the FEBS Journal

Homooligomerization is needed for stability: a molecular modelling and solution study of *Escherichia coli* purine nucleoside phosphorylase

Branimir Bertoša[1], Goran Mikleušević[2], Beata Wielgus-Kutrowska[3], Marta Narczyk[3], Matea Hajnić[1], Ivana Leščić Ašler[2], Sanja Tomić[2], Marija Luić[2] and Agnieszka Bzowska[3]

1 Division of Physical Chemistry, Faculty of Science at University of Zagreb, Croatia
2 Division of Physical Chemistry, Ruđer Bošković Institute, Zagreb, Croatia
3 Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Poland

Keywords
analytical ultracentrifugation; CD spectra; homooligomeric proteins; molecular modelling; obligate (obligatory) oligomer; purine nucleoside phosphorylase

Correspondence
B. Bertoša, Division of Physical Chemistry, Department of Chemistry, Faculty of Science at University of Zagreb, Horvatovac 102, HR-10000 Zagreb, Croatia
Fax: +385 1 4606 131
Tel: +385 1 4606 132
E-mail: branimir.bertosa@irb.hr
A. Bzowska, Division of Biophysics, Institute of Experimental Physics, University of Warsaw, Żwirki i Wigury 93, 02-089 Warsaw, Poland
Fax: +48 22 554 0771
Tel. +48 22 554 0789
E-mail: abzowska@biogeo.uw.edu.pl

(Received 8 November 2013, revised 8 January 2014, accepted 7 February 2014)

doi:10.1111/febs.12746

Although many enzymes are homooligomers composed of tightly bound subunits, it is often the case that smaller assemblies of such subunits, or even individual monomers, seem to have all the structural features necessary to independently conduct catalysis. In this study, we investigated the reasons justifying the necessity for the hexameric form of *Escherichia coli* purine nucleoside phosphorylase – a homohexamer composed of three linked dimers – since it appears that the dimer is the smallest unit capable of catalyzing the reaction, according to the currently accepted mechanism. Molecular modelling was employed to probe mutations at the dimer–dimer interface that would result in a dimeric enzyme form. In this way, both *in silico* and *in vitro*, the hexamer was successfully transformed into dimers. However, modelling and solution studies show that, when isolated, dimers cannot maintain the appropriate three-dimensional structure, including the geometry of the active site and the position of the catalytically important amino acids. Analytical ultracentrifugation proves that *E. coli* purine nucleoside phosphorylase dimeric mutants tend to dissociate into monomers with dissociation constants of 20–80 μM. Consistently, the catalytic activity of these mutants is negligible, at least 6 orders of magnitude smaller than for the wild-type enzyme. We conclude that the hexameric architecture of *E. coli* purine nucleoside phosphorylase is necessary to provide stabilization of the proper three-dimensional structure of the dimeric assembly, and therefore this enzyme is the obligate (obligatory) hexamer.

Structured digital abstract
• PNP and PNP bind by molecular sieving (1, 2, 3, 4)

Introduction

Many proteins, including enzymes, are homooligomers, which means they are composed of several subunits with the same amino acid sequence [1,2]. In some cases need for such an architecture is directly related to the molecular mechanism, e.g. cooperativity between subunits exists, as in the well-known case of haemoglobin [3], or one subunit donates amino acid residues to the active site of the neighbour, as in thymidylate synthase [4], hexameric helicases [5] and others.

However, in some cases neither the structure itself nor the molecular mechanism of an enzyme gives an indication why a molecule adopts a homooligomeric

Abbreviations
MD, molecular dynamics; PNP, purine nucleoside phosphorylase; WT, wild-type.

structure. One such example is purine nucleoside phosphorylase (PNP, purine nucleoside: orthophosphate ribosyltransferase, EC 2.4.2.1), which participates in the biosynthesis of purine nucleosides through the salvage pathway, catalysing the reversible phosphorolysis of the glycosidic bond of ribonucleosides and deoxyribonucleosides [6]: β-purine nucleoside + phosphate ↔ purine base + α-D-pentose-l-phosphate. Differences in specificity between human and *Escherichia coli* PNPs offer a basis for enzyme-activating prodrug gene therapy by enabling selective killing of tumour cells expressing the *E. coli* PNP gene [7].

E. coli PNP is a homohexamer or more precisely the trimer of dimers [8–10]. In each subunit, the active site is present; however, two subunits composing a dimer mutually donate two amino acids, His4 and Arg43, to the active site of the neighbour (Fig. 1). Two possible dimer conformations have been observed in X-ray studies [10,11], which differ in terms of the accessibility of the active site pocket. In agreement, two binding constant values have been found for phosphate binding [10,12,13], and two sets of kinetic constants are necessary to describe the reaction with phosphate as a variable substrate [12,14]. All these suggest that the dimer might be a minimal catalytic unit. Hence, it seems puzzling why this enzyme is a hexamer with very tightly bound subunits, as indicated by ultracentrifugation analysis [15]. In this paper, we have attempted to address this question by *in silico* and in solution studies of mutants of *E. coli* PNP.

Methods

Molecular modelling

System preparation

Crystal structures of PNP from *E. coli* (1ECP and 1K9S, PDB [16]) were used for building dimeric and hexameric models in apo form and in complexes with substrates (formycin A was replaced by adenosine and sulfate ion by phosphate ion). Water molecules found in crystal structures within 0.3 nm from the protein heavy atoms were retained for calculations. Polar hydrogen atoms were added using the WHAT IF software [17], and non-polar with the *tleap* module of the AMBER 8 program package [18]. For parameterization of protein atoms, an AMBER ff03 force field was used [19]. Parameters for purine nucleoside and phosphate ion were obtained using the *antechamber* module of the AMBER 8 [18]. The protein was centred in the rectangular parallelepiped box filled with TIP3P water molecules [20]. Na^+ ions were added in order to neutralize the system [21].

Using the described procedure, eight systems were prepared for simulations: (a) two hexameric structures, one without substrates and another with substrates in all active sites (each of these two systems consisted of around 185 000 atoms); (b) six dimeric structures – two wild-type (WT) structures, one without substrates and one with substrates in both active sites, and four apo dimeric structures of mutants (mutants **13**, **6**, **0** and **16**, see Table 1; each of these six systems had around 64 000 atoms).

Simulations and geometry optimizations

All systems were energy minimized and geometry optimized in five cycles. Each cycle consisted of 1000 steps of the steepest descent algorithm followed by 4000 steps of a conjugated gradient algorithm. In the first cycle of minimization, a constraint of 418.4 kJ·(mol·Å^2)$^{-1}$ (100 kcal·(mol·Å^2)$^{-1}$) was applied on the substrate and protein atoms, while the solvent molecules were unconstrained. In the second cycle, constraint was applied on the substrate and protein heavy atoms. In the third cycle, the protein backbone and substrate atoms were constrained using a 418.4 kJ·(mol·Å^2)$^{-1}$ (100 kcal·(mol·Å^2)$^{-1}$) force constant. In the fourth cycle, the same atoms as in the third cycle were constrained but the force constant was decreased to 209.2 kJ·(mol·Å^2)$^{-1}$ (50 kcal·(mol·Å^2)$^{-1}$).

