Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Conformational changes in human prolyl-tRNA synthetase upon binding of the substrates proline and ATP and the inhibitor halofuginone

Jonghyeon Son,[a]‡ Eun Hye Lee,[a]‡ Minyoung Park,[b,c] Jong Hyun Kim,[b,c] Junsoo Kim,[a] Sunghoon Kim,[b,c] Young Ho Jeon[d] and Kwang Yeon Hwang[a]*

[a]Department of Biosystems and Biotechnology, Korea University, Anam-dong 5, Seongbuk-gu, Seoul 136-701, Republic of Korea, [b]Medicinal Bioconvergence Research Center, Advanced Institutes of Convergence Technology, Suwon, Gyeonggi 443-759, Republic of Korea, [c]College of Pharmacy, Seoul National University, Seoul 151-742, Republic of Korea, and [d]College of Pharmacy, Korea University, Sejong-ro 2511, Sejong, Chungnam 136-713, Republic of Korea

‡ These authors made equal contributions.

Correspondence e-mail: chahong@korea.ac.kr

Received 10 May 2013
Accepted 24 July 2013

PDB References: *h*PRS-apo, 4k86; *h*PRS-sub, 4k87; *h*PRS-HF, 4k88

Aminoacyl-tRNA synthetases recognize cognate amino acids and tRNAs from their noncognate counterparts and catalyze the formation of aminoacyl-tRNAs. Halofuginone (HF), a coccidiostat used in veterinary medicine, exerts its effects by acting as a high-affinity inhibitor of the enzyme glutamyl-prolyl-tRNA synthetase (EPRS). In order to elucidate the precise molecular basis of this inhibition mechanism of human EPRS, the crystal structures of the prolyl-tRNA synthetase domain of human EPRS (*h*PRS) at 2.4 Å resolution (*h*PRS-apo), of *h*PRS complexed with ATP and the substrate proline at 2.3 Å resolution (*h*PRS-sub) and of *h*PRS complexed with HF at 2.62 Å resolution (*h*PRS-HF) are presented. These structures show plainly that motif 1 functions as a cap in *h*PRS, which is loosely opened in *h*PRS-apo, tightly closed in *h*PRS-sub and incorrectly closed in *h*PRS-HF. In addition, the structural analyses are consistent with more effective binding of *h*PRS to HF with ATP. Mutagenesis and biochemical analysis confirmed the key roles of two residues, Phe1097 and Arg1152, in the HF inhibition mechanism. These structures will lead to the development of more potent and selective *h*PRS inhibitors for promoting inflammatory resolution.

1. Introduction

The aminoacylation of tRNAs, which involves the attachment of amino acids to a specific tRNA, is catalyzed by aminoacyl-tRNA synthetases (ARSs; Schimmel, 1987). These highly specific enzymes catalyze the ligation of amino acids to their cognate tRNAs (Favorova, 1984). ARSs are well known as housekeeping enzymes; however, recent studies have demonstrated that they are also drug targets for the treatment of cancer, inflammatory diseases and autoimmune diseases (Park *et al.*, 2005, 2008). Chordate ARSs have evolved distinctive features that are not present in the ancestral forms, including compartmentalization in a multisynthetase complex (MSC), noncatalytic peptide appendages and ancillary functions unrelated to aminoacylation (Kim *et al.*, 2011; Sampath *et al.*, 2004).

ARSs are divided into two distinct classes based on their primary structures and the folds of their catalytic domain as identified by structure determination (Eriani *et al.*, 1990; Cusack *et al.*, 1990). Crystal structures of most ARSs have been determined for a variety of substrate complexes, which have yielded detailed information regarding the substrate specificity, enzyme mechanisms and evolution of the two classes (Martinis *et al.*, 1999; Arnez & Moras, 1997). Prolyl-tRNA synthetases (PRSs) are class II ARSs that show the

highest structural resemblance to threonyl-tRNA, histidyl-tRNA and glycyl-tRNA synthetases, which are all members of subclass IIa (Sankaranarayanan *et al.*, 1999; Arnez *et al.*, 1995; Logan *et al.*, 1995). Although enzymes in this family have been extensively studied, only two structures of eukaryotic proteins have been reported to date: that of *Giardia lamblia* and that of human PRS domain complexed with HF and an ATP analogue (Zhou *et al.*, 2013; Larson *et al.*, 2012).

Human glutamyl-PRS (*h*EPRS), a bifunctional ARS part of an MSC, shows regulated noncanonical activity that blocks the synthesis of a specific protein (Sampath *et al.*, 2004). The enzyme glutamyl-prolyl tRNA synthetase (EPRS) was identified as a component of the interferon-γ (IFN-γ) activated inhibitor of translation (GAIT) complex by RNA affinity chromatography using the ceruloplasmin (Cp) GAIT element as a ligand (Ray *et al.*, 2007; Mukhopadhyay *et al.*, 2009). Thus, EPRS shows divergent functions in protein synthesis; in the MSC its aminoacylation activity supports global translation, whereas translocation of EPRS to an inflammation-responsive messenger ribonuclear protein (mRNP) causes gene-specific translational silencing (Sampath *et al.*, 2004). Recently, febrifugine and halofuginone (HF) have been shown to inhibit TH17-driven autoimmunity by activating the amino-acid starvation response pathway (Sundrud *et al.*, 2009). These molecules directly inhibited the PRS activity of EPRS and competed with proline for the PRS active site, resulting in the accumulation of uncharged tRNA and a state mimicking reduced cellular proline availability (Keller *et al.*, 2012). Here, we present the crystal structures of the PRS domain of human EPRS (*h*PRS): the apo form (*h*PRS-apo) at 2.4 Å resolution, *h*PRS complexed with ATP and its substrate proline at 2.3 Å resolution (*h*PRS-sub) and the enzyme complexed with HF at 2.62 Å resolution (*h*PRS-HF). These structures reveal how the *h*PRS protein binds prolyl-adenylate, proline and HF, and reveal structural information regarding the key position of the HF binding site.

2. Materials and methods

2.1. Cloning and protein expression

The cDNA encoding the C-terminal region of *h*PRS, including PRS from residues 1001 to 1512, was amplified by PCR using the primer pair forward, 5′-AAAAA-**GCTAGC**GGAGCAGGAGAAGGGCAGGGG-3′, and reverse, 5′-AA**CTCGAG**TCAGTAGCTGCGACCAAATAAGGTGTA-GTACTTG-3′. The amplified *h*PRS gene was cloned into the pET28a vector using DNA fragments that had been pre-digested with the restriction nucleases *Nhe*I and *Xho*I. The recombinant plasmid was transformed into *Escherichia coli* Rosetta 2 (DE3) pLysS cells, which were grown at 310 K in LB medium until late log phase (OD_{600} = 0.6–0.8). Protein expression was induced using isopropyl β-D-1-thiogalactopyranoside at a final concentration of 1 m*M* at 291 K for 18 h. The cells were harvested by centrifugation at 277 K at 4392*g* for 20 min.

