsion angle (Table 2). The orientation of the terminal OH group, defined by the C11-C12-C13-O13 angle is, however, variable, not only *cis* ($\sim 10^\circ$) or *trans* ($\sim 170^\circ$) type, but also nearly perpendicular to the C13-C12-C14 plane. The disposition of the isoprenoid chain (N1-C6-N6-C10) at the exoamino N6 atom is uniformly distal to the imidazole ring, i.e. the substituent is directed towards the N1 atom of the purine system. The remaining two torsion angles (C6-N6-C10-C11, N6-C10-C11-C12) are quite variable, endowing the isoprenoid chain with a range of conformations, of which only one, in ZEA2 (Fig. 5c), is all-*trans* (extended). On comparison with the conformation of the *trans*-zeatin ligands reported in the VrPhBP (Pasternak *et al.*, 2006) and LIPR-10.2B (Fernandes *et al.*, 2008) complexes, one can conclude that this phytohormone is rather flexible, easily adapting its conformation to the requirements of the binding partner, especially using rotations around the N6-C10, C10-C11 and C12-C13 single bonds. A constant feature is the *cis* conformation of the N1-C6-N6-C10 torsion angle.

Structural adaptation of LIPR-10.1A on binding of *trans*-zeatin

Superposition of the three forms of LIPR-10.1A (Fig. 1, 4, 5b-d) illustrates their differences, which can be viewed as a gradual adaptation of the protein structure to the *trans*-zeatin ligands.

When it comes to the secondary structure elements, the binding events mainly affect the $\alpha 3$ helix, which is unwound in the free form at a segment comprising eight residues (two turns). It is important to note that the R.M.S.D. value of the structural alignment between the C α atoms of the free form **F** and the unsaturated form **U** is 0.86 Å. This value is somewhat higher (1.03 Å) for the alignment of the free and saturated forms (**F-Z**). However, if only the $\alpha 3$ helices are aligned, this value is ~1.3 Å in both cases, indicating that the changes are mostly localized to this structural element. On the other hand, the C α R.M.S.D. value between the two liganded forms (**U-Z**) is only 0.47 Å for all residues, and 0.40 Å for helix $\alpha 3$ alone.

When comparing the protein conformation of the three forms, one can identify several residues that significantly change their position upon interaction with the ligand molecules, leading in general to the increased order of the $\alpha 3$ element. The majority of those residues interact with the ZEA3 molecule but the first stimulus that triggers the helix formation seems to be repulsion between ZEA2 and Leu139, which occupies this binding site in the ligand-free form (Fig. 6). Leu139 caps the E2 entrance in the free form **F** where it forms a van der Waals contact with the C β atom of Asn7, pulling down the $\alpha 3$ helix towards $\beta 1$. In the **Z** complex, the C $\delta 2$ atom of Leu139 in this conformation would be located less than 2 Å from the N9 atom of ZEA2. In the unsaturated form **U** this side chain is slightly pushed away, to be totally moved outside of the E2 entrance in the saturated form **Z**. This change is also mediated by Asn7 and Gln9, both located in strand $\beta 1$, below Leu139. They create two hydrogen bonds in the complex, one between the O $\delta 1$ atom of Asn7 and the N9 atom of ZEA2, and another one between the N $\epsilon 2$ atom of Gln9 and the O $\delta 1$ atom of Asn7, which in effect directs their hydrophilic side chains toward the ZEA2 molecule. At this conformation, these two hydrophilic side chains push the hydrophobic Leu139 outside of the E2 entrance, facilitating the ordering of helix $\alpha 3$. All these conformational changes are very well supported by electron density (Fig. 3a).

The formation (ordering) of helix $\alpha 3$ favors in the end the binding of the ZEA3 molecule as the last internal ligand. In this docking site in the saturated complex **Z**, Arg131 and Lys135 hold the aliphatic tail of ZEA3 in a sort of a “hydrophobic vise”, whereas in the

free form **F** Arg131 is directed away from entrance E2 and Lys135 is in the disordered region (Fig. 6). In the unsaturated complex **U**, Arg131 is close to but not quite yet at the position occupied in the saturated complex, and the side chain of Lys135 is disordered, which can be interpreted as indicating that the unsaturated crystal structure **U** represents an intermediate state leading to the docking of ZEA3 in its binding site. This assumption is supported by the fact that Asp132, which forms a hydrogen bond with the hydroxyl group of ZEA3, is moved farther away from the ZEA3 binding site in the unsaturated complex **U** and is directed totally outside of E2 in the free form **F**. These three side chains (Arg131, Asp132 and Lys135) seem to act like “tentacles” that reach out of the protein surface to draw a ligand molecule to the ZEA3 binding site.

Another residue whose conformational change could stabilize the $\alpha 3$ helix in the complex **Z** is Leu141, which forms van der Waals interactions with ZEA1, while in the free form **F** it penetrates the lumen of the cavity. The interaction with ZEA1 pushes the Leu141 side chain outside of the cavity, straightening at the same time the entire helix $\alpha 3$. Ordering of $\alpha 3$ additionally induces closing up of the E1 entrance by Lys137, sealing off ZEA1 in its binding site.

Another difference at the level of secondary structure is the pronounced change of the conformation (“curvature”) of loops L3, L5 and L7 (the fingertips of the “gripping hand”), which surround the entrance E1, and which upon complex formation change their orientation to tighten the grip of helix $\alpha 3$. Specifically, loop L3 in complex **Z** is moved by ~ 2 Å in the direction of helix $\alpha 3$ relative to the free form **F**. This bending creates van der Waals contacts between Val33-Ile34 and Ala134 from helix $\alpha 3$, clogging the lumen of the E1 entrance.

Although loop L5 in the **F** form is disordered, precluding a detailed conformational comparison, one can observe that the strands $\beta 3$ and $\beta 4$ (connected by L5), which are straightened out in the **F** form, get significantly bent towards the cavity in the **Z** and **U** forms. In this bent conformation, there are van der Waals contacts of residues from the $\beta 3$ and $\beta 4$ strands with ZEA1 in both protein-ligand structures. Moreover, His62 from loop L5 interacts with the external ZEA4 molecule in the saturated structure **Z**, while this interaction is absent in the unsaturated complex **U**, which explains the slightly different conformation of this loop in these two structures.

Loop L7 is also tilted towards helix $\alpha 3$ and the cavity in both complexed forms, but simultaneously it is also stretched towards loop L5 as a consequence of a hydrogen bond

between the N δ 2 atom of His62 from L5 and O ϵ 2 of Glu88 (in L7), which is formed in both protein-ligand complexes (**U** and **Z**), but not in **F**.

The remaining loops seem to be insensitive to the presence of the ligand molecules, as illustrated by their close superposition (not shown). The exception is loop L4, which has a different conformation in each of the LIPR-10.1A forms (Fig. 1b). However, this conformational variability seems to be the effect of crystal packing and is not influenced in a direct way by the *trans*-zeatin cargo molecules.

Finally, it is of note that upon *trans*-zeatin binding there is a pronounced bulging of the strands β 5, β 6 and β 7. These structural elements are cambered most significantly in the saturated complex **Z**, whereas in the free form they are almost collapsed into the interior of the protein (Fig. 1c).

These gradual conformational changes are correlated with the increasing volume of the largest chamber of the internal cavity. Interestingly, the volume changes, which reflect the cargo content, are not related to the cavity surface area, which remains constant at $\sim 5600 \text{ \AA}^2$. Such an independence can be achieved by means of the above-mentioned caving-in of the β 5- β 7 sheet and by the presence of cavity-exposed side chains of the β -sheet that pad the walls of the cavity in the free form **F** but are displaced upon ligand binding. These residues, especially His68, Tyr80 and Ile115, protrude into the interior of the ZEA1 and ZEA2 binding sites in the free form **F** but are tucked away from these sites in the complex (Fig. 5b,c).

Multiple *trans*-zeatin ligands: the sequence and mechanism of binding

Analyzing the orientation of the internal zeatin molecules (Fig. 1), the quality of the $F_c - F_o$ electron density of the ZEA sites in the unsaturated (**U**) and saturated (**Z**) complexes (Fig 3), as well as the shape of the protein cavity (Fig. 4) and the conformation of the protein side chains in all three forms (Fig. 5b-d), one can try to deduce the order and mechanism of zeatin binding.

The ZEA1 and ZEA2 molecules have a similar orientation inside the cavity, suggesting that they could enter the protein cavity one after another through the same E1 entrance. The main protein cavity has a diverticulum in both complex structures (although in the unsaturated complex **U** it is smaller and partly divided by Asn7, Tyr99, Val115 and Ile117), which extends to the E2 entrance and ultimately becomes the binding pocket for ZEA3. Interestingly, in the

free form **F** there is no such a binding pocket at all and in the unsaturated form **U** it is not fully formed, which correlates with the very poor $F_o - F_c$ electron density at the ZEA3 site in the unsaturated complex (Fig. 3c). These observations could suggest that ZEA3 is not present at its binding site before the ZEA1 and ZEA2 molecules are properly docked at their respective binding sites. With this caveat in mind, the ligand binding scenario might be as follows:

- (1) ZEA 2 and ZEA1 enter the cavity through the large entrance E1;
- (2) ZEA2 is docked via hydrogen bonding with Asn7, leaving room for ZEA1;
- (3) The Gln9 side chain rotates towards Asn7 and a hydrogen bond is created between O δ 1 of Asn7 and N ϵ 2 of Gln9, stabilizing Asn7 for firm docking of ZEA2; ZEA2, Asn7 and Gln9 jointly push the side chain of Leu139 outside of the E2 entrance, initiating the ordering of helix α 3 (Fig. 6);
- (4) A turn of the Asn7 side chain from the N-terminal towards the C-terminal part of β 1 after ZEA2 binding opens the E2 entrance enabling docking of the ZEA3 molecule;
- (5) ZEA3 enters the newly open binding site, fitting with its aliphatic tail within the vise of Arg131 and Lys135, and forming a hydrogen bond with Asp132 via the hydroxyl group; this finally stabilizes helix α 3; moreover, N7 of ZEA3 interacts with N δ 2 of Asn7 via a water molecule, additionally stabilizing the network of interactions;
- (6) Binding of ZEA3 and the formation of the ZEA3 binding pocket lead to a widening of the internal cavity and disappearance of the separation from the ZEA1 binding site (Fig. 4); this way the ZEA1 pocket becomes large and this ligand molecule gains additional degrees of freedom; this is why the electron density for ZEA1 is the weakest in the saturated complex **Z** (Fig. 3b).

In summary, ZEA3 binds to the protein only after ZEA2 has been docked first. When in place, ZEA3 blocks the E2 entrance, constricting the space around the ZEA2 molecule. At the same time, ZEA3 binding expands the chamber occupied by ZEA1, endowing this ligand with additional freedom within its otherwise sealed-off cavity.

Crystallographic interpretation of the unsaturated form **U**

While the interpretation of the **F** and **Z** forms is quite clear, because they illustrate, respectively, a ligand-free and 100% *trans*-zeatin-bound protein conformations, the case of **U** is more tricky. We observe here a situation that apparently represents an intermediate protein

conformation, and structurally we claim that this reflects partial saturation with the *trans*-zeatin ligand. However, even if we assume that the ligand binding sites in **U** have fractional occupancy, physically this would mean that some protein molecules in the crystal lattice are empty, while others have the ligands at full occupancy, and that the diffraction experiment gives us a weighted superposition of these two ultimate situations (corresponding to **F** and **Z**). But apparently the protein structure in **U** does not look like a superposition of the states known from **F** and **Z**. If this were the case, we would expect to see a lot more of disorder in the central part of helix $\alpha 3$, and should even be able to model parts of the backbone in double conformation, where the main chain in **U** gets unwound. This is evidently not the case. The most plausible explanation of this logical puzzle is that the structure **U** represents in fact a partial saturation, where the term “partial” refers to occupation of only *some* of the internal binding sites (which in their own right may also have fractional occupancy in the crystal structure). In this interpretation, the initial binding of the first two ligand molecules at sites ZEA1 and ZEA2, would already induce a major switch in the protein conformation, getting it ready for the docking of the final ligand, ZEA3, whose binding is the most tight, but occurs only after this docking site has been preformed by the “avant-garde” molecules ZEA1 (which ultimately is bound very loosely) and ZEA2.

Ordering of helix $\alpha 3$ in the presence of ligands, monitored by CD

The main structural changes of the LIPR-10.1A protein accompanying *trans*-zeatin binding were additionally monitored in solution by circular dichroism measurements in the presence of increasing concentration of the ligand. Circular dichroism is a very sensitive and rapid method for detecting changes in secondary structure content that occur in the presence of denaturing agents or cofactors (Whitmore & Wallace, 2008). Different secondary structures of the protein chain give rise to different CD spectra. The α helix, for example, gives rise to two negative bands at 222 and 208 nm and a positive band at 193 nm, whereas β structures are manifested by one negative band at 218 nm and one positive band at 195 nm (Greenfield, 2006). A change in secondary structure content is typically reflected in a corresponding change of band intensity. Inspection of the 222 nm peak in the CD spectra of LIPR-10.1A recorded in the presence of 0-, 1-, 2-,... and 10-fold molar excess of *trans*-zeatin revealed a concomitant increase of the intensity of this peak, which is associated with the increase of α -helix content in the protein structure. This conclusion was confirmed by a

detailed analysis of the CD spectra with the *DICHROWEB* server (Whitmore & Wallace, 2004) for the detection of the changes in α -helix content (Fig. 7), which additionally showed a saturation (i.e. a plateau of α -helix content) reached at five-fold molar excess of *trans*-zeatin. These experiments in solution confirm the positive effect of *trans*-zeatin on helix $\alpha 3$ ordering, as deduced from the crystallographic data.