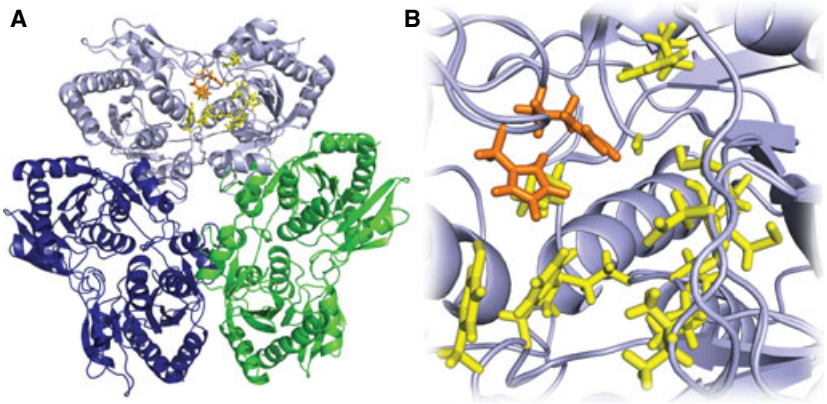


Fig. 1. (A) Structure of an apo PNP homohexamer that can be considered a trimer of dimers (each dimer is represented in a different colour). (B) One active site of the apo PNP hexamer. The active site of one monomer (whose amino acids are shown in yellow) is completed by two amino acids from the adjacent monomer (His4 and Arg43, indicated in orange) from a dimer.

Finally, in the fifth cycle of minimization, no constraints were applied.

After geometry optimization, systems were subjected to 6 ns of molecular dynamics (MD) simulations (only the hexameric system with substrates in all active sites was simulated up to 4.5 ns). Temperature was linearly increased from 0 to 300 K during the first 250 ps, followed by 50 ps of simulation at a constant temperature (300 K) using a Berendsen thermostat [22]. During the first 300 ps, protein and substrate atoms were constrained with a force constant of 104.6 kJ·(mol·Å^2)$^{-1}$ (25 kcal·(mol·Å^2)$^{-1}$) and volume was kept constant. From 300 ps to 6 ns, no constraints were applied and simulations were conducted at a constant temperature (300 K) and constant pressure (101 325 Pa) using a Berendsen thermostat and a barostat [22]. The time step was 1 fs, and the structures were sampled every picosecond. Periodic boundary conditions were applied. Particle mesh Ewald was used for calculation of electrostatic interactions. The cut-off value for non-bonded interactions was set to 1 nm. Structures were sampled every 1 ps over trajectories 6 ns long. All simulations were performed using AMBER 8 and GROMACS 4.5.1 simulation packages [18,23]. All trajectories were analysed using the VMD program [24] and GROMACS analysing tools.

In silico mutants

The hexameric structure of apo PNP obtained after 2 ns of MD simulation was used as a template for introducing mutations. Hexameric structures of the WT and constructed mutants were geometry optimized using the procedure described above and subjected to the *anal* module of AMBER 8, PISA and PROTORP web servers [18,25,26]. A similar procedure for investigating interactions between non-covalently bound subunits of oligomeric proteins had already been applied on a different system [27]. Mutants **13**, **6**, **0** and **16** (Table 1) were built in dimeric forms, geometry optimized and subjected to MD simulations.

Site-directed mutagenesis and sequencing

Site-directed mutagenesis was employed using a QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA, USA) and the set of specific primers for each amino acid substitution (listed in Table S2). Plasmid pSE380 with an inserted *deoD* gene (pSE380*deoD*) was used as a template in each PCR reaction. PCR amplification, selection of positive bacterial transformants and sequence confirmation of inserted mutations were performed as described previously [14]. Three PNP mutants were prepared: **0**, **6** and **13** (Table 1).

Overexpression and purification of PNP mutants

PNP mutants were overexpressed in the *E. coli* BL21 (DE3) strain according to [28]. A two-step purification procedure was employed with Q-Sepharose Fast Flow anion-exchange resin (GE Healthcare, Little Chalfont, UK) and Sephacryl S-200 Superfine gel filtration medium (Pharmacia, Uppsala, Sweden), as described for WT hexameric PNP [14]. The purity of the samples was verified using SDS/PAGE. The approximate molecular mass of each mutant was estimated from a gel filtration (Fig S1), using the following calibrants: chymotrypsinogen A (25 kDa), bovine serum albumin (67 kDa), aldolase (158 kDa) and catalase (232 kDa). In all experiments protein concentration was measured spectrophotometrically using $\varepsilon^{1\%}_{280\ nm} = 2.7\ cm^{-1}$ [10], and when given in molar units it always refers to the monomer concentration. The monomer masses of 25 832 Da, 25 630 Da and 25 685 Da for mutants **0**, **6** and **13** respectively were used in the calculations.

Circular dichroism analysis

CD spectra measurement was performed on a J-815 spectropolarimeter (Jasco, Tokyo, Japan), in the wavelength

Table 1. Interactions of dimers in the hexameric PNP structure, and stability of the hexameric WT and different mutants estimated using the *anal* module of AMBER 8 software and web servers PISA [25] and PROTORP [26]. Mutants are ordered according to their estimated potential to result in stable dimers in solution. Only the most promising mutants are shown here; the full list is shown in Table S1. VDW + ELE interactions are van der Waals and electrostatic interactions between two adjacent dimers. ΔG is the solvation free energy gain upon formation of the interface. NHB is the number of hydrogen bonds at the interface: PISA [25] values and PROTORP [26] values in parenthesis. NSB is the number of salt bridges at the interface: PISA [25] values and PROTORP [26] values in parenthesis. N_{mut} is the total number of mutations inserted.

Mutant	VDW + ELE interactions (kcal)	ΔG (kcal·mol)	Interface area (nm^2)	NHB	NSB	N_{mut}	Introduced mutations
WT	−17177	−21.8	16.1	22 (19)	11 (80)	0	None
Mutant **13**	−16422	−11.7	12.5	9	0	6	R117E, K121E, D139R, F120D, I128S, F131G
Mutant **6**	−16467	−20.4	13.9	9 (9)	0 (18)	4	R117E, K121E, D139R, F120D
Mutant **0**	−16842	−26.8	14.1	6 (6)	0 (21)	3	R117E, K121E, D139R
Mutant **16**	−16824	−22.7	13.1	6	0	4	R117E, K121E, D139R, Y173S

range 185–240 nm at 25 °C using a cuvette with a 0.1-mm path length, PNP samples, 0.55 mM of WT and 1.05 mM of mutants in 10 mM Tris/HCl buffer pH 7.0. Data analysis was performed on the DichroWeb using the SELCON 3 program [29,30].

Analytical ultracentrifugation

Analytical ultracentrifugation was conducted with an Optima XL-I (Beckman-Coulter Inc., Indianapolis, IN, USA) using the An-50Ti and An-60Ti analytical rotors and double-sector 1.2-cm cells with Epon-charcoal centrepieces and sapphire windows. Partial specific volume of mutants (from their amino acid composition), densities and viscosities of buffers were calculated using the SEDNTERP program [31].