2.2. Protein production

Harvested cells were resuspended in buffer *A* (20 m*M* Tris, 100 m*M* NaCl, 5 m*M* $MgCl_2$, 5 m*M* β-mercaptoethanol pH 8.0) and disrupted by sonication. Cell debris was pelleted by centrifugation (24 878*g* for 50 min) and the resulting supernatant was purified using Ni–NTA affinity chromatography. A linear imidazole concentration gradient was applied to elute the bound proteins. Fractions were analyzed using SDS–PAGE and those containing protein were pooled. The protein was further purified by ion-exchange chromatography using a linear NaCl gradient in buffer *B* [20 m*M* Tris, 5 m*M* $MgCl_2$, 5 m*M* dithiothreitol (DTT) pH 8.0]. The protein was concentrated using Amicon Ultra centrifugal filters (Milipore) as a final purification step. To ensure that all of the material was dimeric, a size-exclusion chromatography step was performed using a Superdex 200 26/60 column (GE Healthcare, Little Chalfont, England) equilibrated in the final buffer (20 m*M* Tris, 100 m*M* NaCl, 5 m*M* $MgCl_2$, 1% glycerol, 5 m*M* DTT pH 8.0). The protein was concentrated to 49 mg ml^{-1} for subsequent crystallization trials.

2.3. Crystallization

Initial crystal screening was performed by the sitting-drop vapour-diffusion method using commercial screening kits from Hampton Research in 96-well Intelli-Plates (Aliso Viejo, California, USA). Crystallization trays were prepared by mixing 0.5 µl of protein sample with an equal volume of screening solution. An initial crystallization hit was identified in a saturating solution of 0.1 *M* sodium cacodylate, 1.0 *M* sodium citrate dibasic trihydrate pH 6.5 at 293 K. Suitable crystals for X-ray diffraction were obtained by mixing 1 µl protein solution with 1 µl reservoir solution by the hanging-drop vapour-diffusion method. Cocrystallization was executed to obtain complex crystals containing substrates and the inhibitor HF. The substrate complex was obtained by incubation with 2 m*M* ATP and 5 m*M* L-proline. The inhibitor complex was obtained by incubating this solution with 2 m*M* HF at room temperature for 1 h before crystallization. Briefly, three types of crystals were obtained: *h*PRS-apo, *h*PRS-sub (the complex with ATP and L-proline) and *h*PRS-HF (the complex with HF). For cryoprotection, crystals were transferred to reservoir solution containing 15%(*v*/*v*) 3.4 *M* sodium malonate for several minutes before flash-cooling them in liquid nitrogen.

2.4. Data collection and structure determination

A data set (*h*PRS-apo) for single-wavelength anomalous diffraction (SAD) phasing was collected to 2.8 Å resolution using synchrotron radiation at a zinc peak wavelength of 1.2829 Å on beamline BL-1A at the Photon Factory, Tsukuba, Japan. The images were indexed and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). Using the zinc SAD data set, the initial phase of *h*PRS was obtained and improved using the *SOLVE*/*RESOLVE* module in *PHENIX* (Adams *et al.*, 2010), resulting in approximately 60% of the modelled residues being automatically built. The partial structure of *h*PRS-apo

was determined by molecular replacement using a PRS structure from *Thermus thermophilus* (PDB entry 1hc7; Yaremchuk *et al.*, 2001) as a search model with the *Phaser* v.2.5.1 module in *PHENIX*. Next, we combined the molecular-replacement phase with the Zn-SAD phase using *PHENIX*. Automated model building was carried out using *ARP*/*wARP* (Langer *et al.*, 2008). Further model building was performed using *Coot* (Emsley & Cowtan, 2004) and refinement was performed using *REFMAC*5 in *CCP*4 (Winn *et al.*, 2011) and *PHENIX* (Adams *et al.*, 2010). The final model of *h*PRS-apo was validated using *MolProbity* (Chen *et al.*, 2010) and had *R* values of $R_{\rm cryst}$ = 23.9% and $R_{\rm free}$ = 27.7% at 2.4 Å resolution. The complex structures of *h*PRS were determined by molecular replacement using the *h*PRS-apo structure and had *R* values of $R_{\rm cryst}$ = 23.0% and $R_{\rm free}$ = 27.1% for *h*PRS-sub at 2.3 Å resolution and $R_{\rm cryst}$ = 23.7% and $R_{\rm free}$ = 27.5% for *h*PRS-HF at 2.62 Å resolution. A summary of the statistics for the detailed final data collection and structure refinement is provided in Table 1.

2.5. Site-directed mutagenesis

A series of constructs for mutations in *h*PRS were produced using PCR. The QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, California, USA) was used to create the variants F1097A, F1097W, R1152K and R1152L. Expression and purification of the mutant variants were performed using the same protocol as used for wild-type *h*PRS.

2.6. Aminoacylation assay

His-tagged *h*PRS was expressed in *E. coli* strain Rosetta 2 (DE3) pLysS, harvested in phosphate-buffered saline (PBS), lysed with lysis buffer (20 m*M* Tris, 100 m*M* NaCl, 5 m*M* $MgCl_2$, 5 m*M* β-mercaptoethanol pH 8.0) and purified using the procedure described in §2.3. His-tagged *h*PRS was eluted using 200 m*M* imidazole pH 6.0 and dialyzed against PBS. The aminoacylation activity of *h*PRS was determined in reaction buffer {20 m*M* potassium phosphate pH 7.4, 6 m*M* magnesium acetate, 5 m*M* ATP, 400 µg ml^{-1} tRNA, 0.5 m*M* DTT, 20 Ci