Comparison of LIPR-10.1A with the previous model and other yellow lupine isoforms in the free state

The model of the free-form LIPR-10.1A protein determined at the resolution of 1.95 Å available thus far in the PDB under the accession code 1icx, is very similar to the **F** structure presented in this work, with an R.M.S.D. of 0.28 Å for the C α superposition. This value is quite low when compared to the R.M.S. deviation of 0.47 Å calculated for the present LIPR-10.1A complexes at different *trans*-zeatin saturation (**Z** vs **U**). The *a/b/c* cell parameters are also very similar: 35.0/56.6/61.8 Å for the 1icx model and 35.1/56.5/61.9 Å in the present crystal. Although the segments His59-Gly61 and Ala134-Lys135, which are disordered in **F** (and thus omitted from the model), were included in the 1icx model without main-chain interruptions, their ambiguous electron density and high average *B* factors (57 and 76 Å², respectively) indicate poor order of atoms in these regions and make their location doubtful. Apart from this difference, no other significant conformational disparities are observed.

The disorder of helix $\alpha 3$ in the free form of LIPR-10.1A seems to be a unique feature among the yellow lupine PR-10 isoforms for which ligand-free protein structures are available, namely LIPR-10.1B (1ifv, Biesiadka *et al.*, 2002) and LIPR-10.2A (1xdf, Pasternak *et al.*, 2005) (Fig. 8). Comparison of the corresponding sequences in terms of the residues that take part in *trans*-zeatin binding of LIPR-10.1A, explains the absence of extensive helix $\alpha 3$ deformation in LIPR-10.1B and LIPR-10.2A. In neither of those proteins is there a triad of residues, Asn7/Gln9/Leu139, which in LIPR-10.1A, after interacting with the ligand molecules, trigger the series of movements that lead to the ordering of helix $\alpha 3$. Although in LIPR-10.1B there are negatively and positively charged Asp7 and His9 residues at the corresponding positions, Leu139 is replaced by Thr139, whose shorter and polar side chain precludes the structural rearrangements described above.

In LIPR-10.2A (1ivf) the residues at positions 7 and 9 are Asp7 and Ser9. Most importantly, the carbonyl group of Gly139 forms a short hydrogen bond with Tyr82 at the

bottom of the cavity, leading to a sharp kink in the middle of helix $\alpha 3$ and its collapse into the protein interior. Interestingly, Tyr82 is a ligand-interacting residue in most of the PR-10 proteins, while in LIPR-10.2A it is responsible for the helix deformation. Using for comparison the ligand-free model of LIPR-10.2A and a *trans*-zeatin complex of LIPR-10.2B, Fernandes *et al.* (2009) suggested that this collapse of helix $\alpha 3$ into the protein core could be reverted upon ligand binding. However, in the **Z** and **F** forms of LIPR-10.1A (Fig. 1b) the degree of this collapse is almost the same. The differences between the free and liganded forms of LIPR-10.1A are focused mainly on the degree of order of the $\alpha 3$ helix and on its distance from the $\beta 1$ strand (Fig. 1a). The mode of adaptation of the C-terminal helix, therefore, seems to be specific for each PR-10 case and depends on the particular (generally highly variable) helix $\alpha 3$ sequence.

Comparison of the saturated complex with other PR-10 complexes of *trans*-zeatin

C α superposition of the available models of PR-10 proteins (VrPhBP, MtN13 and LIPR-10.2B) in complex with *trans*-zeatin (2flh, 4jhg and 2qim, respectively) with the present complex **Z** resulted in R.M.S.D. values of 1.41, 1.63 and 1.72 Å, respectively. Interestingly, the corresponding levels of sequence identity/similarity in pairwise comparisons with LIPR-10.1A (Rice *et al.*, 2000) vary significantly and are equal to 20/43, 36/55 and 60/77%, respectively, indicating that the primary structure does affect, even if only slightly, the conformational (dis)similarity of the complexes.

In terms of conformation, the most conserved structural elements among the available structures of PR-10/*trans*-zeatin complexes are the β -sheet and loops L1 and L6 (Fernandes *et al.*, 2013). On the other hand, the secondary structure elements with highest conformational variability are helix $\alpha 3$ and loop L8, as well as the loops surrounding the E1 entrance, i.e. L3, L5, L7 and L9 (Fig. 9).

In particular, the C-terminal helix $\alpha 3$ has a variable distance to the $\beta 1$ strand (Fig. 9a) as well as different degree of collapse (inward kinking) (Fig. 9b) into the protein hydrophobic core, in different PR-10/*trans*-zeatin complexes. These structural properties of $\alpha 3$ seem to depend on the ligand distribution within the protein cavity. The $\alpha 3$ helix of form **Z** of LIPR-10.1A has the longest distance to the $\beta 1$ strand because two of the ZEA ligands (ZEA2 and ZEA3) are accommodated between these structural elements. In the LIPR-10.2B/ZEA

complex (2qim, Fernandes *et al.*, 2009), this distance is somewhat shorter since the (three) *trans*-zeatin molecules are located more centrally in a huge protein cavity and (Fig. 9c), whereas in the cases of the VrPhBP and MtN13 proteins, which have a rather small cavity and where there is no functional E2 entrance (Sliwiak *et al.*, 2013), helix $\alpha 3$ is even closer to the $\beta 1$ strand. The inward kinking of helix $\alpha 3$ into the hydrophobic core is the most prominent in the case of the LIPR-10.1A/ZEA complex, and the smallest in the LIPR-10.2B/ZEA complex where the big cargo has a central location in the cavity (Fig. 9c). These observations confirm the dominating effect of the $\alpha 3$ helix on the shape of the internal cavity of the PR-10 proteins.

The cavity volumes for the above complexes are 1979 (LIPR-10.1A/ZEA), 2434 (LIPR-10.2B/ZEA), 1080 (VrPhBP/ZEA) and 726 Å³ (MtN13/ZEA), with shapes as presented in Fig. 9c. We note that the two yellow lupine isoforms are examples of PR-10 proteins with type 2 cavity, whereas MtN13 and VrPhBP have type 1 cavity (Sliwiak *et al.*, 2013).

With regard to *trans*-zeatin binding, neither the ligand position and orientation, nor even the stoichiometry are conserved (Fig. 9c). A multiple sequence alignment (Goujon *et al.*, 2010) reveals that there are no conserved residues in the LIPR-10.1A sequence that would be constantly involved in ligand interactions, except Tyr82, which interacts, however, with different *trans*-zeatin atoms and via different interactions. Nevertheless, if only the LIPR-10.1A and 2B lupine isoforms are compared, it is apparent that they share a set of common residues (Phe5, Leu22, Ile115, Val117, Phe142) that form van der Waals interactions with the ligands (even if with their different fragments).

In each of the PR-10/ZEA complex structures available to-date, at least one *trans*-zeatin molecule was bound in a “fork-like” manner at the N6/N7 atoms by a Glu, Gln or Asp residue. In the present complex of LIPR-10.1A (Z), none of the five ligand molecules is bound in this way. Moreover, the N6 and N7 atoms of ZEA2 and ZEA3 are hydrogen bonded with one water molecule in each case, which mediates their hydrogen-bond interactions with the protein atoms.

Conclusions

Uncharacteristically, the LIPR-10.1A protein from yellow lupine exhibits no promiscuity in ligand binding, in contrast to other PR-10 proteins with large type 2 cavity, such as the yellow lupine LIPR-10.2B isoform or Bet v 1. The fact that LIPR-10.1A could be only complexed with *trans*-zeatin, suggests that it might be a specific binder of this

phytohormone. Based on the presented models of LIPR-10.1A in the free (**F**), unsaturated (**U**) and saturated (**Z**) forms, one can detect several structural changes that accompany *trans*-zeatin binding. They are visible mainly in (i) the conformation of helix $\alpha 3$ (unwinding-winding of the central two coils), (ii) the curvature of the loops (“fingertips”) surrounding the E1 cavity entrance (gripping of helix $\alpha 3$), (iii) the bulging out of the $\beta 5$ - $\beta 7$ sheet, and (iv) the orientation of the side chains lining the walls of the cavity. The helix formation effect (i) was additionally observed in solution by monitoring the CD spectra at increasing *trans*-zeatin concentration. It was found that at five-fold molar excess the ligand effect reaches a plateau and does not increase the helical content anymore. LIPR-10.1A is capable of accommodating three *trans*-zeatin molecules in its internal cavity. ZEA1 is held rather loosely in the largest hilum by sporadic contacts, whereas ZEA2 and ZEA3 have numerous hydrophobic contacts with the protein as well as hydrogen bonds with residues from the $\alpha 3$ and $\beta 1$ structural elements, which trigger the ordering of the $\alpha 3$ helix. The sequence of the binding events seems to begin with the docking of ZEA1 and ZEA2, which drives the main structural changes and prepare the ultimate binding site of ZEA3, which is bound most tightly, but only in the presence of ZEA1 and ZEA2. The structural adaptation of helix $\alpha 3$ seems to control and orchestrate the whole binding process: the unwound helix facilitates ligand access to the cavity, whereas the fully formed helix blocks the ligand escape routes. The three crystal structures of LIPR-10.1A presented in this work are the first complete set of the same PR-10 protein in different ligand-binding states to enable an in-depth discussion of the structural transformations of plant proteins with PR-10 fold upon binding of physiologically-relevant natural ligands.

Acknowledgments

This publication was supported by the Polish Ministry of Science and Higher Education, under the KNOW program.

Table 1. Data collection and refinement statistics.

Crystal form	F	U	Z
Data collection			
Temperature (K)	100	100	100
Space group	$P2_12_12_1$	$P2_12_12_1$	$C2$
Cell dimensions			
a, b, c (Å)	35.1, 56.5, 61.9	35.5, 57.9, 62.5	70.1, 63.9, 47.6
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 126.7, 90
Resolution (Å)	30.0-1.32 (1.37-1.32) ¹	50.0-1.50 (1.59-1.50)	99.0-1.38 (1.43-1.38)
R_{merge}	0.059 (0.401)	0.090 (0.829)	0.054 (0.209)
$\langle I/\sigma I \rangle$	19.70 (2.00)	21.96 (2.45)	24.74 (6.48)
Completeness (%)	99.9 (99.7)	99.5 (98.2)	99.2 (99.6)
Redundancy	3.7 (3.3)	7.2 (2.4)	4.1 (3.5)
Refinement			
Resolution (Å)	28.27-1.32	42.48-1.50	42.49-1.38
Reflections work/test	28503 / 1059	21204 / 1001	33856 / 896
R_{work}/R_{free} (%)	15.3 / 21.2	18.1 / 24.3	14.3 / 17.0
No. atoms			
Protein	1248	1243	1239
Ligand / ion	0 / 4	0 / 0	80 (ZEA) / 5 (SO ₄ ²⁻)
Water	178	148	222
$\langle B \rangle$ factors (Å ²)			
Protein	19.98	27.10	14.55
Ligand / ion	0 / 33.99	-	23.83 / 36.10
Water	31.03	37.92	30.72
R.M.S. deviations from ideal			
Bond lengths (Å)	0.019	0.018	0.015
Bond angles (°)	1.8	1.8	2.0
Ramachandran statistics (%)			
Favored	99.4	99.4	99.4
Outliers	0	0	0
PDB code	4Y31	5C9Y	4RYV

¹Values in parentheses correspond to the last resolution shell.

Table 2. Conformation of the *trans*-zeatin molecules in structure **Z**. The internal ligands (ZEA1, ZEA2, ZEA3) are separated from the superficial ones (ZEA4, ZEA5) with a heavy vertical line. The atom numbering scheme is given in Fig. 2.

Torsion angle (°)	ZEA1	ZEA2	ZEA3	ZEA4	ZEA5
N1-C6-N6-C10	-13.1	-0.6	5.9	1.5	-11.4
C6-N6-C10-C11	104.4	-175.8	112.6	-179.8	65.2
N6-C10-C11-C12	167.2	-169.7	101.2	9.0	90.0
C10-C11-C12-C13	178.1	179.2	179.5	-177.7	179.0
C11-C12-C13-O13	93.6	13.8	9.3	-62.8	-167.6

Figure captions

Fig. 1. (a) Superposition of the three LIPR-10.1A structures determined in this work, with annotation of the canonical structural elements of the PR-10 fold. The free form **F** is shown in green, the unsaturated form **U** in yellow, and the saturated form **Z** in red (with the external *trans*-zeatin molecules in gray). (b) A different view of the same superposition. (c) A zoom-in view illustrating the variable degree of caving of the $\beta 5$ - $\beta 7$ fragment of the β -sheet.

Fig. 2. The chemical structure of *trans*-zeatin, with the correct (Jaskolski, 2013) atom-numbering scheme.