For the sedimentation velocity studies, protein samples (0.6–3.7 mg·mL^{-1}, 390 μL) in 50 mM Tris/HCl buffer pH 7.6 or 7.4 were equilibrated at 700 ***g*** (3 000 r.p.m.) and 20 °C and accelerated to 142 000 ***g*** (at the cell bottom – 72 mm; 42 000 r.p.m.); radial scans of the protein absorption profile in the cell were measured at 5- or 7-min intervals. Sedimentation coefficients s and standard sedimentation coefficients $s_{20,w}$ were calculated, using the continuous $c(s)$ distribution model incorporated in the SEDFIT program [31], from which the molecular mass of the species present in the solution was derived, assuming that they have the same friction ratio (obtained as a parameter in the SEDFIT program).

For the sedimentation equilibrium studies, solutions of PNP mutants (2.2 mg·mL^{-1}) in 50 mM Tris/HCl buffer pH 7.6 (120 μL) were equilibrated at 15 000 r.p.m. and 25 000 r.p.m. (18 100 ***g*** and 50 400 ***g*** at the cell bottom, respectively) at 4 °C for 64 h. Protein concentration gradients in equilibrium were determined from absorbance at 280 nm, and were analysed by fitting a monomer–dimer self-association model with the SEDFIT and SEDPHAT programs [31].

Activity versus natural substrates and substrate analogue

The activity of all mutants was determined in 50 mM phosphate buffer pH 7.0, at 25 °C, with 0.5 mM for inosine, 0.3 mM for adenosine and 0.12–0.35 mM for 7-methylguanosine. For adenosine and 7-methylguanosine a direct spectrophotometric method was used, while for inosine a coupled assay with xanthine oxidase [32] was used. The concentration of mutants was in the range 0.025–0.5 mg·mL^{-1} (1–20 μM of monomers).

The progress of the reaction was followed in the scanning kinetic mode on a double-beam spectrophotometer (Cary 100; Agilent, Santa Clara, CA, USA) equipped with a Peltier thermostat for the experimental and reference cuvette holders. The reaction mixture, 0.5 mL, was placed in 5-mm path length quartz cuvettes (Hellma, Müllheim, Germany), experimental and reference. The reaction was started by adding a mutant to the experimental cuvette and the same volume of buffer to the reference cuvette in order to keep the nucleoside concentration identical. The progress of the reaction was followed by differential absorption spectra collected every 5 min. Initial spectra and final spectra for the experimental and the reference cuvettes versus buffer were also measured to determine if the overall changes in the spectra in the experimental cuvette corresponded to the phosphorolytic reaction, and to control the stability of the substrates in the absence of the mutant (reference cuvette). For each mutant, at least two experiments were performed with different protein concentrations.

The plot of absorbance versus time at wavelengths specific for each nucleoside was obtained as a cross-section of the absorption spectra. The initial velocity was determined by fitting a linear function to the initial fragment of the absorbance versus time plot. For calculating activities, the differential extinction coefficients were used (Table S3).

Results

MD simulations of WT dimers and hexamers in apo forms and in complexes with adenosine and phosphate in the active site

The instability of the WT PNP dimeric apo form was revealed during the 6 ns of MD simulation using the explicit solvent model (Figs 2, 3, S2 and S3). Large conformational changes, due to dimer exposure to the solvent, were observed, as measured by the root mean square deviation (RMSD) of Cα atoms (Fig. S2). In the apo form of the hexamer, the RMSD value is stable around 1 Å from the initial structure during 6 ns simulations, while in the apo dimeric form it increases up to 3.5 Å during this period. The highest differences in amino acid fluctuations (root mean square fluctuations) between hexamer and dimer were found for regions formed by residues 165–174 (monomers consist of 237 amino acid residues and the sequence 165–174 of the second monomeric subunit forming a dimer is equivalent to residues 402–411) (Fig. S3). Partly hydrophobic H6 helices constituted of these residues are situated at the hydrophobic interface between two dimers in the hexameric structure (Fig. 2) and during the 6 ns of MD simulations they remain in their initial positions (red line in Fig. 4B), while in the dimeric form they are exposed to solvent which has a destabilizing effect on the whole structure. Thus, during the simulations of dimers, they moved towards the centre of the dimer where they established interactions with

hydrophobic amino acid residues (Fig. 3), e.g. after the movement of H6 helices, Tyr173 is stabilized through interactions with Phe154, Asp124 and Cys110.

In silico mutations were introduced to check if it is possible to obtain stable dimeric forms

To check if stable dimeric forms may be obtained from the hexameric structure, amino acid mutations were proposed. The two main goals were: (a) to break the hexameric structure into dimers by destabilizing interactions at the interface between two dimers in a hexamer and (b) to stabilize these dimers in solution. Each proposed mutation, or combination of mutations, was built *in silico* and their influence on the hexamer and the dimer stability was computationally tested (Tables 1 and S1). Mutations were chosen according to the results of the MD simulations of the WT forms of the protein described in the previous section (dimer and hexamer in apo forms and in complexes with adenosine and phosphate). Mutations were introduced into the hexameric structure obtained after 2 ns of MD simulation of the WT apo form, since this structure was considered to be equilibrated in solution.

In the first step, hexameric structures of mutants were energy optimized through the optimization procedure described in Methods and subjected to the *anal* module of the AMBER 8 software [18] and web servers (PISA and PROTORP) [25,26]. These calculations were used for the rapid estimation of: (a) the energies of electrostatic and van der Waals interactions between the two dimers in the hexamer; (b) the difference in the free energy of solvation upon formation of the hexamer from the dimers; (c) the number of salt bridges and hydrogen bonds formed at the interface between two dimers; (d) the surface of a dimer–dimer interface in a hexamer. The values obtained for each mutant were compared with the respective values for the WT hexamer (Tables 1 and S1). We considered mutants with less favourable energy of non-covalent interactions between two dimers in a hexamer, a lower number of salt bridges and hydrogen bonds, less favourable solvation free energy gain upon formation of the interface and smaller interface area compared with the WT as the most likely to obtain stable dimers. In addition, the smallest possible number of mutations was also one of the criteria for choosing the mutants for further computational and experimental research. With this approach, mutants **13** (R117E, K121E, D139R, F120D, I128S, F131G), **6** (R117E, K121E, D139R, F120D), **0** (R117E, K121E, D139R) and **16** (R117E, K121E, D139R, Y173S) were identified as the most promising and were subjected to MD simulations.

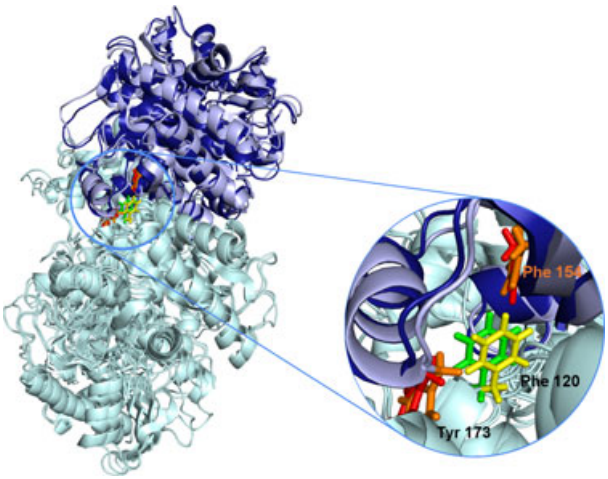


Fig. 2. Superimposed structures of the apo PNP hexamer: initial (blue) and after 5 ns (azure) of MD simulation. Amino acids Phe154 and Tyr173 are outlined and coloured red (0 ns) and orange (5 ns); Phe120 from an adjacent dimer is coloured green (0 ns) and yellow (5 ns).