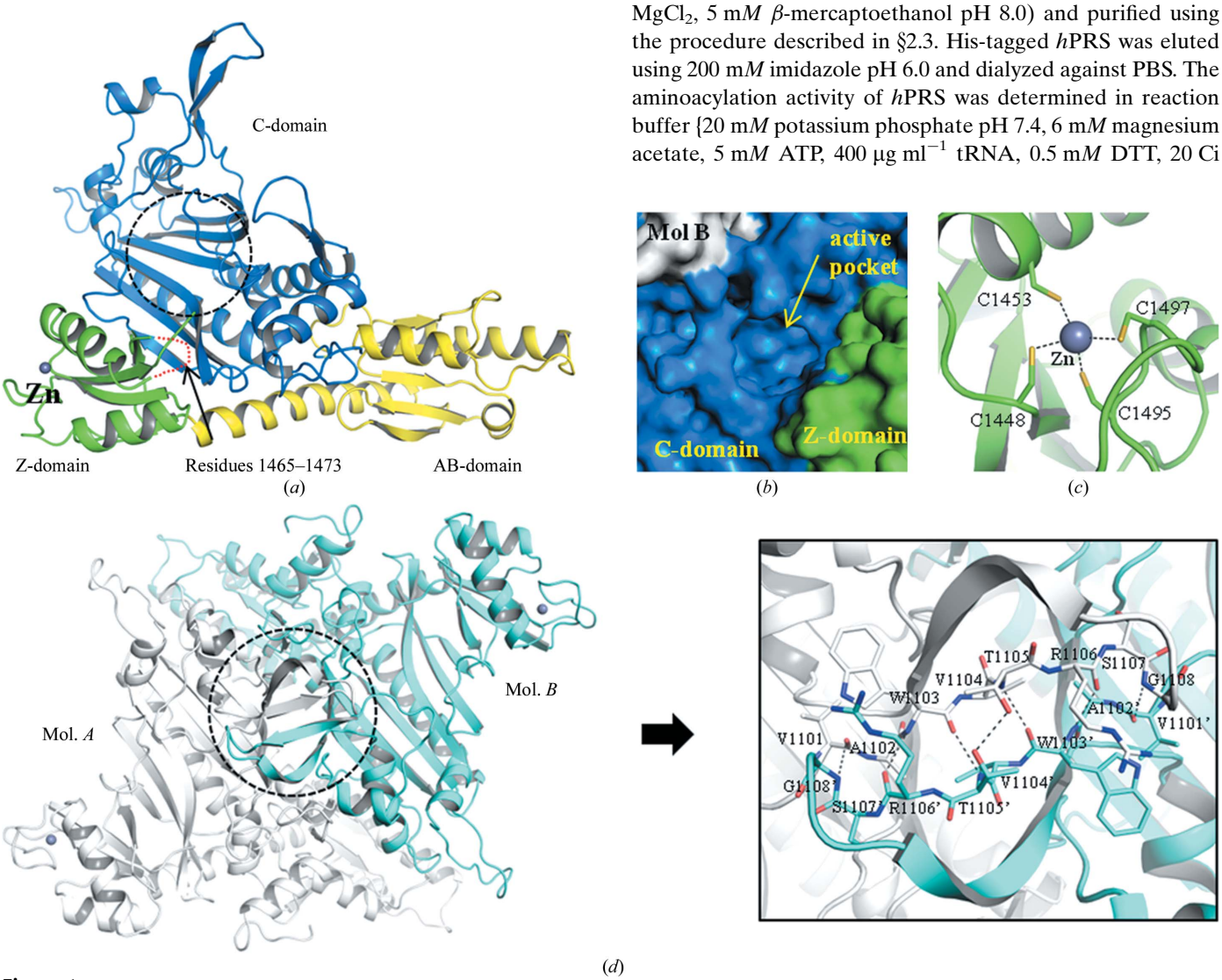

Figure 1
Overall structure of *h*PRS-apo. (*a*) *h*PRS-apo can be divided into three domains. The conserved class II catalytic core domain (C-domain) is shown in blue, the anticodon-binding domain (AB-domain) is shown in yellow and the C-terminal zinc-binding domain (Z-domain) is shown in green. The zinc ion (Zn) is represented as a silver sphere and the disordered region is indicated by a black arrow. The dashed circle indicates the active site. (*b*) The active site of *h*PRS is displayed on the surface model and the active pocket is indicated by an arrow. (*c*) The zinc ion is coordinated by four cysteines located in the Z-domain. (*d*) The dimer structure is represented by a cartoon model and the two molecules (Mol. *A* and Mol. *B*) forming a tight interaction are indicated by a dashed circle. Detailed interactions at the dimer interface are represented in the box on the right.

Table 1
Statistics of X-ray data collection and refinement.

Values in parentheses are for the highest resolution shell.

	*h*PRS-apo	*h*PRS-sub	*h*PRS-HF
Data collection			
X-ray source	Beamline 4A, PAL	Beamline 4A, PAL	Beamline 1A, Photon Factory
Wavelength (Å)	0.9793	0.9793	0.9800
Space group	$P3_121$	$P3_121$	$P3_121$
Unit-cell parameters			
$a = b$ (Å)	119.85	120.16	120.81
c (Å)	95.72	107.09	104.71
$\alpha = \beta$ (°)	90	90	90
γ (°)	120	120	120
Resolution (Å)	50–2.40 (2.45–2.40)	50–2.30 (2.34–2.30)	50–2.62 (2.67–2.62)
R_{merge} (%)	0.085 (0.398)	0.068 (0.417)	0.102 (0.440)
$\langle I/\sigma(I)\rangle$	26.7 (3.8)	38.7 (2.7)	18.9 (2.9)
Completeness (%)	99.6 (99.6)	96.2 (96.9)	96.9 (89.5)
Average multiplicity	11.5	7.3	5.8
Refinement			
Observed reflections	408419	281707	149486
Unique reflections	31308	38511	26051
R_{cryst}/R_{free} (%)	23.9/27.7	21.7/26.7	23.7/27.5
No. of atoms			
Total	3922	4021	4029
Protein	3914	3929	3929
Ligand/ion	1	28	25
Water	7	64	75
Average B factor (Å^2)			
Protein	65.095	71.262	59.894
Ligands	—	80.962	71.271
Water	57.868	60.469	50.934
R.m.s. deviations			
Bond lengths (Å)	1.031	1.382	1.168
Bond angles (°)	0.006	0.011	0.008
Ramachandran plot (*MolProbity*†)			
Favoured (%)	94.4	95.1	91.0
Allowed (%)	5.0	4.5	8.4
Outliers (%)	0.6	0.4	0.6
PDB code	4k86	4k87	4k88

† Chen *et al.* (2010).

(1 mCi ml^{-1}) [^{3}H]-proline} in the presence or absence of HF (100 n*M* or 1 µ*M*) at 310 K. The aminoacylation reaction was quenched on 3M filter paper pre-wetted with 5% trichloroacetic acid. After drying the filter paper, radioactivity was detected by a liquid-scintillation counter.

3. Results and discussion

3.1. Overall structure of the *h*PRS domain of human EPRS

The crystal model of *h*PRS-apo belonged to the trigonal space group $P3_121$, with unit-cell parameters $a = b = 119.85$, $c = 95.72$ Å. One *h*PRS molecule is included in the asymmetric unit, which includes residues 1016–1512 with the exception of residues 1465–1473 (Fig. 1*a*). The overall fold of *h*PRS shows a similar three-dimensional structure to those of its homologues from *T. thermophilus* (*Th*PRS) and *Giardia lambia* (*Gl*PRS) (Yaremchuk *et al.*, 2001; Larson *et al.*, 2012). The three domains are the catalytic domain (C-domain), the anticodon-binding domain (AB-domain) and the C-terminal zinc-binding domain (Z-domain). The C-domain contains residues 1016–1296 and forms a core consisting of six antiparallel β-strands surrounded by α-helices and three β-sheets. The active pocket consists of the C-domain, which is able to bind three ligands: proline, ATP and HF (Fig. 1*b*). The AB-domain consists of residues 1297–1423 and forms a central β-sheet and four helices, with the long C-terminal helix connected to the Z-domain by crossing one side of the C-domain. The Z-domain contains residues 1424–1512; it tightly interacts with the C-domain and coordinates a zinc ion through four cysteine residues: Cys1448 (2.5 Å), Cys1453 (2.5 Å), Cys1495 (2.4 Å) and Cys1497 (2.6 Å; Fig. 1*c*).