Fig. 3. (a) $2F_o - F_c$ electron density contoured at 1.0σ , illustrating the modeling of Asn7, Gln9 and Leu139 in the three crystal forms, color-coded as in Fig. 1a. (b) Omit $F_o - F_c$ maps contoured at 3.0σ for the internal ZEA molecules in the saturated complex **Z**. (c) $F_o - F_c$ maps contoured at 3.0σ , generated for the unsaturated structure **U** at the locations of the internal ZEA molecules, presented as ball-and-stick models taken from the saturated complex **Z**. *trans*-Zeatin molecules were not modeled in **U** and are shown in (c) only for orientation.

Fig. 4. The shape of the internal cavity in the three forms of LIPR-10.1A, color-coded as in Fig. 1a, with additional coloring of the surfaces at nitrogen (blue) and oxygen (red) atoms. The *trans*-zeatin molecules (ZEA), taken from the saturated structure **Z**, were superposed to illustrate the formation of the binding sites.

Fig. 5. (a) ZEA molecules in space-filling representation within the surface of the LIPR-10.1A protein in the saturated complex **Z**. The binding sites of ZEA1 (b), ZEA2 (c) and ZEA3 (d) with residues forming direct hydrogen bonds with the ligand molecules shown in red and those forming van der Waals contacts or water-mediated hydrogen bonds shown in salmon. The corresponding residues from the ligand-free (**F**, green) and unsaturated (**U**, yellow) forms were superposed to illustrate their gradual adaptation.

Fig. 6. Cooperativity between the ZEA2 and ZEA3 binding sites, illustrated as a sequence of residue movements that create the ultimate ZEA3 binding site and lead to ordering of helix $\alpha 3$ (light green, free form **F**; yellow, unsaturated form **U**; salmon, saturated form **Z**). The gray arrows illustrate the direction of ligand-induced conformational adaptations.

Fig. 7. (a) CD spectrum of the LIPR-10.1A protein without a ligand. (b) Plot of the intensity of the peak at 222 nm (●), or change (%) in α helix content (□) generated in *DICHROWEB* (Withmore & Wallace, 2004) from the CD spectra collected for protein samples at different ZEA:protein ratio.

Fig. 8. Superposition of three yellow lupine PR-10 isoforms in their ligand-free state: LIPR-10.1A (F form, green) LIPR-10.1B (2qim, orange) and LIPR-10.2A (1ifv, blue).

Fig. 9. (a) Superposition of the available PR-10/*trans*-zeatin complexes of MtN13 (4jhg; green), VrPhBP (2flh, chain A; blue), LIPR-10.2B (2qim; orange) and of the present model **Z** (red), emphasizing the different distances (gray arrow) of helix $\alpha 3$ to strand $\beta 1$. (b) The same superposition shown in a view illustrating the degree of collapse of helix $\alpha 3$ into the hydrophobic core of the protein (white arrow). (c) Ligand positions in the internal cavities of the PR-10/*trans*-zeatin complexes, colored as in (a) and (b).

References

- Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* 68, 352-367.
- Ames, B.D., Korman, T.P., Zhang, W., Smith, P., Vu, T., Tang, Y., Tsai, S.C., 2008. Crystal structure and functional analysis of tetracenomycin ARO/CYC: Implications for cyclization specificity of aromatic polyketides. *PNAS* 105, 5349-5354.
- Berkner, H., Schweimer, K., Matecko, I., Rösch, P., 2008. Conformation, catalytic site, and enzymatic mechanism of the PR10 allergen-related enzyme norcoclaurine synthase. *Biochem. J.* 413, 281-290.
- Biesiadka, J., Bujacz, G., Sikorski, M.M., Jaskolski, M., 2002. Crystal structures of two homologous pathogenesis-related proteins from yellow lupine. *J. Mol. Biol.* 319, 1223-1234.
- 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. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* 66, 12-21.
- Cohen, G.H., 1997. *ALIGN*: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Crystallogr.* 30, 1160-1161.
- Chwastyk, M., Jaskolski, M., Cieplak, M., 2014. Structure-based thermodynamic and mechanical stability of plant PR-10 proteins with cavities. *FEBS J.* 281, 416-429.
- Emsley, P., Lohkamp, B., Scott, W., Cowtan, K., 2010. Features and Development of Coot. *Acta Crystallogr. D* 66, 486-501.
- Fernandes, H., Pasternak, O., Bujacz, G., Bujacz, A., Sikorski, M.M., Jaskolski, M., 2009. *Lupinus luteus* Pathogenesis-Related protein as a reservoir for cytokinin. *J. Mol. Biol.* 378, 1040-1051.
- Fernandes, H., Bujacz, A., Bujacz, G., Jelen, F., Jasinski, M., Kachlicki, P., Otlewski, J., Sikorski, M.M., Jaskolski, M., 2009. Cytokinin-induced structural adaptability of a *Lupinus luteus* PR-10 protein. *FEBS J.* 276, 1596-1609.

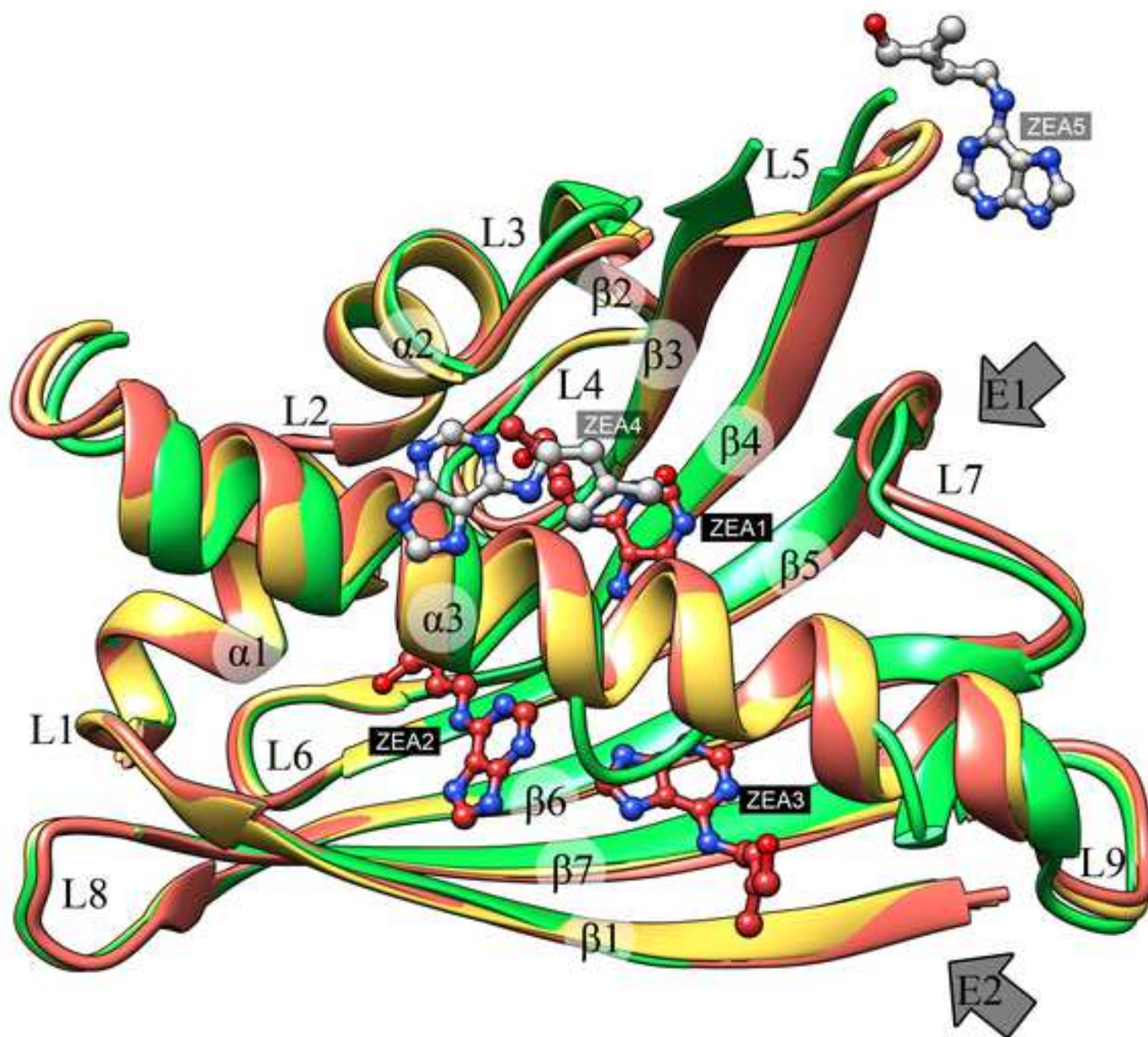
- Fernandes, H., Michalska, K., Sikorski, M.M., Jaskolski, M., 2013. Structural and functional aspects of PR-10 proteins. *FEBS J.* 280, 1169-1199.
- Fujimoto, Y., Nagata, R., Fukasawa, H., Yano, K., Azuma, M., Iida, A., Sugimoto, S., Shudo, K., Hashimoto, Y., 1998. Purification and cDNA cloning of cytokinin-specific binding protein from mung bean (*Vigna radiata*). *Eur. J. Biochem.* 258, 794-802.
- Greenfield, N.J., 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* 1, 2876-2890.
- Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., Lopez, R., 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* 38, W695-W699.
- Hurlburt, B.K., Offermann, L.R., McBride, J.K., Majorek, K.A., Maleki, S.J., Chruszcz, M., 2013. Structure and function of the peanut panallergen Ara h 8. *J. Biol. Chem.* 288, 36890-36901.
- Hwang, I., Sheen, J., Müller, B., 2012. Cytokinin signaling networks. *Ann. Rev. Plant Biol.* 63, 353-380.
- Jaskolski, M., 2013. On propagation of errors. *Acta Crystallogr.* D69, 1865-1866.
- Kabsch, W., 2010. XDS. *Acta Crystallogr.* D66, 125-132.
- Kofler, S., Asam, C., Eckhard, U., Wallner, M., Ferreira, F., Brandstetter, H., 2012. Crystallographically mapped ligand binding differs in high and low IgE binding isoforms of birch pollen allergen bet v 1. *J. Mol. Biol.* 422, 109-123.
- Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774-797.
- Loon, L.C.van, Rep, M., Pieterse, C.M., 2006. Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol.* 44, 135-162.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658-674.
- Merritt, E.A., 2012. To *B* or not to *B*: a question of resolution? *Acta Crystallogr.* D68, 468-477.