MD simulation of dimeric forms of mutants

Although the **13**, **6**, **0** and **16** mutants showed greater stability compared with the WT dimer, changes in their structures compared with the WT dimer in a hexamer were still noticed and were rather pronounced. In the hexameric structure of the WT PNP, the RMSD of Cα atoms of the dimers did not vary during the 6 ns of simulation and had a constant value of around 1 Å compared with the crystal structure, while in the case of the apo WT dimer and mutant **16** the RMSD constantly increased and at the end of 6 ns of MD simulations reached the value of 3.5 Å (Fig. 4A). In the case of dimers **0** and **6** the RMSD increased slightly less, to about 2.5–3.0 Å (Fig. 4A). During the simulations, the RMSD of dimer **13** showed a behaviour most closely resembling that observed for a dimer in a hexamer, and at the end of the 6 ns simulation was about 1.5 Å (Fig. 4A).

The structural changes of the mutated dimers in the solution during 6 ns simulations was also measured using the RMSD of Cα atoms (Fig. 4A) and using translocation of H6 helices (measured as distances between the centres of mass of two H6 helices in a dimer) (Fig. 4B). Again, in comparison with the WT dimer, a slight stabilization of dimers **0** and **6** and substantial stabilization of dimer **13** were noticed (Fig. 4). Thus, MD simulations indicated the mutant **13** as the most promising combination of mutations for obtaining stable dimers in solution.

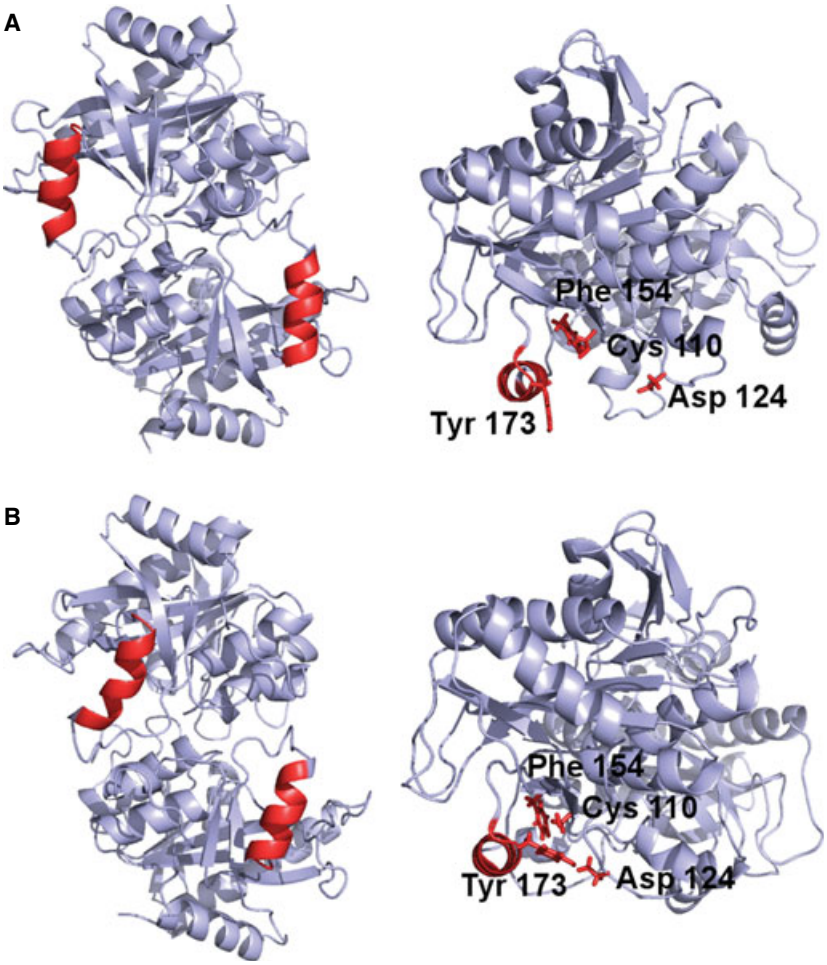


Fig. 3. Structure of a dimeric PNP apo form: (A) at the beginning; (B) after 6 ns of MD simulation. The right-hand panels show the respective structure of a dimer along the helix H6 (red). Amino acid residues Cys110, Asp124, Phe154 and Tyr173, which stabilize helix H6 in the dimeric form, are labelled. During the simulation, the distance between the centres of mass of helices H6 belonging to two monomers forming a dimer decreased from 39.4 Å to 32.0 Å.

MD simulations showing the influence of substrate presence on the stability of the active site

The tendency to lose the original shape of the active site was noticed in simulated apo forms of all studied dimers, both WT and mutated. In the absence of substrates, some of the amino acids forming the active site moved away from each other during the 6 ns of simulations (Fig. 5). In order to quantify this change, we measured the radius of gyration of the active site during the simulations (Fig. 4C). The most pronounced changes in the active site occurred for the apo form of the WT dimer. The largest deviation from the initial position was noticed for the following residues that form the active site: His4, Arg43 from adjacent monomers (Fig. 5, group 1), Phe159, Tyr160 (Fig. 5, group 2) and Asp204, Arg217 (Fig. 5, group 3). In all dimers, group 1 shifted the most from its original position in the hexamer, followed by residues from group 2, which interact in the complex with substrate molecules through two π–π interactions. In contrast, during the simulation of the WT dimer with nucleoside and phosphate, the active site retained the appearance found in the crystal structure.

Preparation of *in vitro* mutants

Mutants **13**, **6** and **0**, identified via *in silico* studies as the most likely to obtain stable dimers from the hexameric structure, were experimentally prepared and studied *in vivo*. The mutations were inserted into the *deoD* gene (coding for *E. coli* PNP) using site-directed PCR mutagenesis and mutants were overexpressed in the *E. coli* BL21(DE3) strain according to [28]. The mutants were purified in a two-step procedure used previously for active site mutants of *E. coli* PNP [14]. The elution volume from a Sephacryl S-200 gel filtration column for mutants **13**, **6** and **0** (Fig. S1) leads to a molecular mass of 35–37 kDa, which is between the values predicted for the monomer and dimer and hence indicates quick equilibrium of these two forms, which was verified by analytical ultracentrifugation (see below).

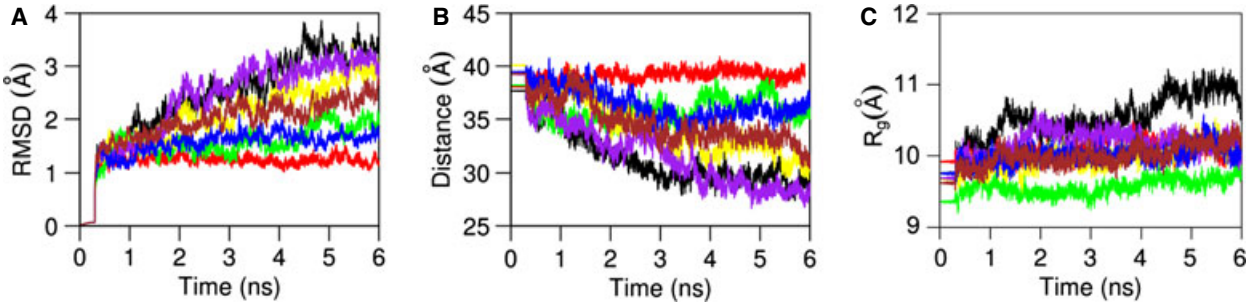


Fig. 4. RMSD of Cα atoms (A), distances between the centres of mass of helices H6 (B) and the radius of gyration of the active site (C) were monitored during 6 ns of MD simulations for the apo form of the WT dimer (black), the apo form of the WT hexamer (red), the WT dimer with substrates in the active site (green), mutant dimers in the apo form, mutant **0** dimer (yellow), mutant **6** dimer (brown), mutant **13** dimer (blue) and mutant **16** dimer (violet).