*h*PRS elutes as a homodimer in size-exclusion chromatography, which coincides with the functional form of other PRSs, which is also a dimer (Delarue & Moras, 1993). The dimeric structure is related by crystallographic symmetry (Fig. 1*d*). The dimer interface buries 2087.0 Å^2 of solvent-accessible surface from each molecule as calculated using *PISA* (Krissinel & Henrick, 2007), which is comparable to the interfaces observed for the homologues *Th*PRS and *Gl*PRS. The dimeric structure is highly conserved in class II tRNA synthetases. The interface is mainly composed of C-domains, which form salt bridges, hydrophobic patches and polar interactions. At the centre of the dimeric structure, a β-sheet from each C-domain has an antiparallel pattern and is located in the immediate region of the active site, possibly resulting in an allosteric conformational change (Figs. 1*b* and 1*d*). Between the two β-sheets, six main-chain hydrogen bonds are formed by residues 1101–1108. The side chains of Thr1105 and Thr1105* (where the asterisk indicates the dimerized form of molecule *B*) face each other and form a hydrogen bond.

The crystal structure of *h*PRS was inferred by using a combined approach of molecular replacement using *Th*PRS (PDB entry 1hc7) as a search model and Zn-MAD at 2.4 Å resolution. *h*PRS shares 41% identity with *Th*PRS and yielded a root-mean-square deviation (r.m.s.d.) of 1.01 Å for 356 C$^\alpha$-atom pairs; it shares 16% identity with *Gl*PRS and yielded an r.m.s.d. of 1.56 Å for 339 C$^\alpha$-atom pairs (Fig. 2). The overall structures of the PRS enzymes from the three organisms are similar and have conserved active pockets in the C-domain. The C-domain is highly conserved between species, whereas the AB-domain and Z-domain are less conserved. The zinc-binding site of the Z-domain is conserved between *Th*PRS and *h*PRS, whereas *Gl*PRS has no binding site for zinc. A disordered region from residues 1465 to 1473 is found in the Z-domain which is only present in *h*PRS and is absent in *Th*PRS and *Gl*PRS (Supplementary Fig. S1[1]). This region is located close to the tRNA-binding site based on the tRNA-complex structure of threonyl-tRNA synthetase from *E. coli* (PDB entry 1qf6; Sankaranarayanan *et al.*, 1999). This enzyme is the most similar class II aminoacyl-tRNA synthetase to *h*PRS (Supplementary Fig. S2). The nine flexible residues (1465–1473) may form an ordered structure by participating in tRNA binding.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: DW5057). Services for accessing this material are described at the back of the journal.

3.2. Proline- and adenosine-binding sites

To elucidate the exact mechanism of *h*PRS, we determined the complex structure of *h*PRS with substrates (the *h*PRS-sub structure) at a resolution of 2.3 Å. The complex crystal was obtained by cocrystallization with proline and ATP. The prolyl-adenylate is enzymatically formed in the active site. The electron density of proline was clearly defined, but that for the adenylate was weak because of incomplete reaction (Fig. 3*a*). Only the adenosine moiety of adenylate was observed at the active site, without the phosphate moiety. Proline is bound in the P-pocket, which is mainly formed by the conserved residues Thr1121, Glu1123, Arg1152, Glu1171, His1173, Thr1240 and His1242 and the hydrophobic residues Phe1097, Tyr1169 and Phe1216 (Fig. 3*a*). The N atom of the proline ring is positioned to form a hydrogen bond to the O^{γ} atom of Thr1121 or the O^{ε} atom of Glu1123, which is well matched in other complexes with prolyl-adenylate in other homologous structures (Larson *et al.*, 2012). The carboxyl group of proline is stabilized by hydrogen bonds to the $N^{\eta 2}$ atom of Arg1152 and the N^{ε} atom of His1242, residues that are strictly conserved in PRS homologues. The Trp1169 residue is located inside the P-pocket, and Phe1097 and Phe1216 bury the bound proline inside the pocket. The adenosine moiety is bound to the A-pocket, which is made up of Arg1152, Glu1154, Arg1163, Thr1164, Phe1167, Gln1237, Thr1240, Thr1276 and Arg1278 (Fig. 3*b*). Overall, the binding mode of the adenosine moiety is similar to that of prolyl-adenylate in homologous structures (Supplementary Fig. S3). The adenine moiety is located between Phe1167 and Arg1278. The N6 atom of the amino group forms a hydrogen bond with the carbonyl O atom of Thr1164. The ribose moiety is stabilized by hydrogen bonds contributed by several residues. The O2′ atom forms two hydrogen bonds with the N^{ε} atom of Arg1278 and the carbonyl O atom of Gln1237. A water molecule stabilized by Gly1238 and Thr1276 contributes a hydrogen bond to the O2′ atom of the ribose moiety. The O3′ atom of the ribose moiety is stabilized by hydrogen bonds contributed by the carbonyl O atom of Gly1274 and a water molecule. The $N^{\eta 1}$ atom of Arg1152 forms a hydrogen bond with the O5′ atom of the ribose moiety.

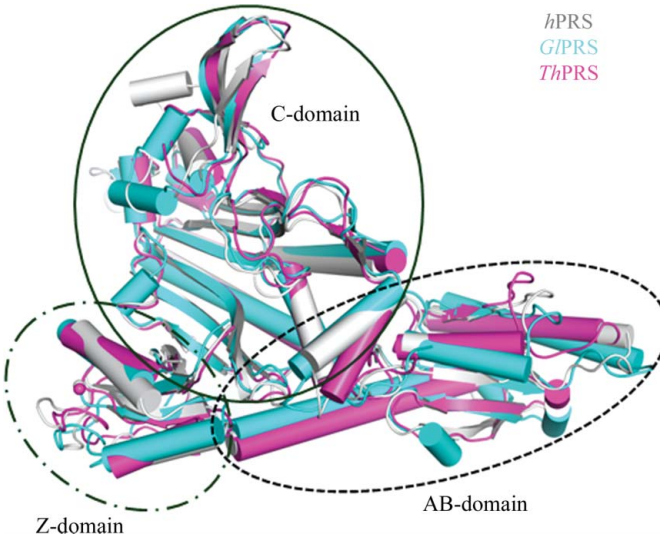

Figure 2
Superposition of overall PRS structures. Two PRS structures from *T. thermophilus* (*Th*PRS; pink) and *G. lamblia* (*Gl*PRS; cyan) are superimposed on *h*PRS (grey). The three corresponding domain regions are indicated by circles with the domain names.