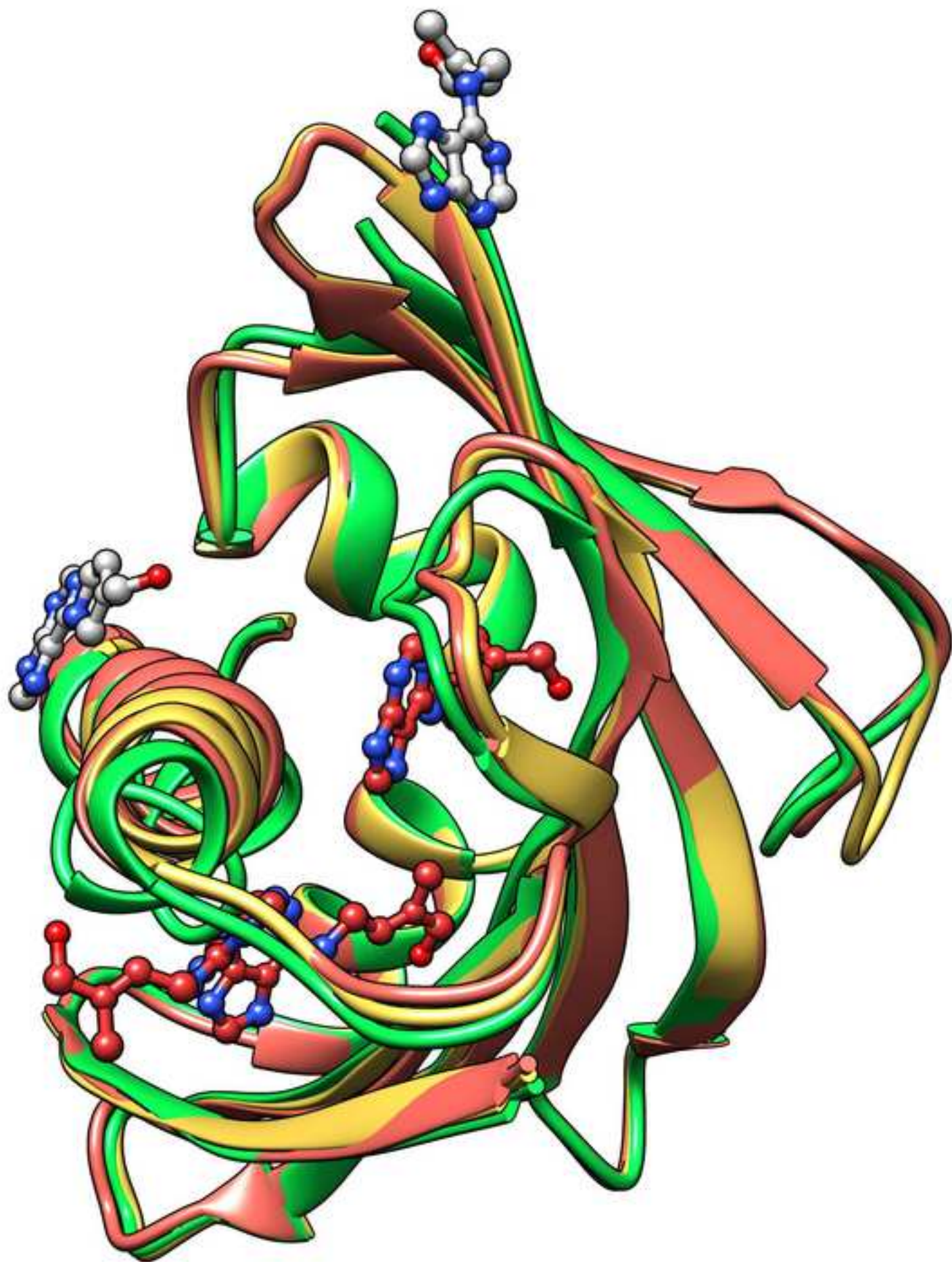
- Miyazono, K., Miyakawa, T., Sawano, Y., Kubota, K., Kang, H., Asano, A., Miyauchi, Y., Takahashi, M., Zhi, Y., Fujita, Y., Yoshida, T., Kodaira, K., Yamaguchi-Shinozaki, K., Tanokura, M., 2009. Structural basis of abscisic acid signalling. *Nature* 462, 609-615.
- Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the Maximum-Likelihood method. *Acta Crystallogr.* D53, 240-255.
- Otwinowski, Z., Minor, W., 1997. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* 276, 307-326.
- Pasternak, O., Biesiadka, J., Dolot, R., Handschuh, L., Bujacz, G., Sikorski, M.M., Jaskolski, M., 2005. Structure of a yellow lupin pathogenesis-related PR-10 protein belonging to a novel subclass. *Acta Crystallogr.* D61, 99-107.
- Pasternak, O., Bujacz, G.D., Fujimoto, Y., Hashimoto, Y., Jelen, F., Otlewski, J., Sikorski, M.M., Jaskolski, M., 2006. Crystal structure of *Vigna radiata* cytokinin-specific binding protein in complex with zeatin. *Plant Cell* 18, 2622-2634.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605-1612.
- Ricci, A., Bertolotti, C., 2009. Urea derivatives on the move: cytokinin-like activity and adventitious rooting enhancement depend on chemical structure. *Plant Biol.* 11, 262-272.
- Rice, P., Longden, I., Bleasby, A., 2000. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* 16, 276-277.
- Ruszkowski, M., Szpotkowski, K., Sikorski, M.M., Jaskolski, M., 2013. The landscape of cytokinin binding by a plant nodulin. *Acta Crystallogr.* D69, 2365-2380.
- Ruszkowski, M., Sliwiak, J., Ciesielska, A., Barciszewski, J., Sikorski, M.M., Jaskolski, M., 2014. Specific binding of gibberellic acid by Cytokinin-Specific Binding Proteins: a new aspect of plant hormone-binding proteins with the PR-10 fold. *Acta Crystallogr.* D70, 2032-2041.
- Sliwiak, J., Dauter, Z., Jaskolski, M., 2013. Hyp-1 protein from St John's wort as a PR-10 protein. *BioTechnologia* 94, 47-50.

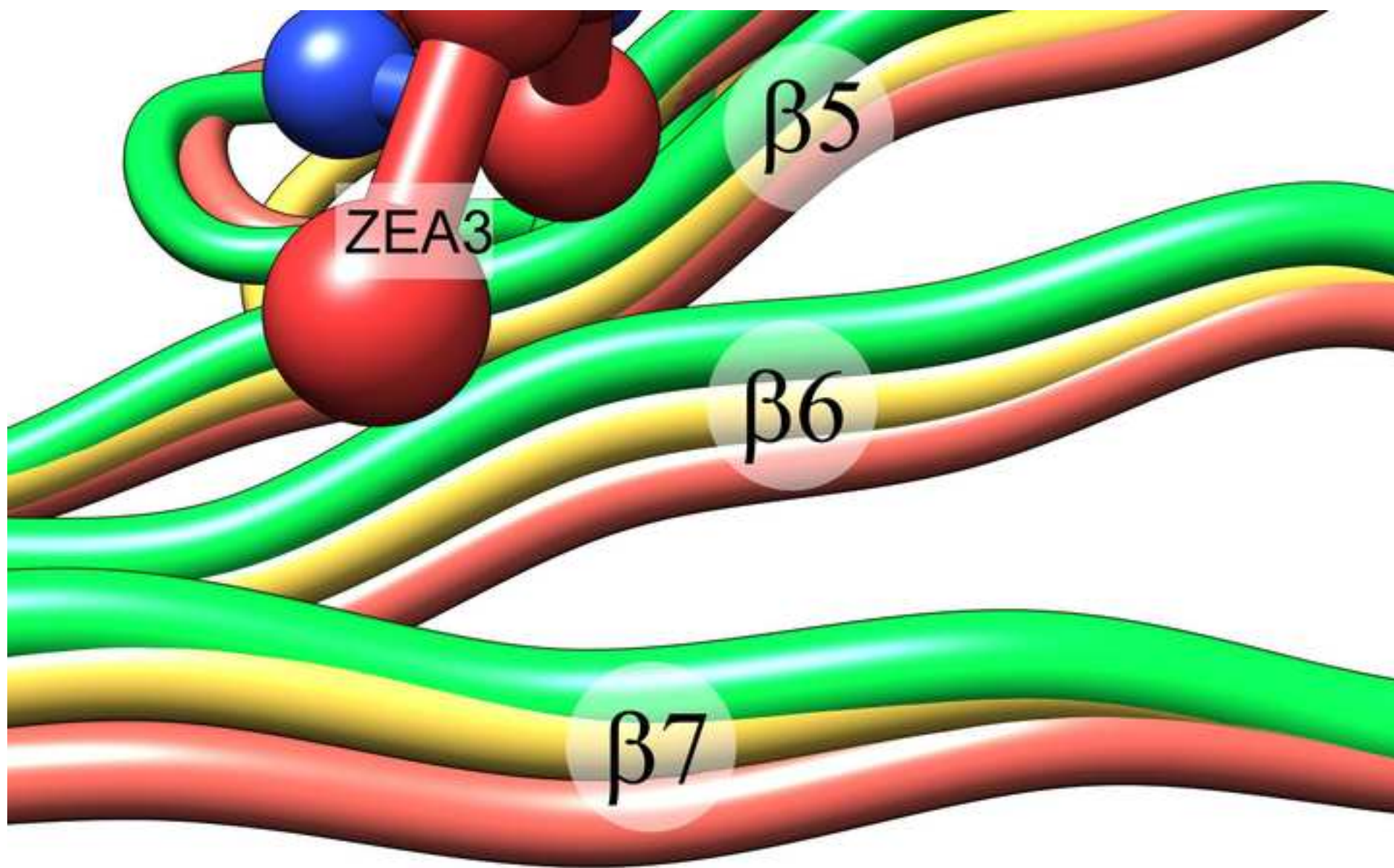
Whitmore, L., Wallace, B.A., 2004. DICHROWEB: an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* 32, 668-673.

Whitmore, L., Wallace, B.A., 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 89, 392-400.

ACCEPTED MANUSCRIPT







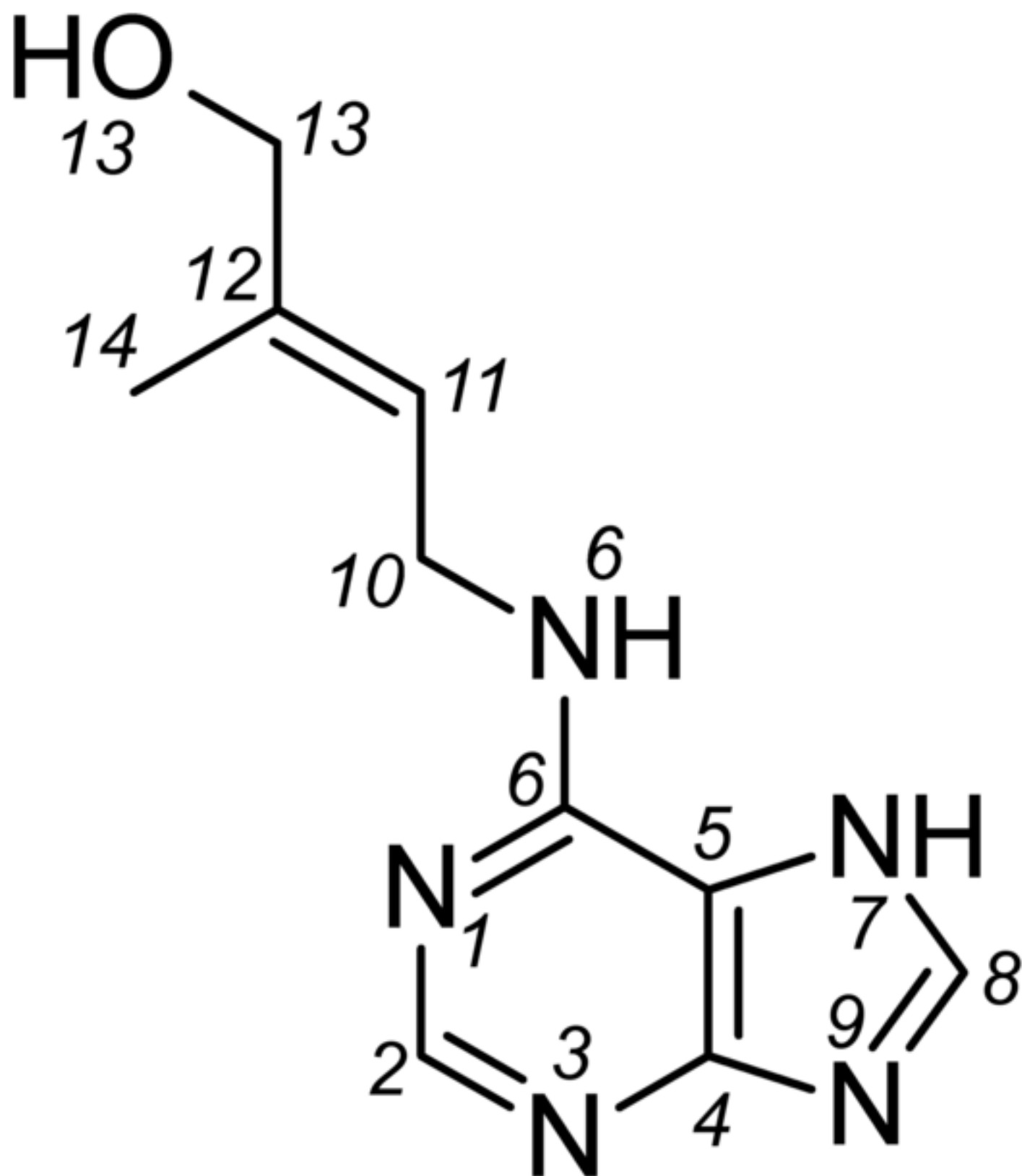
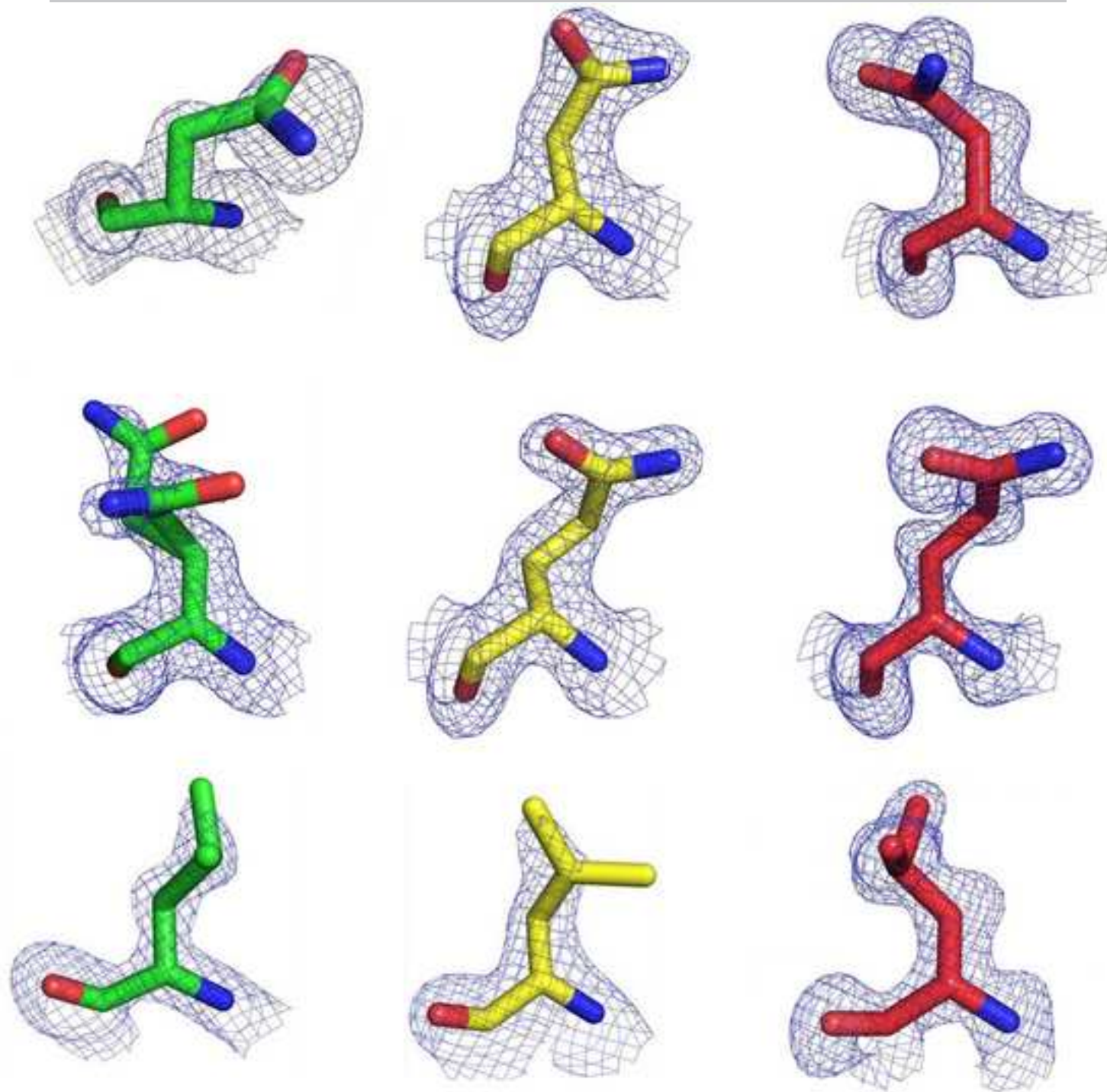
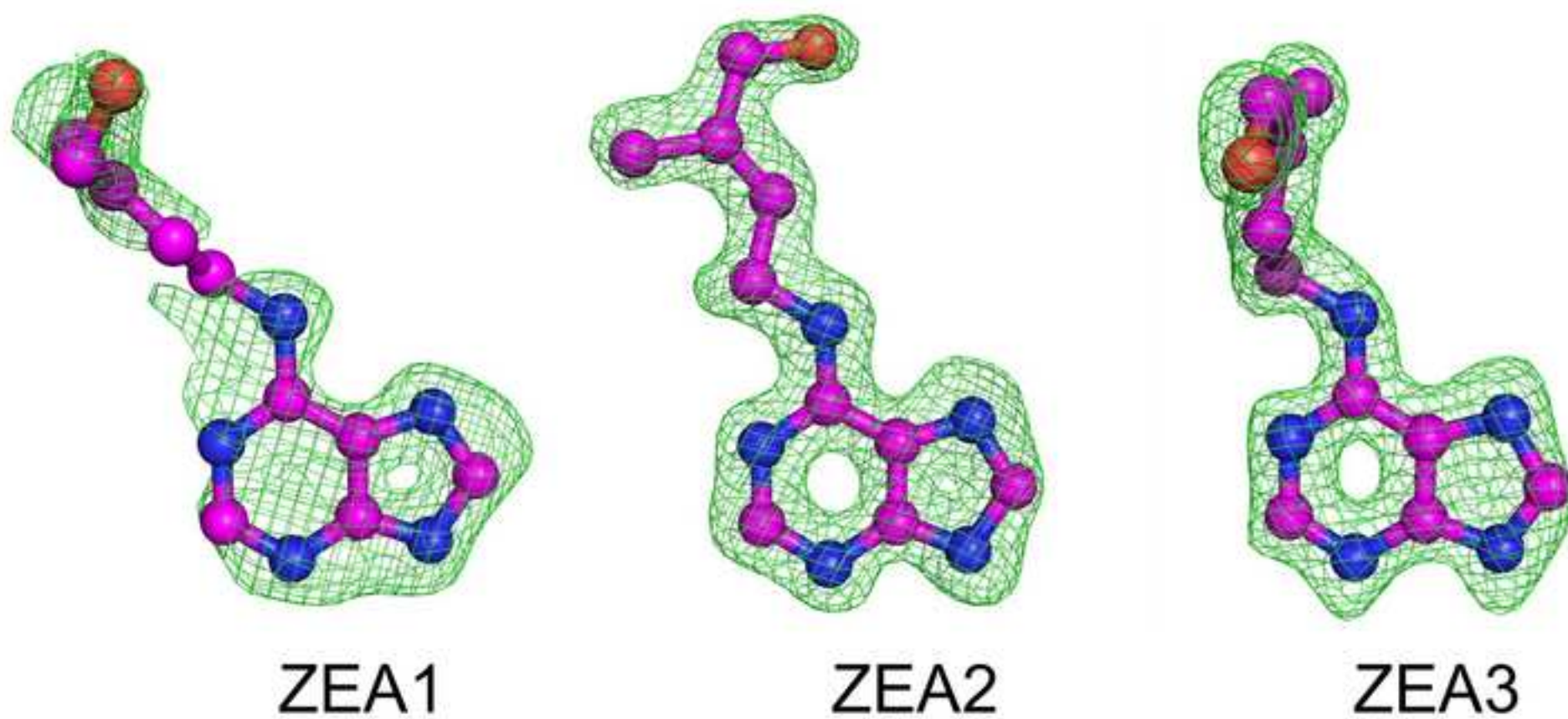