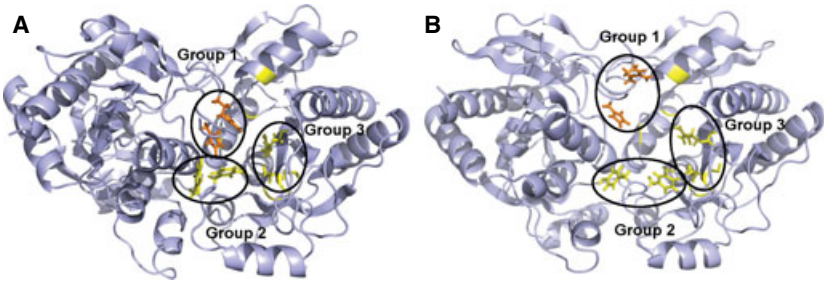


Fig. 5. Structure of the apo PNP dimer: (A) at the beginning and (B) after 6 ns of MD simulation. Amino acid residues of the active site belonging to one monomer are coloured yellow (groups 2 and group 3), amino acids from the adjacent monomer are coloured orange (group 1). All amino acid residues of the active site whose position evidently changed during the MD simulation are explicitly shown using the stick model.

CD spectra reveal some changes in the secondary structure of all three mutants compared with the WT PNP

CD analysis of mutants **13**, **6** and **0** (Fig. 6) shows that there are some small changes in the secondary structure due to the inserted mutations (Table S4). The content of helices increases with the number of mutations. Mutants **13** and **6** have the highest α-helix percentage, 31% compared with the WT PNP 27%, and slightly lower β-strand content, 17% compared with the 20% in the WT PNP. Mutant **0** shows the smallest changes, 29% of α-helices and 21% of β-strands. The small changes observed here were not noticed in the MD simulations so we cannot interpret them further on the basis of the analysis done.

Analytical ultracentrifugation shows that all three mutants are a mixture of monomeric and dimeric forms

In principle very similar results were obtained for **13**, **6** and **0** mutants in analytical ultracentrifugation, so Fig. 7 shows data for mutant **6** as an example. Sedimentation velocity experiments revealed two species present in the samples of mutants **13**, **6** and **0** (Fig. 7, panels A and B). The molecular masses of these species were determined as 25.7 and 51.4 kDa, 25.6 and 51.3 kDa and 25.8 and 51.7 kDa, for mutants **13**, **6** and **0**, respectively, which clearly correspond to the monomeric and dimeric forms. Integration of the area under the two observed peaks shows that most of the protein, about 80% in each case (Table S5), is in the monomeric form, and only about 20% of each mutant is able to form a dimer.

For mutant **0**, the sedimentation velocity experiment was repeated for several protein concentrations spanning the range 0.6–3.7 $mg{\cdot}mL^{-1}$ (25–150 μM of monomers). The results indicate (Table S5) that, for this mutant, the equilibrium dissociation constant K_d of a dimer disintegrating into monomers is in the protein concentration range used in the experiment, because at a lower concentration more protein was found in the monomeric state. Equilibrium dissociation constants for the self-association of monomers and formation of a dimer were therefore obtained from the equilibrium

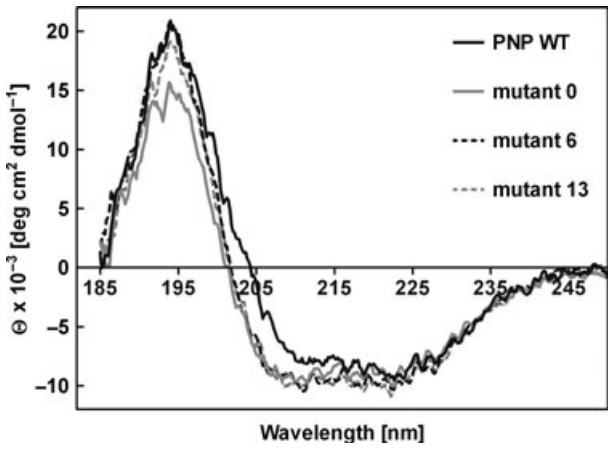


Fig. 6. CD spectra of 0.55 mM *E. coli* PNP WT and of 1.05 mM PNP mutants **0**, **6** and **13** in 10 mM Tris/HCl buffer pH 7.0.

sedimentation experiments (Table S6), as shown for mutant **6** in Fig. 7(C). Equilibrium dissociation constants are 28 ± 6 μM, 48 ± 4 μM and 25 ± 5 μM for mutants **0**, **6** and **13**, respectively. The association rate constants for protein–protein interaction span a wide range, from 10^3 $M^{-1}{\cdot}s^{-1}$ to 10^{10} $M^{-1}{\cdot}s^{-1}$ [33]. However, even with the lower association rate constant limit, 10^3 $M^{-1}{\cdot}s^{-1}$, the dissociation rate constant is about 10^{-2} s^{-1}, which confirms that none of the mutants is able to form a stable dimer.

Kinetic studies show that all three mutants have negligible phosphorolytic activity

The possible phosphorolytic activity of mutants **13**, **6** and **0** towards various nucleosides, natural substrates, adenosine and inosine, and a well-known fluorescent substrate 7-methylguanosine, was studied. No activity was detected under standard conditions for WT *E. coli* PNP [6,10,14], so a much higher protein concentration was used (up to 0.5 $mg{\cdot}mL^{-1}$) and the experiment was conducted for several hours. In order to eliminate the possible small instability of the substrates, the differential absorption spectra of the reaction mixture, with and without each mutant, were collected.

The reaction observed for inosine and mutant **13** is shown as an example in Fig. 8. Changes in the spectra (except for adenosine and mutants **0** and **13**, see below), although very small, mostly occur in those spectral regions characteristic for the phosphorolytic reaction, and show isosbestic points indicating one process. Phosphorolytic activities based on these data, however, are at best about 10^{-5} $U{\cdot}mg^{-1}$ (Table 2). Moreover, we cannot exclude the possibility that the phosphorolytic reactions observed are not partly or entirely due to traces of WT PNP in the purified mutant samples, so values in Table 2 should be regarded as upper limits of the phosphorolytic activity. On the other hand we expect that only the dimeric form can be active, which is present in a maximum of 20% in the conditions used in kinetic studies, so the specific activity will be several-fold higher when calculated only for the protein in a dimeric form. Nevertheless, the final conclusion is that all the investigated mutants show only extremely small, residual phosphorolytic activity, at best 5 orders of magnitude lower than that of the hexameric WT enzyme.