3.3. Halofuginone-binding site

The crystal structure of *h*PRS was recently reported to contain HF and ATP in its active site (Zhou *et al.*, 2013). Most inhibitors of aminoacyl-tRNA synthetases mimic aminoacyl-

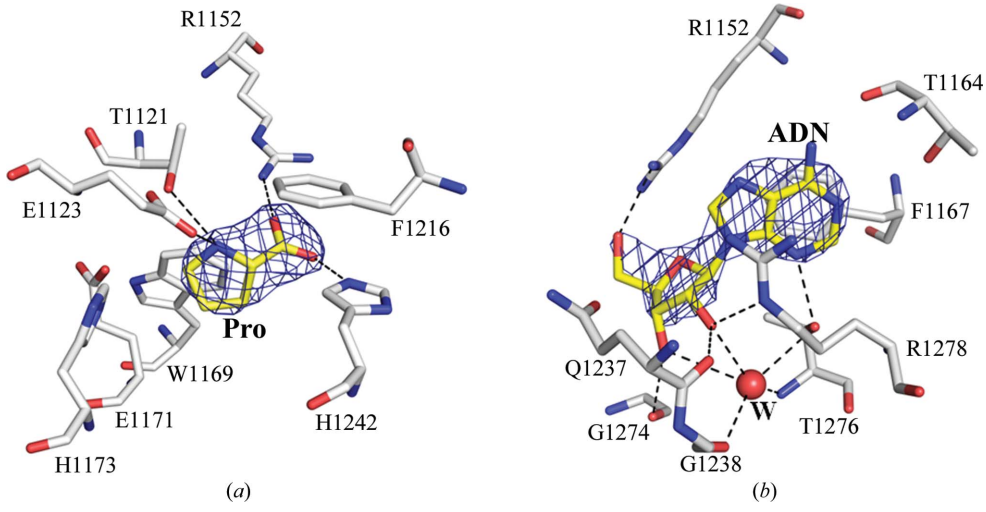

Figure 3
Stick model of proline and adenosine binding in the active site. (*a*) The active-site residues of the proline (Pro) binding site are represented by sticks, and hydrogen bonds between active-site residues and substrates are shown as dashed lines. (*b*) The adenosine (ADN) binding residues are represented by a stick model with bound adenosine. A water molecule (W) coordinated by Thr1276 and Gly1238 is represented as a red sphere. The $2F_o - F_c$ map contoured at 2.0σ shows the electron density for substrates in the active site.

adenylates and block the binding of the substrate ATP, whereas HF blocks the proline- and tRNA-binding sites. We determined the *h*PRS-HF structure at a resolution of 2.62 Å in the absence of ATP. ATP promotes the binding of HF to *h*PRS, although it is not necessary for the binding reaction (Keller *et al.*, 2012). The overall binding of HF is similar and it is bound at a dual site, including a piperidine ring at the P-pocket and a quinazolinone group at the H-pocket, and does not invade the ATP-binding site (A-pocket; Fig. 4*a*). The H-pocket, in which the halogenated 4-quinazolinone group is buried, primarily includes Leu1087, Phe1097, Glu1100, Val1101, Pro1120, Thr1121 and Arg1152 (Fig. 4*b*). The quinazolinone group is stabilized by a hydrophobic stack between Phe1097 and Arg1152, and a hydrogen bond is formed between the $N^{\eta1}$ atom of Arg1152 and the quinazolinone N atom. The $N^{\eta2}$ atom of Arg1152 forms a hydrogen bond with the keto group at the bridge between the quinazolinone group and the piperidine ring of HF. Two hydrogen bonds are contributed by two water molecules bound adjacent to HF and are stabilized by Arg1152 and Thr1240. The piperidine ring is bound at the same site as proline and structurally mimics proline binding. The N atom of the piperidine ring forms hydrogen bonds with Thr1121 and Glu1123, and the N^{ε} atom of His1242 forms a hydrogen bond with the O atom of the piperidine ring.

3.4. Conformational change upon ligand binding

The structures of *h*PRS-sub and *h*PRS-HF were superimposed onto the *h*PRS-apo structure, giving r.m.s.d.s of 0.36 for 396 C^{α}-atom pairs (*h*PRS-sub *versus* *h*PRS-apo) and 0.40 for 390 C^{α}-atom pairs (*h*PRS-HF *versus* *h*PRS-apo). A conformational change in the active pocket was detected upon ligand binding (Fig. 5*a*). Particularly, motif 1 (residues 1084–1100), which is made up of helices and found in class II tRNA synthetases, shows a dramatic change. The proline-binding loop (P-loop) contains residues 1210–1219; it showed specific changes upon proline binding. To distinguish between conformational changes for each ligand, the respective ligand-binding sites were compared with *h*PRS-apo. The residues involved in adenosine binding are almost identical to those in *h*PRS-apo, except for Arg1152, which moves towards the proline-binding pocket (Fig. 5*b*). Because of this movement, Arg1152 can participate in both adenosine and proline binding and can provide room for the adenine moiety to bind adenosine. Several conformational changes were observed upon proline binding (Fig. 5*c*). Motif 1 moves towards the proline-binding pocket and induces the Phe1097 residue to adopt a more helical structure. The Phe1216 residue in the P-loop flips inward to create a buried proline-binding pocket with Phe1097. Two residues, Thr1121 and Glu1123, in the TXE-loop (residues 1121–1123) shift to form hydrogen bonds with the N atom of the proline ring. These movements were also observed in the *Th*PRS homologue and the substrates are essential for organizing active-site residues that recognize proline through an induced-fit mechanism (Yaremchuk *et al.*, 2001). However, motif 1 of *Th*PRS is disordered in the apo form, whereas that of *h*PRS showed an ordered structure even in the absence of substrate. Phe1097 is located near the P-loop of *h*PRS-apo and moves towards the proline-binding pocket in *h*PRS-sub. Arg1152 is located at the centre of the active pocket and contributes to proline and adenosine binding. In the absence of substrates, this residue is located distant from the proline-binding pocket in *h*PRS, whereas that of *Th*PRS is positioned at the centre of the active pocket. *h*PRS-HF showed a conformational change upon HF binding and similar active-site interactions as in proline binding (Fig. 5*d*). The TXE-loop and Arg1152 show a similar position in

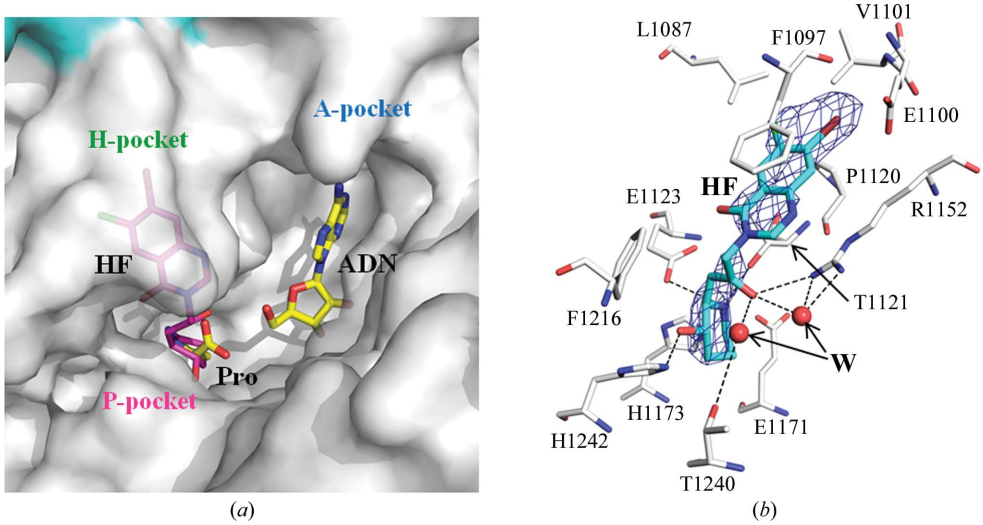

Figure 4
HF binding in the active site. (*a*) The HF binding site is indicated on the surface model and the three pockets are located at the active site. The HF molecule is bound in the H-pocket and P-pocket. The proline (Pro) and adenosine (ADN) are superimposed onto the HF-bound structure for comparison. (*b*) Active-site residues contributing to HF binding are represented by a stick model in the *h*PRS-HF structure. Hydrogen bonds between HF and active-site residues are shown as dashed lines. Two water molecules (W) coordinated by Thr1240 and Arg1152 are represented as red spheres. The $2F_o - F_c$ map contoured at 2.0σ shows the electron density for HF in the active site. A-pocket, ATP-binding pocket; H-pocket, quinazolinone moiety of HF-binding pocket; P-pocket, proline-binding pocket.