Figure 3a





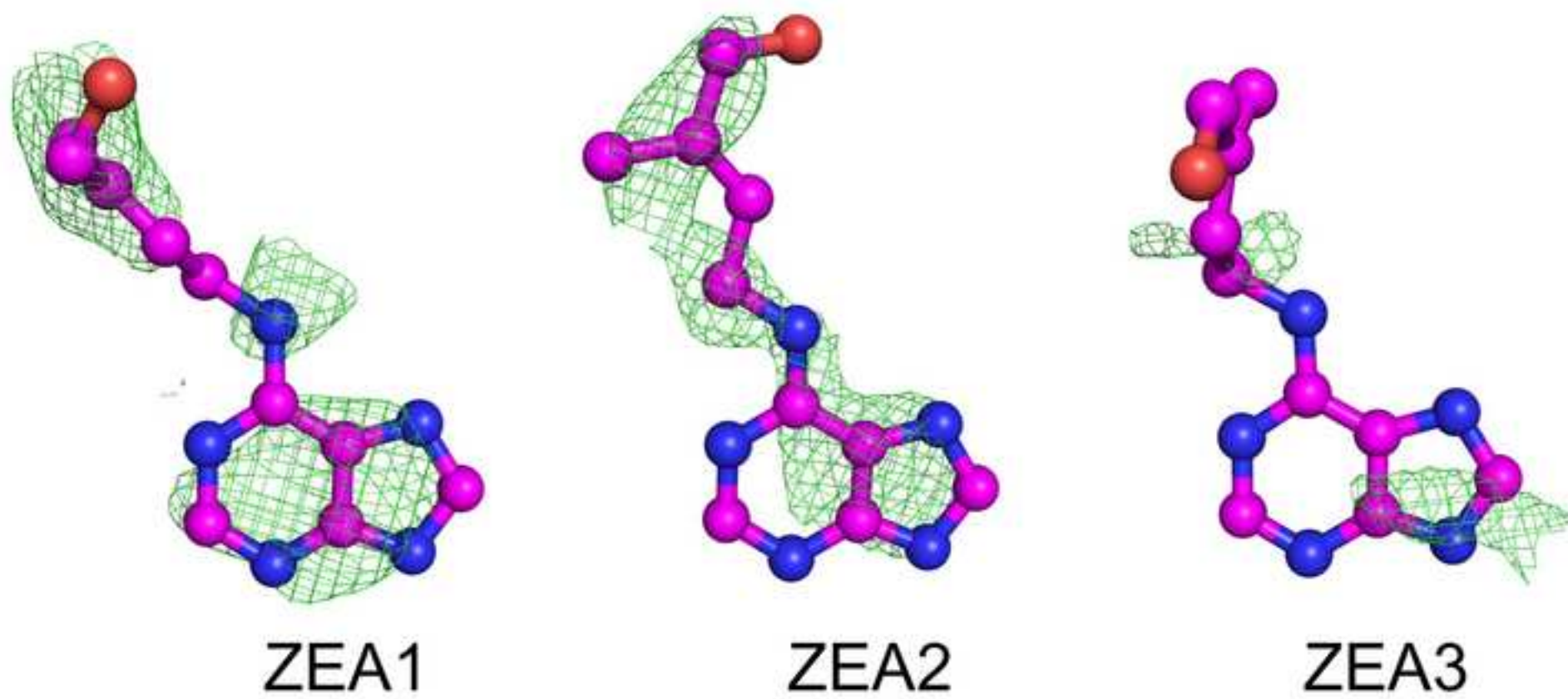


Figure 4

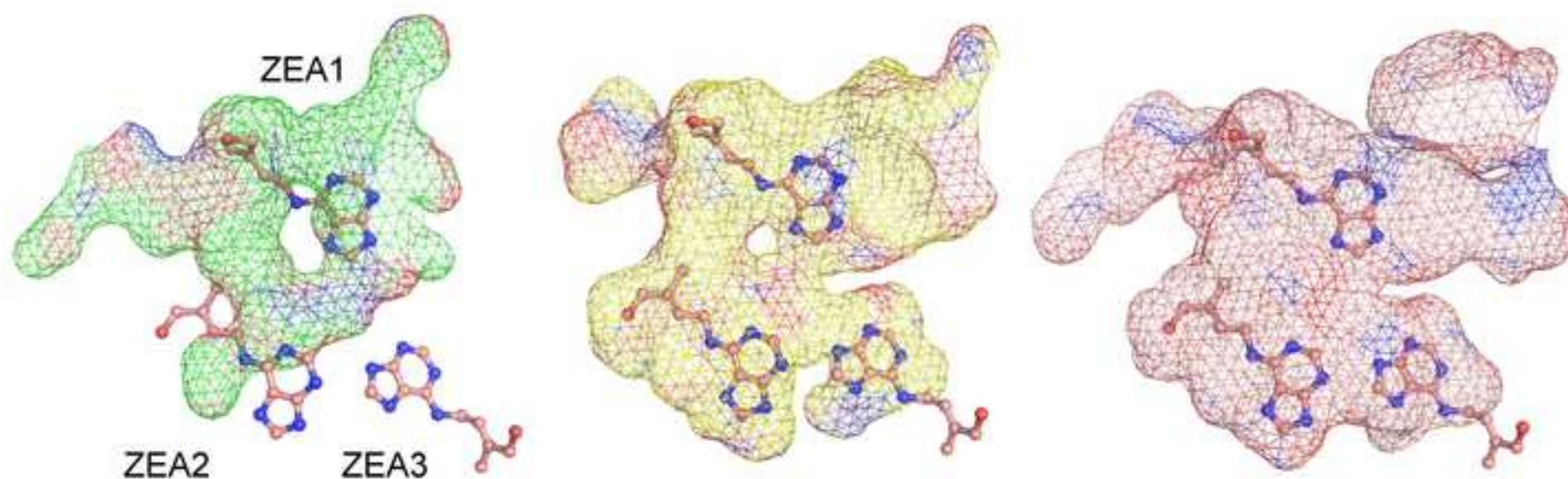


Figure 5a

ACCEPTED MANUSCRIPT

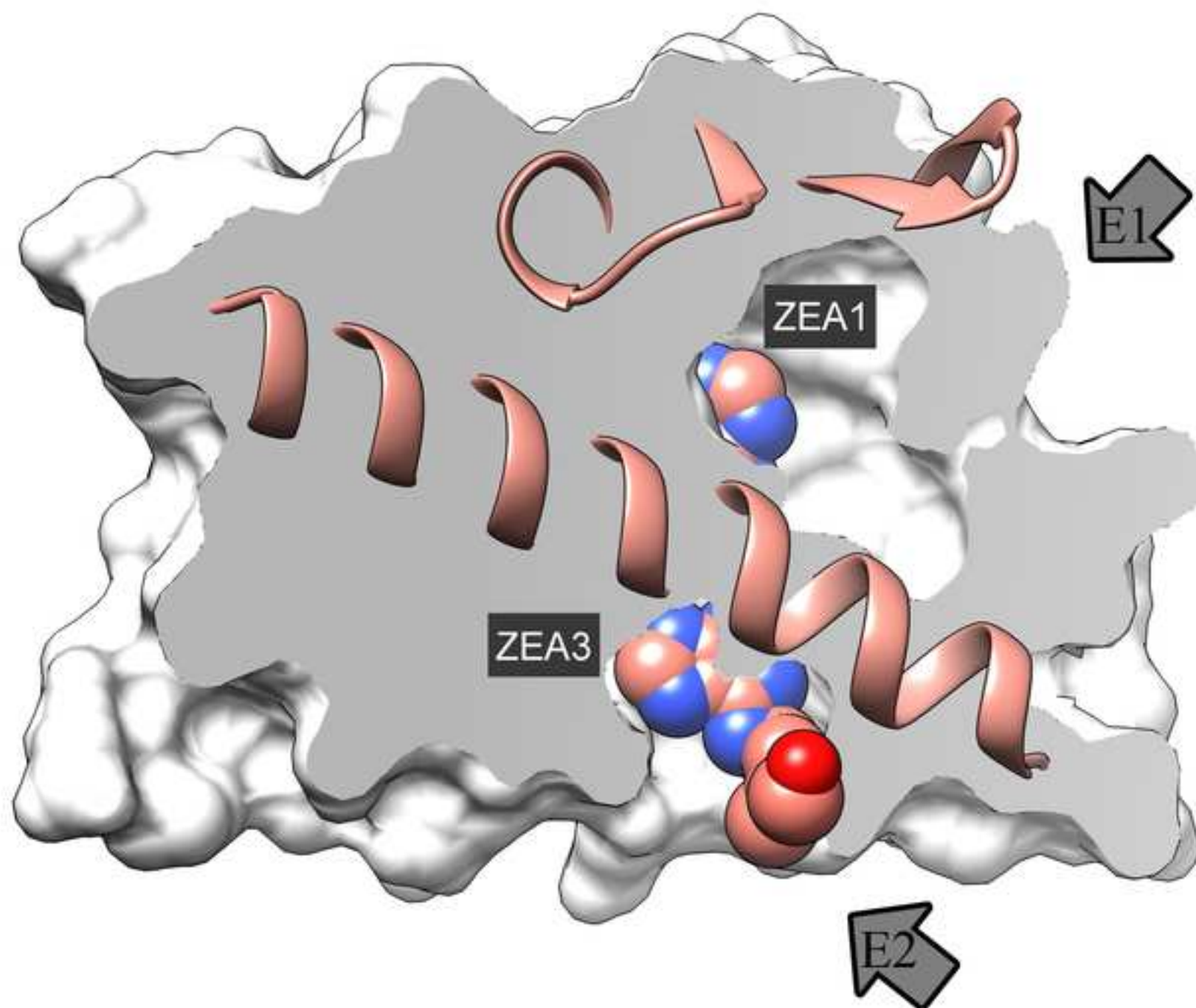