Discussion

As mentioned, the biologically active form of *E. coli* PNP is a hexamer, better described as a trimer of dimers [8]. Crystal structures [8–11,14] and enzyme kinetic and ligand binding studies [10,12–14] have pointed to a dimer as a sufficient catalytic unit. In this study, we tried to understand why the hexameric form is necessary. Mutations into the hexamer were introduced with the goal of obtaining stable dimers. In order to make experimental investigation more efficient and faster, computational methods were used as a guideline.

MD simulations in solution revealed large conformational changes of the WT dimer, mostly due to hydrophobic effects, leading to loss of the original shape of the molecule (observed in the hexamer). These changes consist of the movement of H6 helices (amino acids 165–174), which are mostly hydrophobic and in the hexameric structure are situated in the hydrophobic environment at the interface of two dimers. Due to exposure of H6 to the solvent in the dimeric form, they have a tendency to move towards the more hydrophobic centre of the dimer.

Besides the overall conformational changes in the dimer structure, MD simulations revealed active site deformation, including translocation of catalytically important amino acid residues: Asp204 and Arg217 involved in the protonation of the purine base according to the proposed molecular catalytic mechanism [10], Phe159 and Tyr160 forming π–π interactions with the substrate, and His4 and Arg43 from the neighbouring monomer lining the active site [9]. Hence, according to the simulations, the isolated WT dimer does not have a stable structure in solution and therefore, most probably, is not able to conduct catalysis.

MD simulations of several mutants in dimeric form were conducted to check their stability in solution.

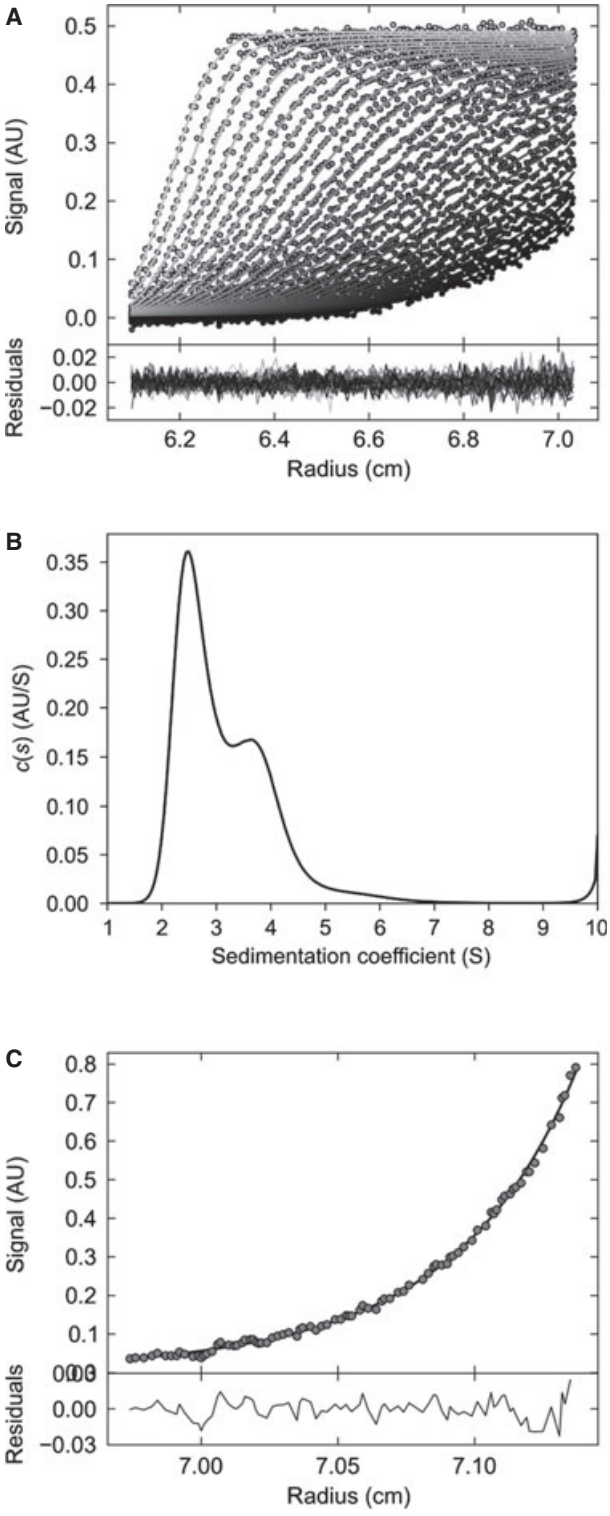


Fig. 7. (A) Determination of the oligomeric state of mutant **6** (2.2 mg·mL^{-1}) via analytical centrifugation in 50 mM Tris/HCl buffer pH 7.6. (A) Sedimentation velocity profiles collected every 15 min during centrifugation at 42 000 r.p.m. (142 000 ***g*** at the cell bottom) at 20 °C. Continuous *c*(*s*) distribution model [31] was fitted (SEDFIT program) to the data, with the residues shown in the lower part of (A). (B) The $A_{280\ nm}(s)$ profile obtained from fitting the continuous *c*(*s*) distribution model to the data presented in (A). Two species with sedimentation coefficients 2.64 S and 4.03 S were detected (standard sedimentation coefficients $S_{w,20} = 2.78$ and $S_{20,w} = 4.24$). This yields, with the best-fit friction ratio = 1.183, molecular masses 27.7 kDa and 52.2 kDa, respectively, and hence the monomeric and dimeric forms. (C) Equilibrium sedimentation profile obtained at 25 000 r.p.m. (50 400 ***g*** at the cell bottom) at 4 °C and the residuals plot for fitting the monomer–dimer self-association model using the SEDFIT program [31] with a dissociation constant of 74 ± 4 μM. Prepared using Gussi [37].

Four mutants with three to six mutations, mutants **13**, **6**, **0** and **16** (Table 1, Fig. S4), were chosen as the most promising in attempts to obtain stable dimers from the hexamer by introducing proper mutations. They have less favourable interaction energy between dimers in the hexamer in comparison with the WT PNP which might lead to a breakdown of the hexameric structure of the dimers in solution. Furthermore, most of them (except for mutant **16**) showed larger stability in comparison with the WT dimer. Moreover, MD calculations show that the presence of substrates in the active site leads to some increases in dimer stability similar to earlier experimental results obtained for the WT *E. coli* PNP [14,34]. However, the computationally observed positive effect on dimer stability due to the presence of substrates was not sufficient to fully stabilize the active site and to prevent conformational changes in the dimeric form.

Thus, the results of *in silico* investigations indicated that the *E. coli* PNP can be stable and active only in the hexameric form. We decided to check this conclusion experimentally. The most stable mutants from *in silico* studies, **13**, **6** and **0**, were constructed using molecular biology methods, purified and investigated via various experimental techniques.

CD spectra of all three mutants disclose some small changes in their secondary structure. The percentage of helices increases upon mutation at the cost of β-sheet content. Although these changes are not very large, and do not exceed 5% of the whole structure, their presence is in line with the *in silico* results indicating that the isolated dimer is not able to fully keep the structure that it has as a part of the hexameric assembly.