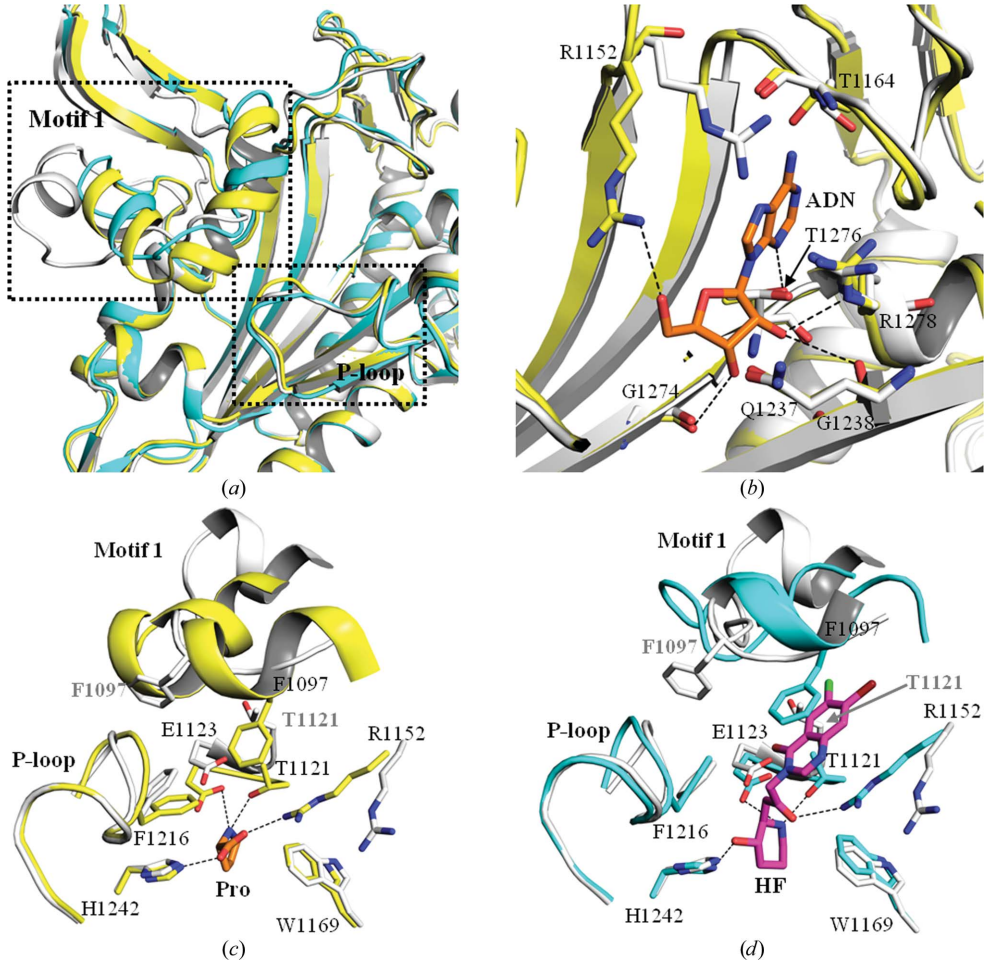

Figure 5
Ligand-binding site comparison of *h*PRS. (*a*) The monomer structures of *h*PRS-apo (grey), *h*PRS-sub (yellow) and *h*PRS-HF (cyan) are superposed. Motif 1 and the P-loop are indicated by dashed squares. *h*PRS-sub (yellow) is superposed onto the *h*PRS-apo (grey) structure. The adenosine-binding pocket is compared in (*b*). The proline-binding pocket is compared in (*c*) and consists of motif 1, the proline-binding loop (P-loop), the TXE-loop and several residues. Two residues of *h*PRS-apo are labelled in grey. (*d*) *h*PRS-HF (cyan) is superposed onto the *h*PRS-apo (grey) structure. The HF-binding pocket consists of motif 1, the proline-binding loop (P-loop), the TXE-loop and several residues. Two residues of *h*PRS-apo are labelled in grey.

*h*PRS-sub because HF uses the proline-binding pocket by binding the piperidine ring. Motif 1 showed a conformational change and contains the Phe1097 residue, which is crucial for stabilizing HF binding in the hydrophobic stack of the quinazolinone ring.

3.5. Active-pocket and dimer formation

The three structures of *h*PRS exhibited differences in their active pockets depending on the presence of bound ligands (Fig. 6). *h*PRS showed significant conformational changes upon ligand binding, inducing differential active-pocket formation. Motif 1 (residues 1084–1100) forms one side of the active pocket and is followed by a β-sheet (residues 1101–1109) which forms a dimer interaction with molecule *B* (Fig. 1*d*). Aside from the antiparallel pattern interaction between two β-sheets, the motif 1 region forms hydrogen bonds with residues 1107–1109 of molecule *B*. A loop region (residues 1152–1165) forming a different side of the active pocket locates Trp1153 at the dimer interface near Ala1102 of the β-sheet and Ser1107* of molecule *B*. In *h*PRS-apo, the carboxyl O atom of Val1101 forms a hydrogen bond with the backbone amide of Gly1108*, and Val1101 interacts with molecule *B* (Fig. 6, upper panel). The side chain of Trp1153 is rotated towards the dimer interface, whereas Arg1152 is located distant from motif 1 and the TXE-loop (residues 1121–1123). In the active pocket, the side chain of Arg1119 is disordered and shows no interaction with other residues. In *h*PRS-sub, the backbone amide of Gly1108* forms hydrogen bonds with the carbonyl O atoms of Pro1099 and Val1101 (Fig. 6, middle panel). Pro1099 moves towards the dimer interface and participates in the dimer interaction. The side chain of Trp1153 is rotated outwards from the interface and Arg1152 moves towards the same side as motif 1 to form a hydrogen bond with the carbonyl O atom of Pro1120. Pro1120 shows a conformational change with Arg1119 and the TXE-loop. Arg1119 moves between a long loop and β-strand, and several hydrogen bonds are formed using the side chain (Supplementary Fig. S4). In *h*PRS-HF, the Arg1119, Pro1120, Arg1152 and Trp1153 residues and the TXE-loop show similar