Figure 5b

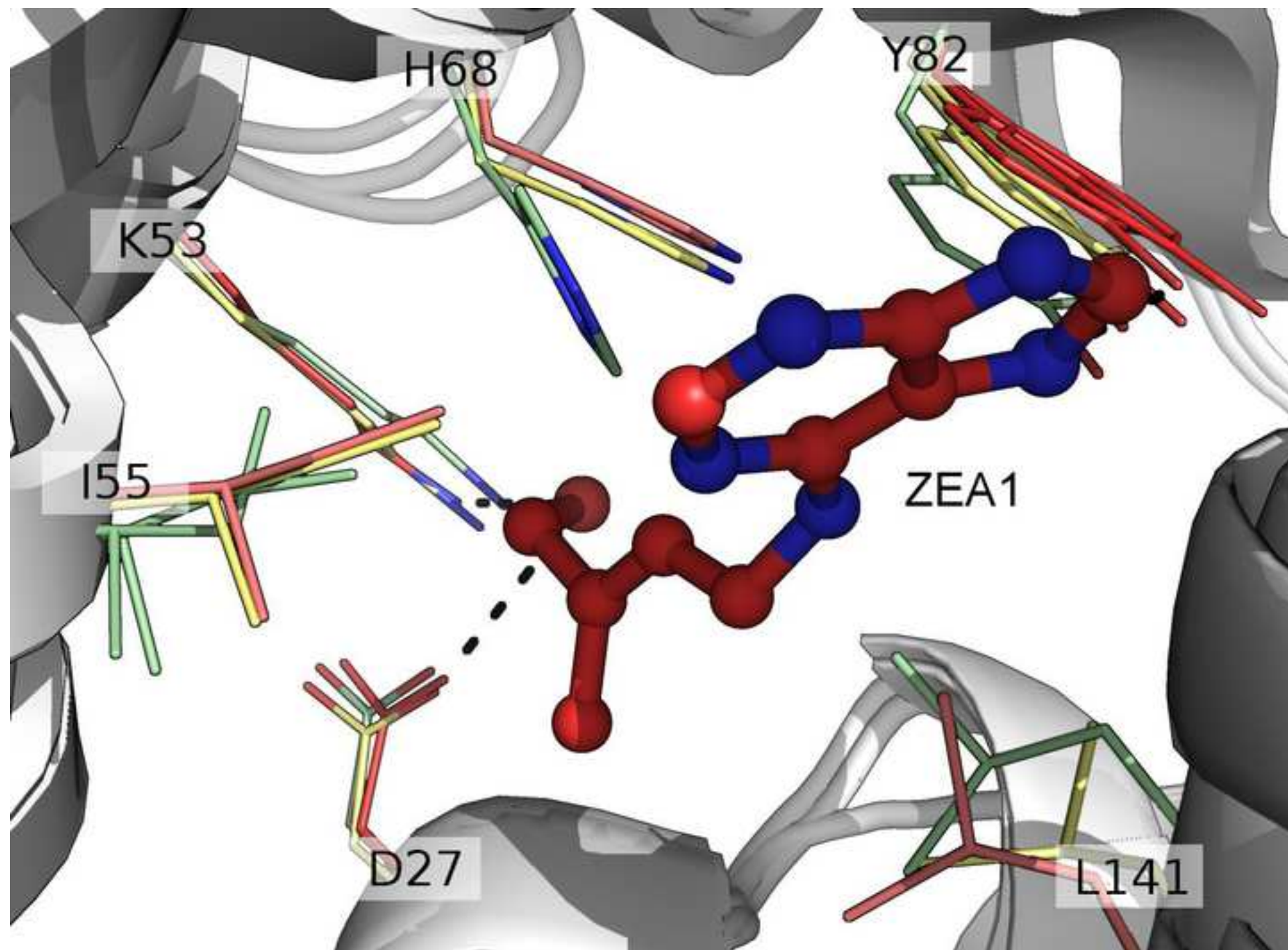


Figure 5c

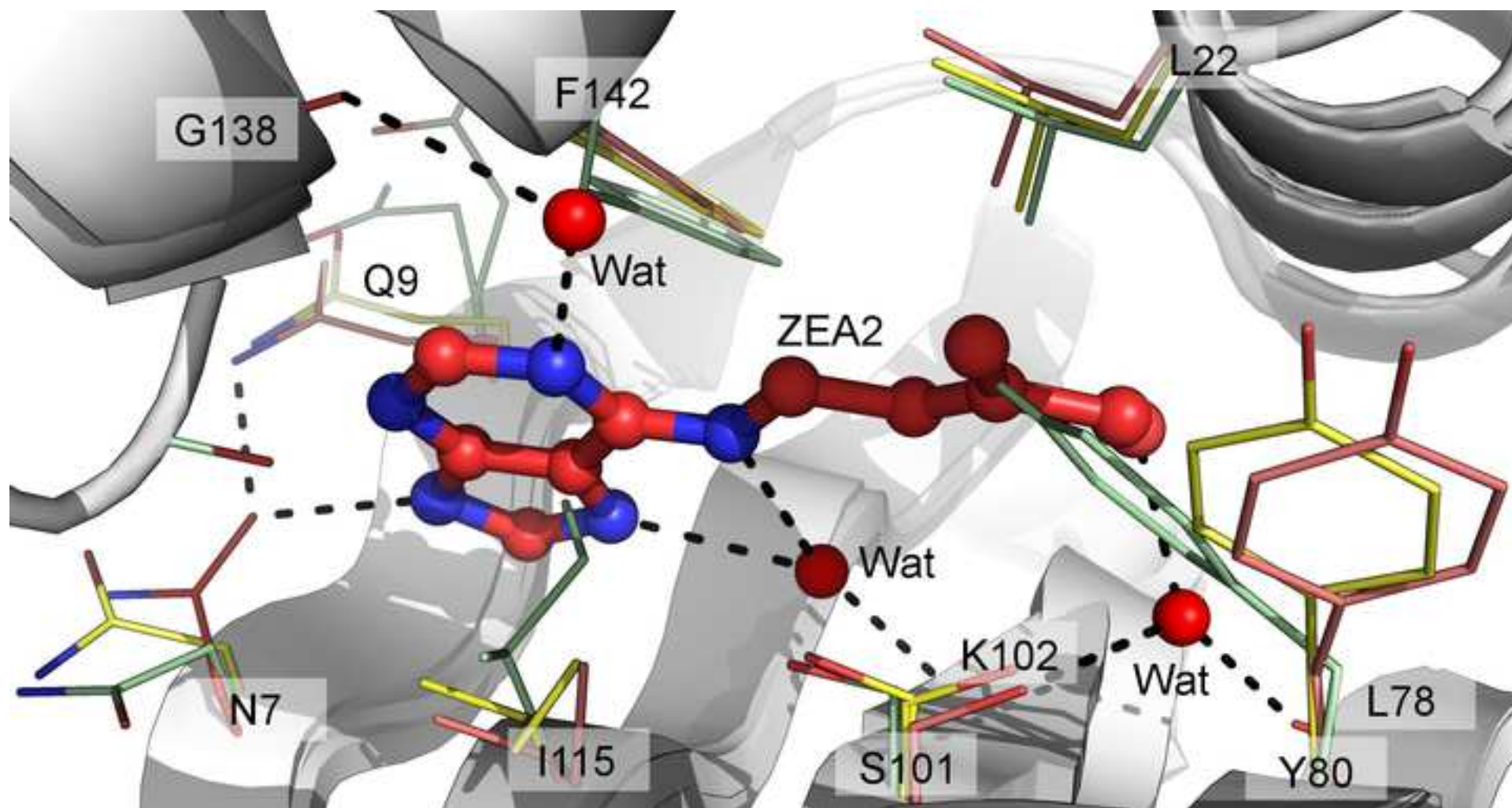


Figure 5d

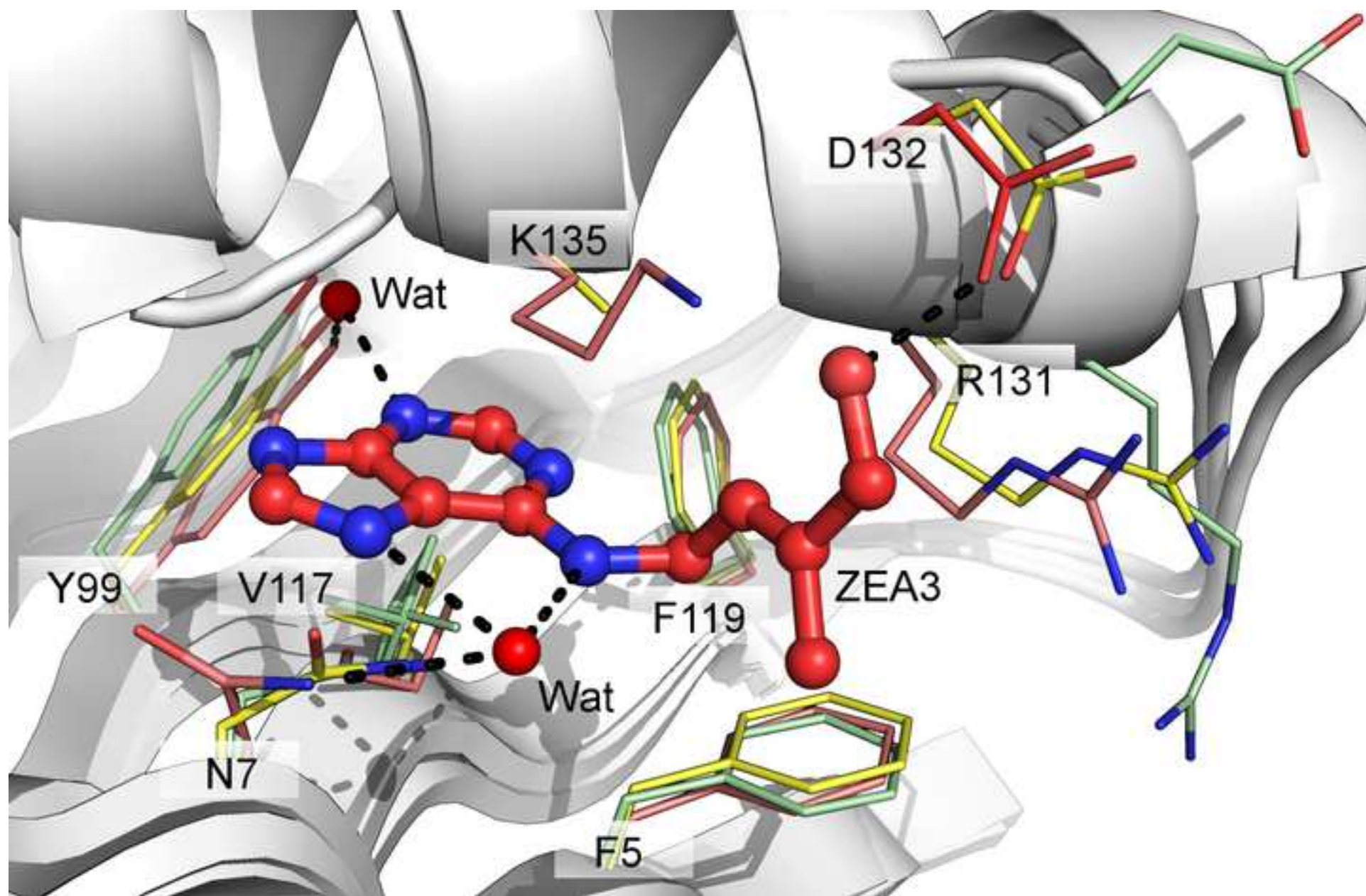
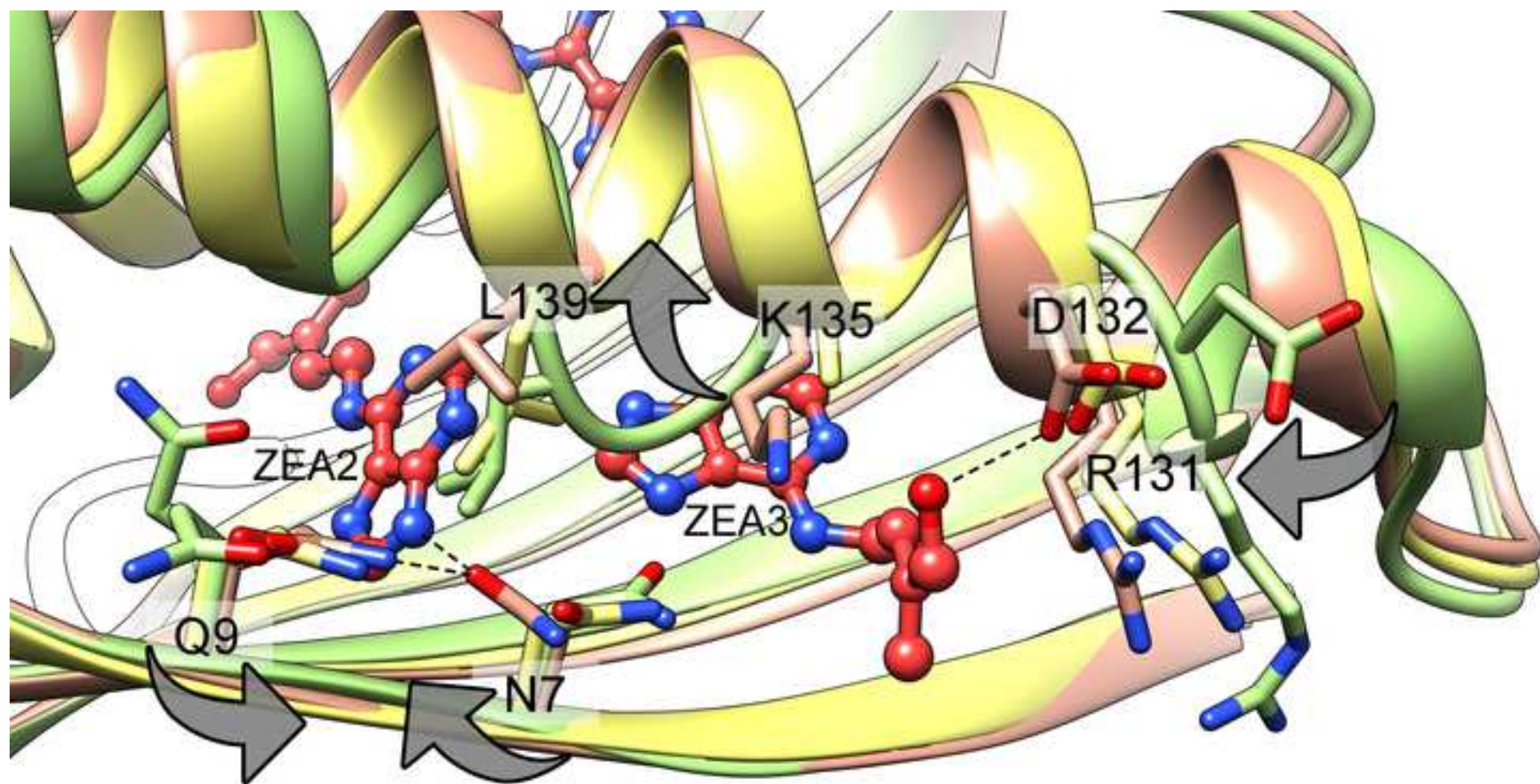