Gel filtration revealed the presence of species with a molecular mass of about 35 kDa, 37 kDa and 35 kDa for the mutants **13**, **6** and **0**, respectively, consistent with a rather quick monomer–dimer equilibrium. In fact, sedimentation velocity experiments confirmed the presence of monomers and dimers for all three

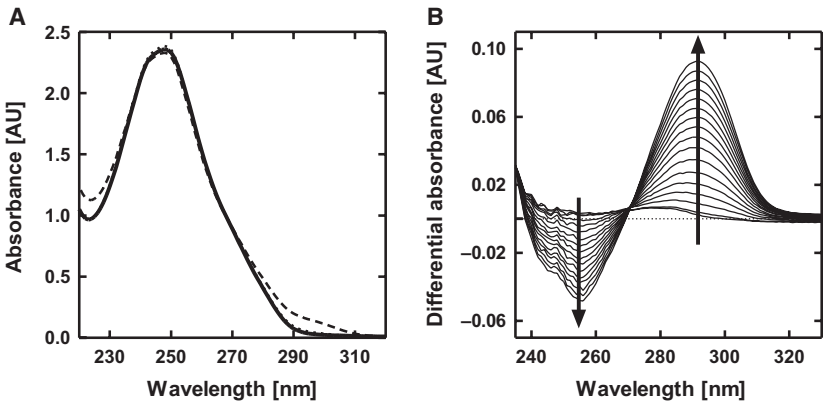


Fig. 8. (A) Scanning kinetics experiment for inosine in 50 mM phosphate buffer pH 7.0 at 25 °C with 0.05 mg·mL^{-1} of mutant **13**. Initial inosine concentration was 500 μM and the path length of the cuvette was 5 mm. Initial spectra of the reaction mixture (solid line), and spectra of the reaction mixture after 16 h with (dashed line) and without (dotted line) mutant **13**. (B) Differential absorbance spectra for the same experiment as in (A). Spectra were collected every hour. Arrows show directions of changes in the two regions of the spectra (maxima of inosine and of uric acid absorbance, the final product of the coupled PNP assay).

mutants, with the monomeric form more abundant, comprising about 80% of all molecules present in solution in our experimental conditions (2.2 mg·mL^{-1} protein). The dissociation constants 28 ± 6 μM, 48 ± 4 μM and 25 ± 5 μM for mutants **0**, **6** and **13**, respectively, confirm that the dimeric form is not fully stable. This behaviour contrasts to that of the WT *E. coli* PNP, which was shown to be completely stable as a hexamer [15] and, moreover, does not dissociate into dimers, since its disruption leads in one step to the unfolded protein with no dimer as an intermediate [15].

Comprehensive activity tests on mutants **13**, **6** and **0** showed that the upper limit of specific activity versus typical substrates is about 10^{-5} U·mg^{-1}, hence almost negligible compared with the WT hexameric form of the enzyme, which has a specific activity versus inosine, adenosine and 7-methylguanisine of 103.2 U·mg^{-1}, 58.5 U·mg^{-1} and 21.4 U·mg^{-1}, respectively [14].

Table 2. The upper limit of the phosphorolytic activity of all the studied mutants. The activity was determined in 50 mM phosphate buffer pH 7.0 at 25 °C in a prolonged kinetic experiment with high mutant concentration (0.025–0.5 mg·mL^{-1}). For comparison, phosphorolytic activity for the hexameric WT *E. coli* PNP is also included [14]. ND, not detected due to the presence of traces of an enzyme, most probably adenosine deaminase, which converted adenosine to inosine, hence making detection of a very low level of phosphorolytic activity impossible.

Mutant	Inosine (U·mg^{-1})	Adenosine (U·mg^{-1})	7-Methylguanosine (U·mg^{-1})
0	(3.7 ± 1.1) × 10^{-5}	ND	(3.5 ± 0.1) × 10^{-5}
6	(0.9 ± 0.2) × 10^{-5}	~0.4 × 10^{-5} [a]	(0.11 ± 0.04) × 10^{-5}
13	~0.2 × 10^{-5} [a]	ND	~7 × 10^{-5} [a]
WT	103.2 ± 4.0	58.5 ± 5.9	21.4 ± 1.6

[a] In this case the error was found to be higher than 30%.

Conclusions

All the presented investigations show that *E. coli* PNP needs a hexameric scaffold to be able to fulfil its biological function. Modelling indicates that the isolated dimer cannot maintain the three-dimensional structure, including the shape of the active site, characteristic for this molecule when it is a part of a hexameric assembly. Mutations that partially stabilize the structure of the isolated dimer were identified, but none of them was able to completely prevent unfavourable conformational changes in the whole molecule and in the active site, as shown by *in silico* and in the experimental, wet, part of this work. In line with these findings, none of the mutants shows a reasonable catalytic activity.

Our studies therefore suggest that the hexameric form of *E. coli* PNP, or more precisely the trimer of dimer architecture, is necessary to allow the individual dimer to be a dimer (otherwise it tends to dissociate into monomers) and to keep its three-dimensional structure and the active site shape necessary for efficient catalysis. This is fully in line with one of the possible functions of homooligomerization already proposed and discussed in the literature [1,2] and demonstrated in numerous studies to be the case for particular proteins. For example, homodimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima* and homotrimeric 2′-deoxyuridine triphosphate pyrophosphatase have been shown to form oligomers for reasons of stability [35,36], and our parallel studies with trimeric PNP led to the same conclusion regarding the trimeric architecture of

this enzyme (B. Wielgus-Kutrowska, A. Modrak-Wojcik, A. Dyzma, M. Zolkiewski & A. Bzowska, unpublished results). Obviously, the homooligomeric form is crucial for the biological function of these enzymes; hence, they may be regarded as obligate (obligatory) oligomers that are defined as assemblies where components of the complex (in most cases these are monomers; in the case of *E. coli* PNP – dimers) are unstable and usually non-functional upon isolation [1,2].

Acknowledgements

Plasmid pSE380 containing the *deoD* gene (pSE380-*deoD*) was a gift from Joanne L. Turnbull, Concordia University, Montreal, Canada. We thank Biserka Žinić and Jacek Jemielity for the synthesis of 7-methylguanosine, Lucyna Magnowska for the excellent technical assistance in all studies in solution, and Nina Jajčanin Jozić for help with collection and analysis of CD spectra.

These studies were supported by the Ministry of Science, Education and Sports of the Republic of Croatia (grant no. 098-1191344-2943, 098-1191344-2860) and the Polish Ministry of Science and Higher Education (grant no. N N301 044939). This work was partially performed in the NanoFun laboratories co-financed by the European Regional Development Fund within the Innovation Economy Operational Program, Project no. POIG.02.02.00-00-025/09/.

References

1 Hashimoto K, Nishi H, Bryant S & Panchenko AR (2011) Caught in self-interaction: evolutionary and functional mechanisms of protein homooligomerization. *Phys Biol* **8**, 035007.

2 Nishi H, Hashimoto K, Madej T & Panchenko AR (2013) Evolutionary, physicochemical, and functional mechanisms of protein homooligomerization. *Prog Mol Biol Trans Sci* **117**, 3–24.

3 Monod J, Wyman J & Changeux JJ (1965) On the nature of allosteric transitions: a plausible model. *J Mol Biol* **12**, 88–118.