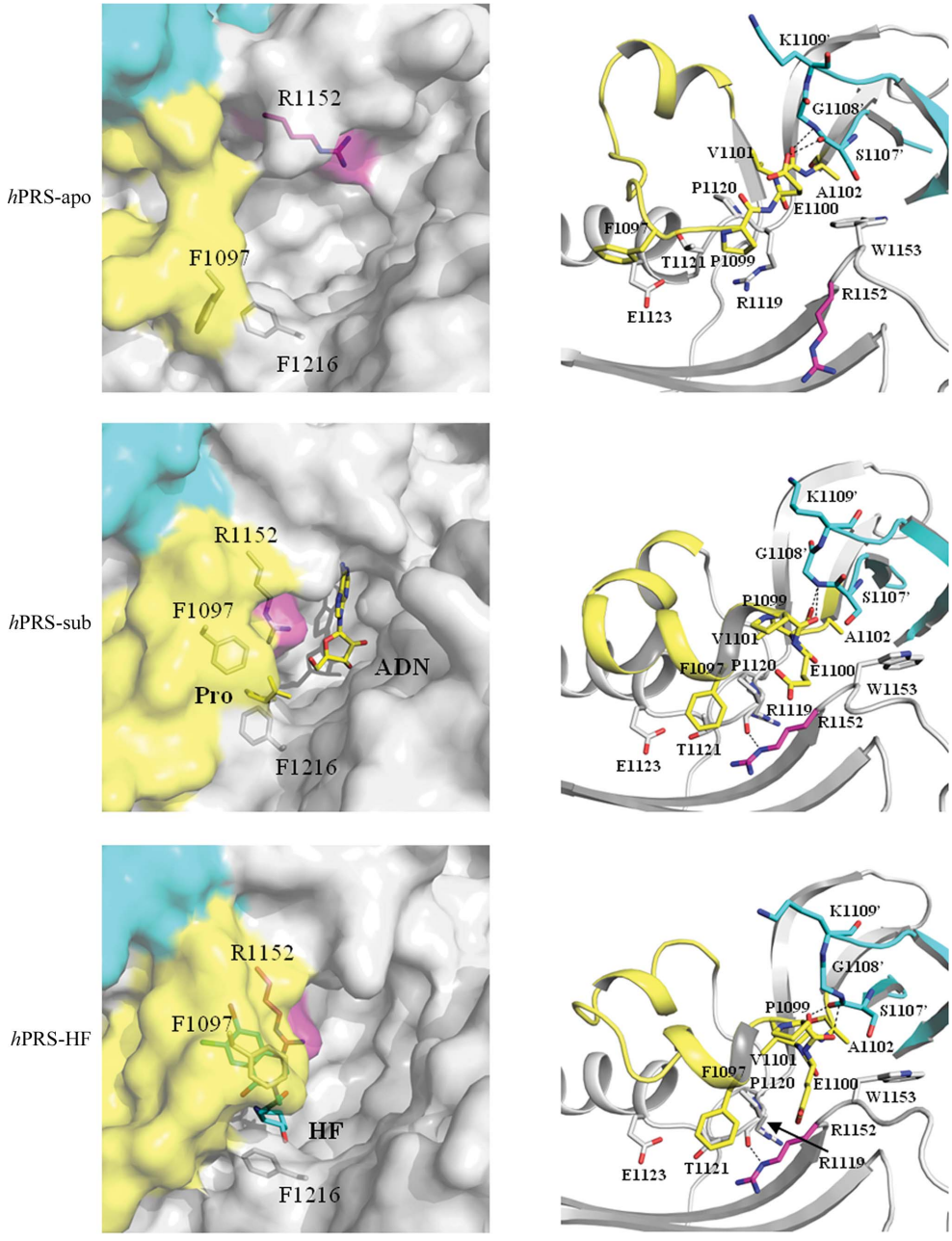

Figure 6
Active pocket and dimer interaction of *h*PRS-apo, *h*PRS-sub and *h*PRS-HF. The active pocket is represented by a surface model with motif 1 in yellow, molecule *B* in cyan and Arg1152 in magenta in the left panels. Three residues, Phe1097, Arg1152 and Phe1216, are displayed as stick models. The dimer interaction of *h*PRS near the active site is represented by a cartoon model in the right panels. Interactions with molecule *B* (cyan) and a specific interaction between Arg1152 and Pro1120 are indicated by dashed lines.

conformational changes to *h*PRS-sub (Fig. 6, lower panel). At the interface, the carbonyl O atom of Ser1107* and the backbone amide of Gly1108* form hydrogen bonds with the carbonyl O atom of Val1101 and the carbonyl O atom of Pro1099, respectively. Motif 1 moves slightly outward from the active pocket to provide space for the bound HF molecule. Overall, the active pocket is loosely organized in the absence of ligand, and ligand binding is mediated by an induced-fit mechanism. Motif 1 functions as a cap in *h*PRS; it is loosely opened in *h*PRS-apo, tightly closed in *h*PRS-sub and incorrectly closed in *h*PRS-HF. The active-pocket formation is closely linked to dimer interaction because changes in the interacting residues at the dimer interface accompany the conformational change of motif 1. The dimer structure may be important in active-pocket formation by maintaining the conformational changes depending on ligand binding.

3.6. Mutational analysis of HF-binding sites

To clarify the residues that are most important for HF binding, we selected Phe1097 and Arg1152 because they are involved in either hydrophobic and hydrophilic interactions

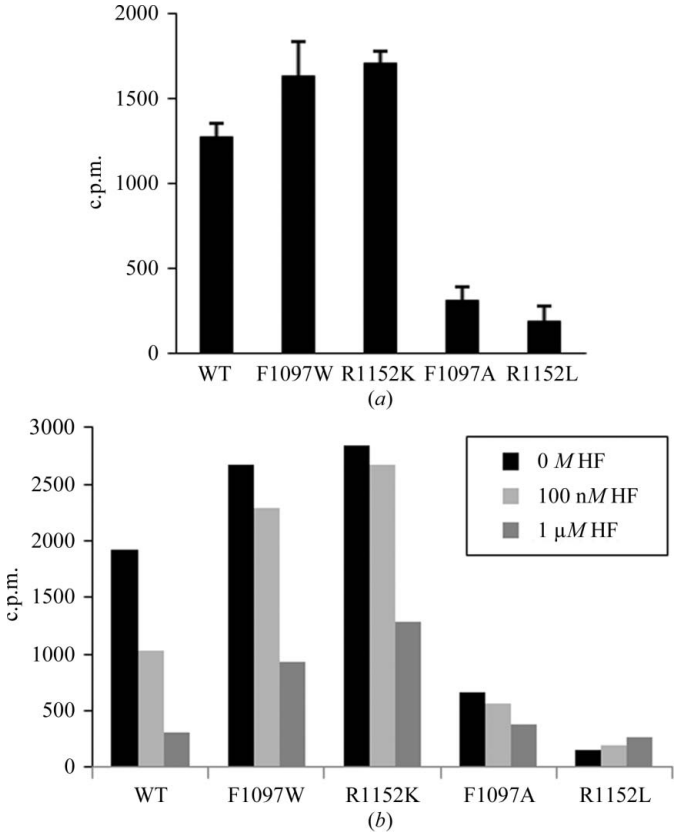

Figure 7
Mutant assays for *h*PRS. (*a*) Aminoacylation assay for *h*PRS. Error bars reflect triplicate determinations. (*b*) Aminoacylation assay in the presence of halofuginone (HF). The inhibitory activity of HF was explored at three different concentrations of HF: 0, 100 n*M* and 1 μ*M*. The assays were performed using wild-type (WT) *h*PRS and four variants: F1097W, R1152K, F1097A and R1152A.