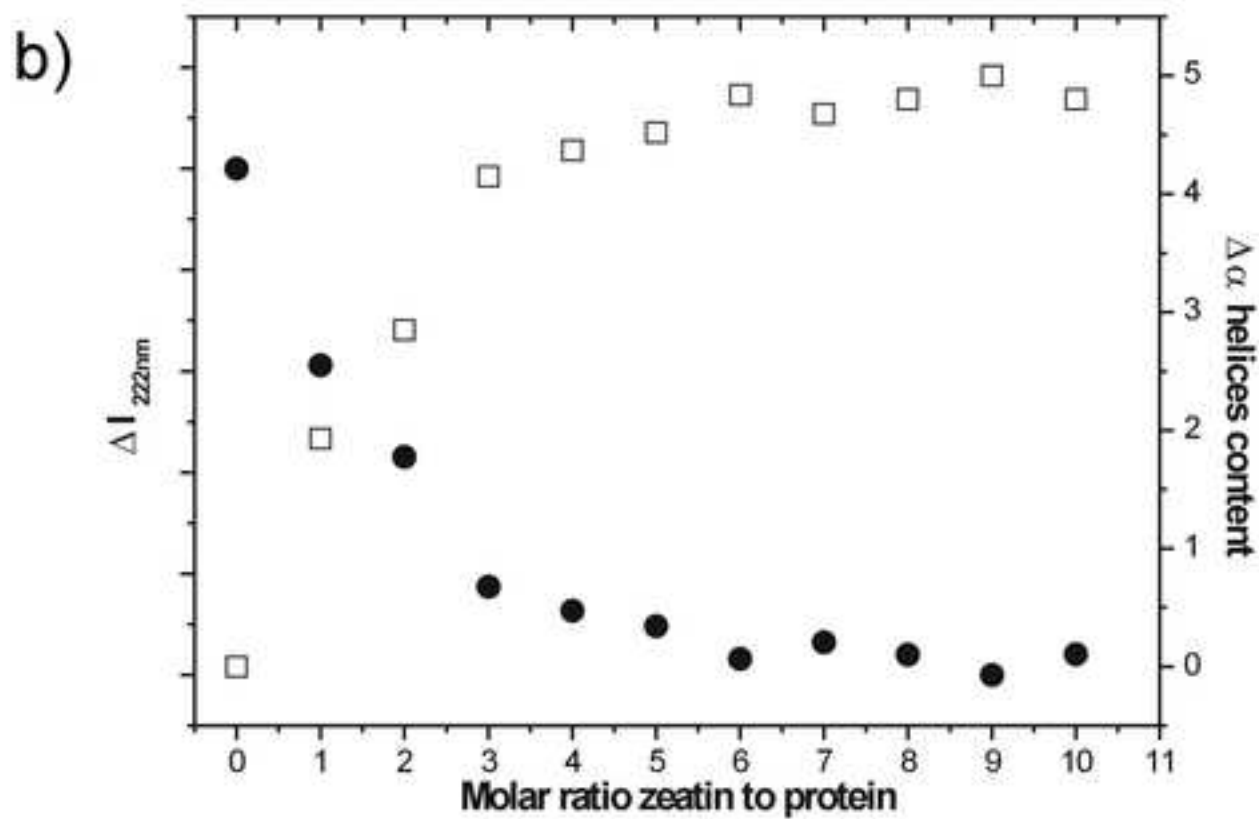
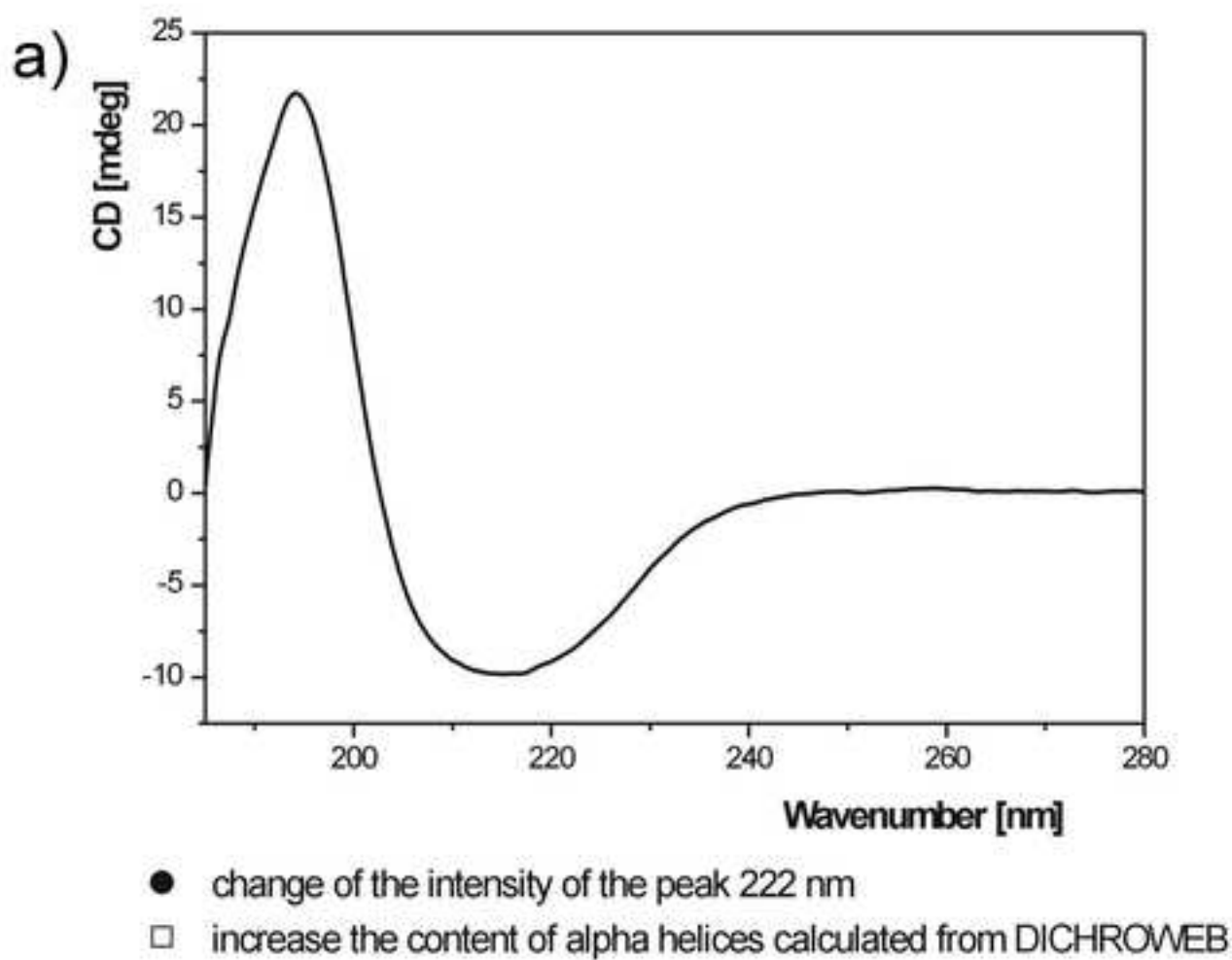
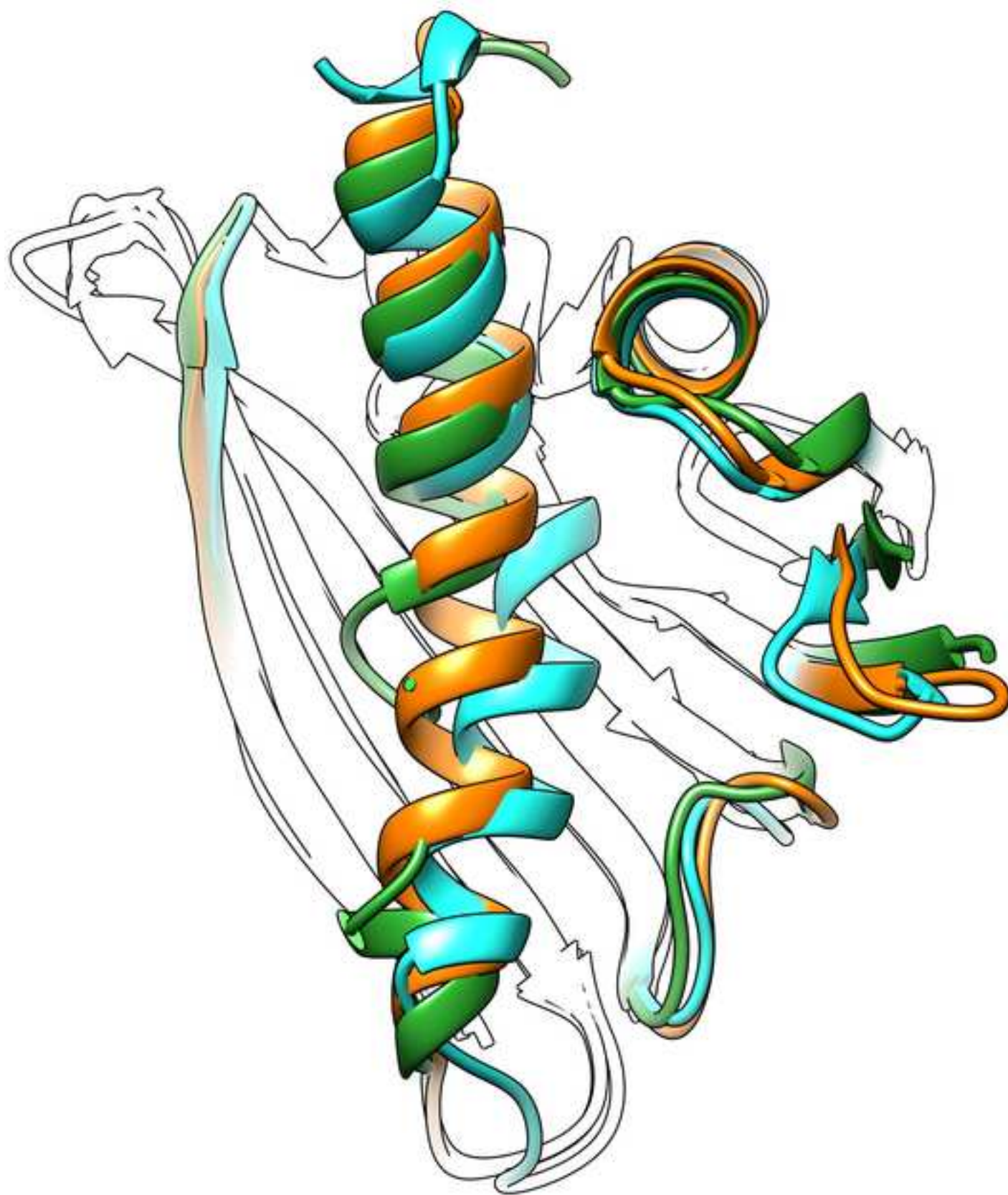
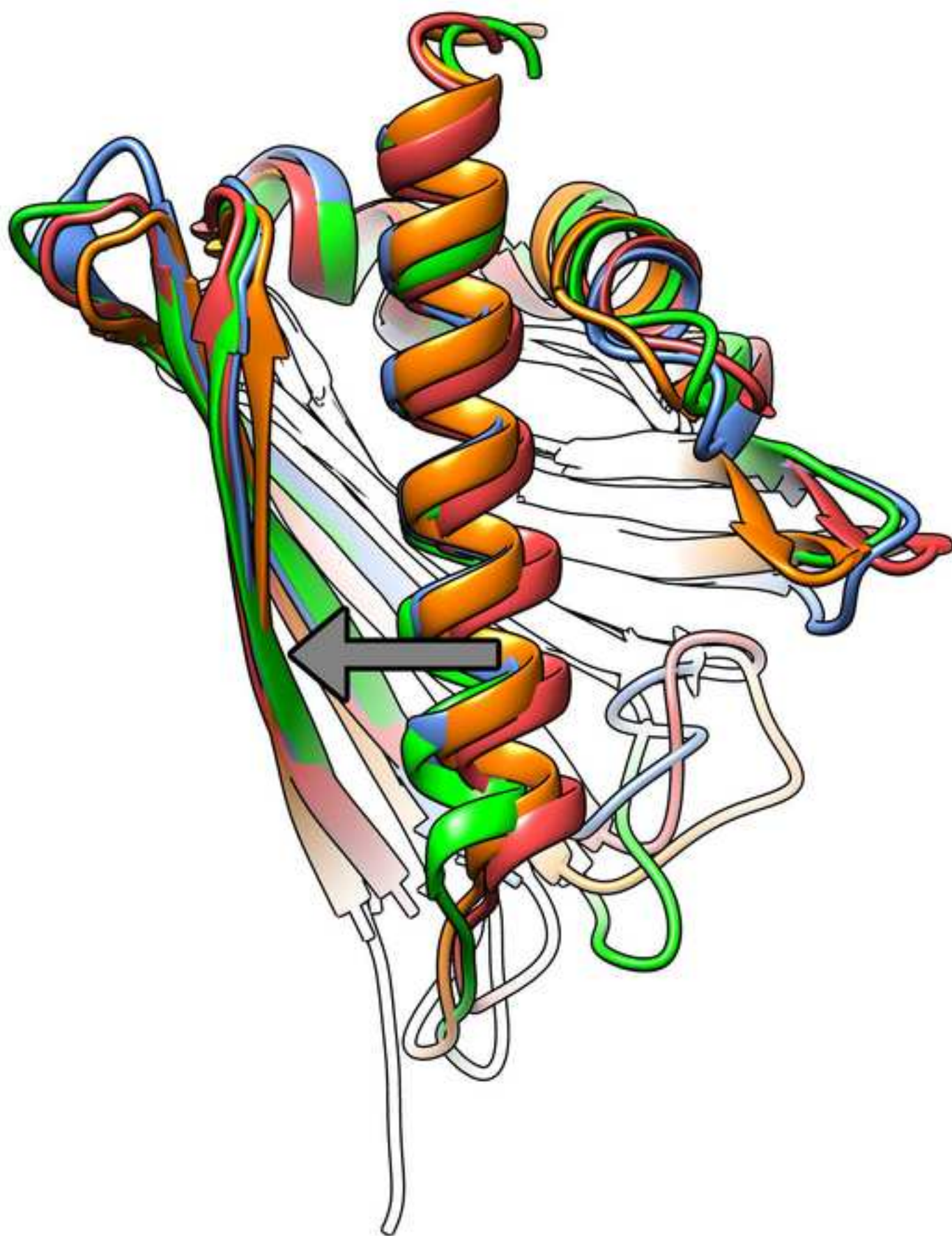