4 Matthews DA, Appelt K & Oatley SJ (1989) Crystal structure of *Escherichia coli* thymidylate synthase with FdUMP and 10-propargyl-5,8-dideazafolate. *Adv Enzyme Regul* **29**, 47–60.

5 Medagli B & Onesti S (2013) Structure and mechanism of hexameric helicases. *Adv Exp Med Biol* **973**, 75–95.

6 Bzowska A, Kulikowska E & Shugar D (2000) Purine nucleoside phosphorylases: properties, functions, and clinical aspects. *Pharmacol Therap* **88**, 349–425.

7 Bennett EM, Anand R, Allan PW, Hassan AE, Hong JS, Levasseur DN, McPherson DT, Parker WB, Secrist JA 3rd, Sorscher EJ *et al.* (2003) Designer gene therapy using an *Escherichia coli* purine nucleoside phosphorylase/prodrug system. *Chem Biol* **10**, 1173–1181.

8 Mao C, Cook WJ, Zhou M, Koszalka GW, Krenitsky TA & Ealick SE (1997) The crystal structure of *Escherichia coli* purine nucleoside phosphorylase: a comparison with the human enzyme reveals a conserved topology. *Structure* **5**, 1373–1383.

9 Koellner G, Luić M, Shugar D, Saenger W & Bzowska A (1998) Crystal structure of the ternary complex of *E. coli* purine nucleoside phosphorylase with formycin B, a structural analogue of the substrate inosine, and phosphate (sulphate) at 2.1 Å resolution. *J Mol Biol* **280**, 153–166.

10 Koellner G, Bzowska A, Wielgus-Kutrowska B, Luić M, Steiner T, Saenger W & Stępiński J (2002) Open and closed conformation of the *E. coli* purine nucleoside phosphorylase active center and implications for the catalytic mechanism. *J Mol Biol* **315**, 351–371.

11 Bennett EM, Li C, Allan PW, Parker WB & Ealick SE (2003) Structural basis for substrate specificity of *Escherichia coli* purine nucleoside phosphorylase. *J Biol Chem* **278**, 47110–47118.

12 Jensen KF (1976) Purine nucleoside phosphorylase from *Salmonella typhimurium* and *Escherichia coli*. Initial velocity kinetics, ligand binding and reaction mechanism. *Eur J Biochem* **67**, 377–386.

13 Kierdaszuk B, Modrak-Wojcik A & Shugar D (1997) Binding of phosphate and sulfate anions by purine nucleoside phosphorylase from *E. coli*: ligand-dependent quenching of enzyme intrinsic fluorescence. *Biophys Chem* **63**, 107–118.

14 Mikleušević G, Štefanić Z, Narczyk M, Wielgus-Kutrowska B, Bzowska A & Luić M (2011) Validation of the catalytic mechanism of *E. coli* purine nucleoside phosphorylase by structural and kinetic studies. *Biochimie* **93**, 1610–1622.

15 Modrak-Wójcik A, Stepniak K, Akoev V, Żókiewski M & Bzowska A (2006) Molecular architecture of *E. coli* purine nucleoside phosphorylase studied by analytical ultracentrifugation and CD-spectroscopy. *Prot Sci* **15**, 1794–1800.

16 Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissing H, Shindyalov IN & Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242.

17 Vriend G (1990) WHAT IF: a molecular modelling and drug design program. *J Mol Graph* **8**, 52–56.

18 Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA *et al.* (2004) AMBER 8. University of California, San Francisco.

19 Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang Y, Cieplak P, Luo R, Lee T *et al.* (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* **16**, 1999–2012.
20 Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926–935.
21 Åqvist J (1990) Ion−water interaction potentials derived from free energy perturbation simulations. *J Phys Chem* **94**, 8021–8024.
22 Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A & Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* **81**, 3684–3690.
23 Hess B, Kutzner C, van der Spoel D & Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* **4**, 435–447.
24 Humphrey W, Dalke A & Schulten K (1996) VMD − visual molecular dynamics. *J Molec Graphics* **14**, 33–38.
25 Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
26 Reynolds C, Damerell D & Jones S (2009) ProtorP: a protein−protein interaction analysis server. *Bioinformatics* **25**, 413–414.
27 Brgles M, Bertoša B, Winkler W, Kurtović T, Allmaier G, Marchetti-Deschmann M & Halassy B (2012) Chromatography, mass spectrometry, and molecular modeling studies on ammodytoxins. *Anal Bioanal Chem* **402**, 2737–2748.
28 Esipov RS, Gurevich AI, Chuvikovsky DV, Chupova LA, Muravyova TI & Miroshnikov AI (2002) Overexpression of *Escherichia coli* genes encoding nucleoside phosphorylases in the pET/Bl21(DE3) system yields active recombinant enzymes. *Prot Exp Purif* **24**, 56–60.
29 Sreerema N & Woody RW (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal Biochem* **209**, 32–44.
30 Sreerama N, Venyaminov SY & Woody RW (1999) Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. *Protein Sci* **8**, 370–380.
31 Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys J* **78**, 1606–1619. (www.analyticalultracentrifugation.com/).
32 Kalckar HM (1947) Differential spectrophotometry of purine compounds by means of specific enzymes. I. Determination of hydroxypurines. *J Biol Chem* **167**, 429–443.
33 Schlosshauer M & Baker D (2004) Realistic protein−protein association rates from a simple diffusional model neglecting long-range interactions, free energy barriers, and landscape ruggedness. *Protein Sci* **13**, 1660–1669.
34 Krenitsky TA & Tuttle JV (1982) Correlation of substrate stabilization patterns with proposed mechanisms for three nucleoside phosphorylases. *Biochim Biophys Acta* **703**, 247–249.
35 Takács E, Grolmusz VK & Vértessy BG (2004) A tradeoff between protein stability and conformational mobility in homotrimeric dUTPases. *FEBS Lett* **566**, 48–54.
36 Thoma R, Hennig M, Sterner MR & Kirschner K (2000) Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*. *Structure* **8**, 265–276.
37 Brautigam CA (2011) Gussi, version 1.0.0. http://biophysics.swmed.edu/MBR/software.html.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Fig. S1. Gel filtration purification step of PNP wild-type and mutants **0**, **6** and **13**.

Fig. S2. Root mean square deviation of Cα atoms during MD simulations of the apo form of the dimer (black) and the apo form of the hexamer (red) during MD simulations.

Fig. S3. Root mean square fluctuations determined for entire residues during 6 ns of MD simulations (residues 1–237 belong to one monomer and residues 238–474 to the other monomer from the same dimer).

Fig. S4. Interface between two adjacent dimers in the apo PNP homohexamer structure.

Table S1. Interactions of dimers in the hexameric structure, and stability of the hexamer for the wild-type and for different mutants estimated using the *anal* module of AMBER 8 software and web servers PISA [25] and PROTORP [26].

Table S2. DNA sequence of specific primers for each amino acid substitution.

Table S3. Spectral data and concentrations of nucleosides used in the activity determination.

Table S4. Content of the secondary structure element in the hexameric WT PNP and three mutants according to the analysis of their CD spectra.

Table S5. Results of the sedimentation velocity experiments for mutants **0**, **6** and **13**.

Table S6. Results of the equilibrium sedimentation experiments for mutants **0**, **6** and **13** (2.2 $mg \cdot mL^{-1}$ protein in 50 mM Tris/HCl buffer pH 7.6 and 4 °C).