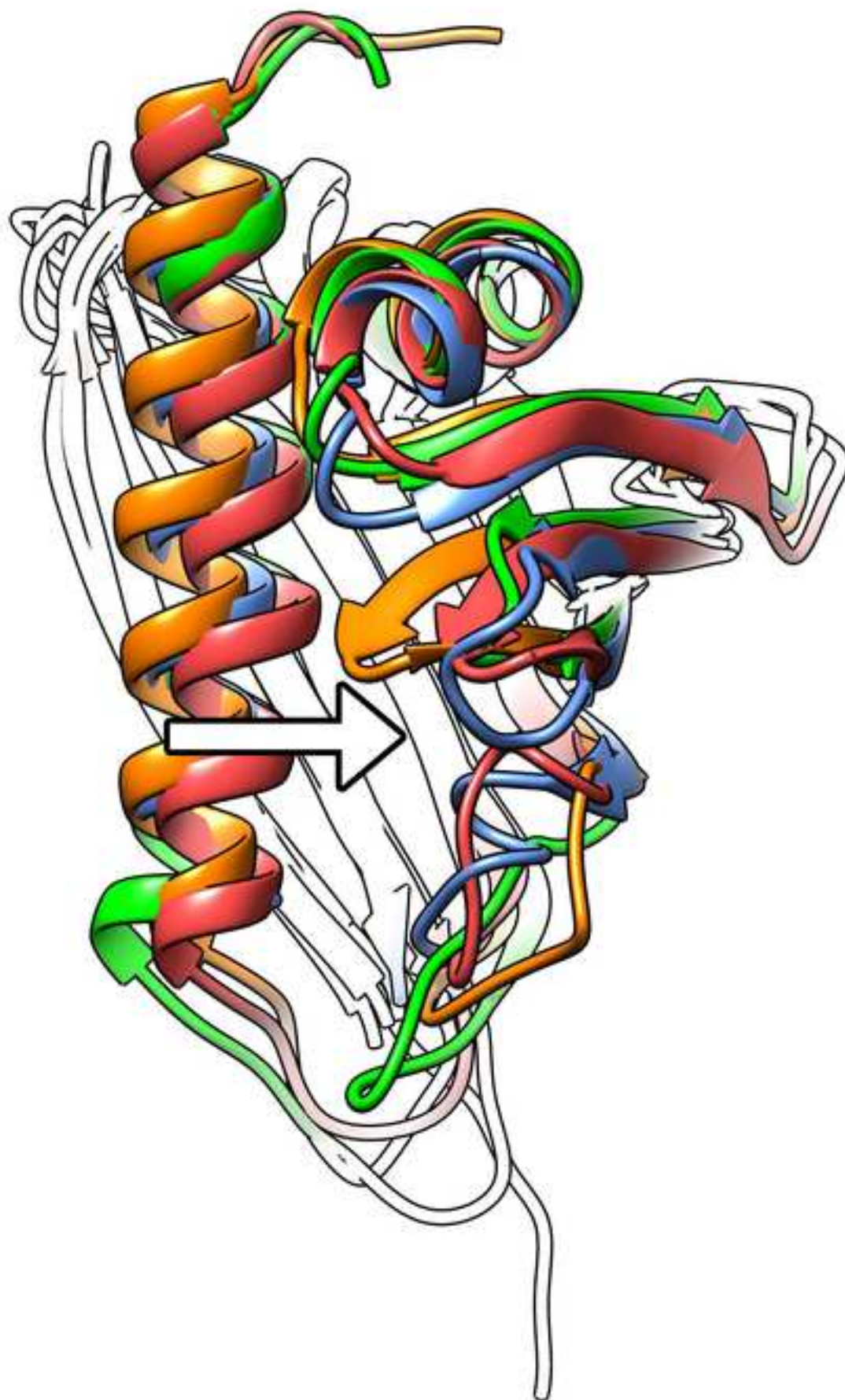
Figure 6

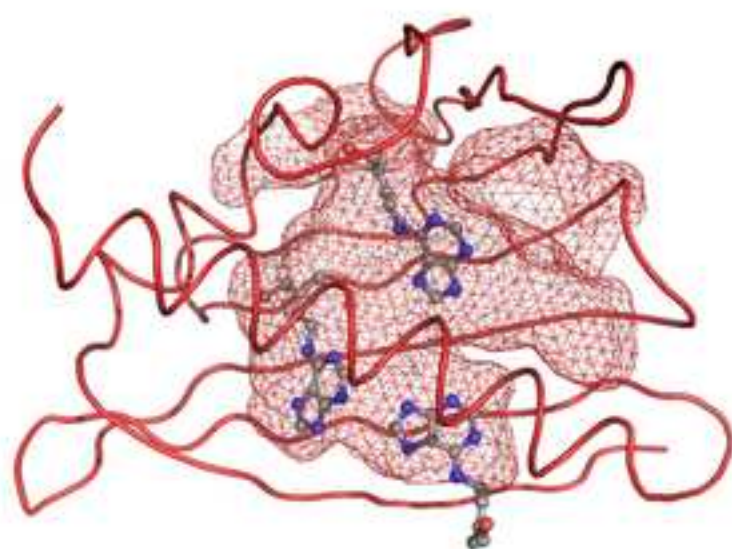




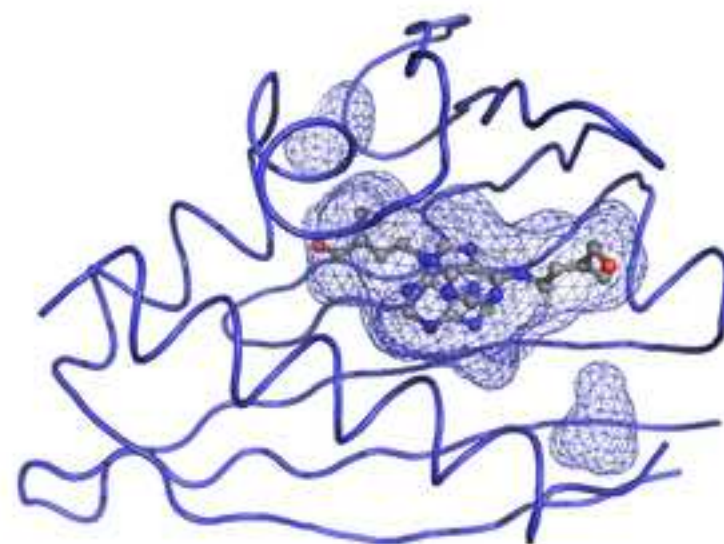




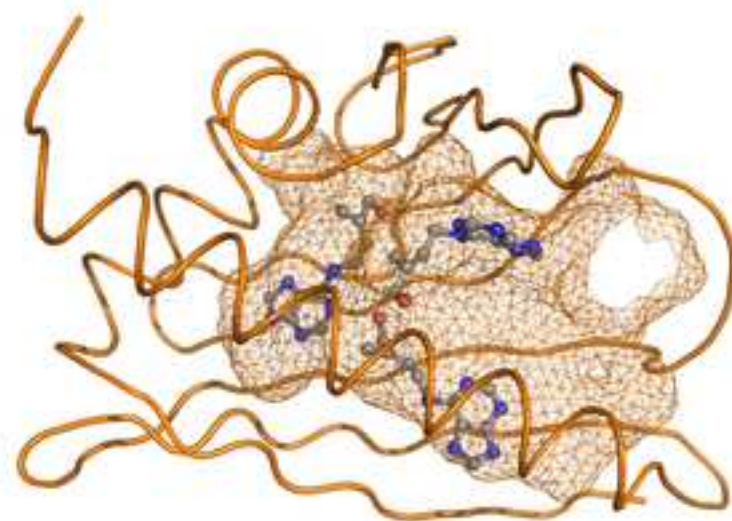




LIPR-10.1A/ZEA



VrPhBP/ZEA



LIPR-10.2B/ZEA



MtN13/ZEA