with HF or in the interaction with adenosine and proline. We performed an aminoacylation assay with the four mutants F1097A, F1097W, R1152L and R1152K, which were generated to determine the roles of specific residues in *h*PRS function and HF binding. Although both the F1097W and the R1152K mutants exhibited elevated basal activity when compared with the wild type (WT), both the F1097A and the R1152L mutants showed no basal activity (Fig. 7*a*). As shown in Fig. 7(*b*), although the activities of the WT and the two mutants F1097W and R1152K decreased in an HF concentration-dependent manner, its inhibitory potency between the WT and mutants was completely different. It is hard to say whether the F1097A and R1152L mutants have sensitivity to HF because these mutants showed no basal activity. The activities of the F1097W and R1152K mutants were less inhibited than those of the WT in the presence of HF (100 n*M* or 1 μ*M*) because these mutants had a lower sensitivity to HF when compared with the WT. Taken together, substituting phenylalanine with tryptophan may interfere in the binding of incoming HF because of the bulky indole ring compared with the benzene ring. The inhibitory potency of HF for the mutants in which arginine was substituted with lysine was not high because the mutants no longer showed a bridge between ATP and HF.

3.7. The structural distortion of HF binding in the absence of ATP

Recently, independently of our studies, Zhou and co-workers published the structure of *h*PRS bound to HF and an ATP analogue (PDB entry 4hvc; Zhou *et al.*, 2013). While they showed a complex structure with both ATP and inhibitor in a molecule of *h*PRS, our structures represents apo, substrate-bound and inhibitor-bound forms. We focused on the active

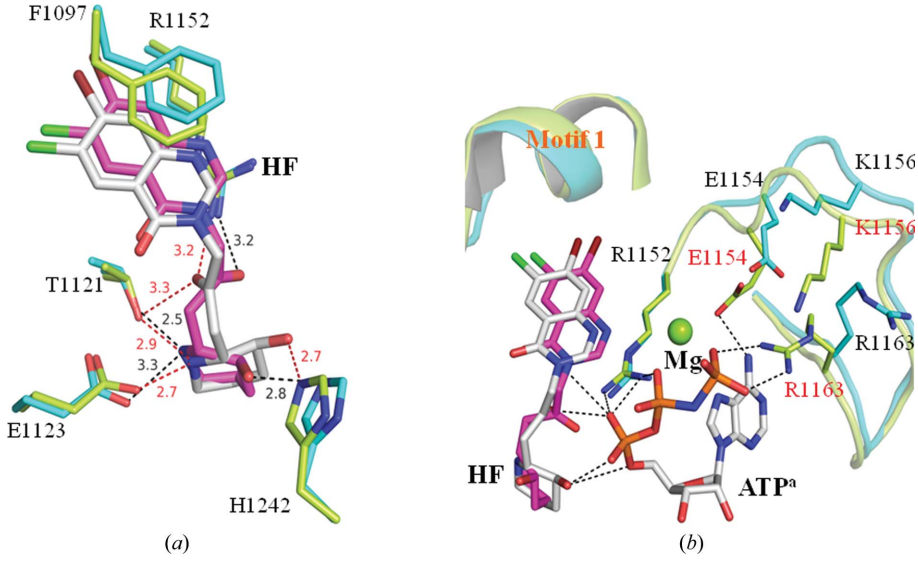

Figure 8
Structural comparison with PDB entry 4hvc. (*a*) The structure of *h*PRS-HF (cyan) is superposed on that of 4hvc (green). The bound molecules of *h*PRS-HF (magenta) and of 4hvc (grey) are shown. The interactions between active-site residues and HF molecules are indicated by black dashed lines in *h*PRS-HF and red dashed lines in 4hvc. The distances between the two interacting atoms are indicated near the dashed lines. (*b*) The ATP analogue (ATP^{a}) forms hydrogen bonds with motif 1 and a neighbouring HF molecule. Three residues of 4hvc are labelled in red.

pocket of the 4hvc structure and compared it with that of *h*PRS-HF. We found that some structural distortion of the HF piperidine ring was induced by the binding of the ATP analogue in the 4hvc structure (Fig. 8*a*). The O atom of the piperidine ring rotated towards ATP through an interaction with the phosphate moiety and the N atom and moved approximately ∼0.4 Å closer to Glu1123 and farther from Thr1121. The interaction with His1242 was maintained and showed a similar distance of 2.7–2.8 Å in the 4hvc structure and *h*PRS-HF. The quinazolinone group of HF shifted towards the H-pocket in the 4hvc structure and the stacking residue Phe1097 is shifted equally to the shift of HF. The Arg1152 residue maintains the same distances to the keto group and the N1 atom of the quinazolinone group, with distances of 3.2 and 3.3 Å, respectively. The keto group in the bridge interacts with the phosphate moiety of ATP, while that of *h*PRS-HF is placed near the ATP analogue, which induces steric collision (Fig. 8*b*). The loop including residues 1152–1165 contributes to ATP binding; it forms a more rigid structure because of the involvement of the Glu1154, Lys1156 and Arg1163 residues in phosphate-moiety binding. The magnesium and phosphate moieties are located at the boundary of three pockets and form hydrogen bonds with HF that maintain the tight binding of HF. Overall, our structures clearly supported the reasons why the active-site residues show tight interactions upon ATP binding in the 4hvc structure and therefore enhance HF binding.

We thank the staff at beamlines 4A at PAL, Republic of Korea and 1A at the Photon Factory, Japan (Proposal No. 2012G189) for technical support. We are grateful to the staff at the Korea Basic Science Institute (KBSI; Daejeon, Republic of Korea) for the use of the Mosquito crystallization robot and the Rigaku MicroMax-007 HF X-ray generator and to Dr Bong Yong Lee for the chemical compounds. KYH was partially supported by KBSI and KIST grants. This work was supported by the Global Frontier R&D Program funded by a National Research Foundation of Korea grant from the Korean Government (MSIP; NRF-2012054227).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Arnez, J. G., Harris, D. C., Mitschler, A., Rees, B., Francklyn, C. S. & Moras, D. (1995). *EMBO J.* **14**, 4143–4155.
Arnez, J. G. & Moras, D. (1997). *Trends Biochem. Sci.* **22**, 211–216.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. & Leberman, R. (1990). *Nature (London)*, **347**, 249–255.
Delarue, M. & Moras, D. (1993). *Bioessays*, **15**, 675–687.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). *Nature (London)*, **347**, 203–206.
Favorova, O. O. (1984). *Mol. Biol. (Mosk.)*, **18**, 205–226.
Keller, T. L., Zocco, D., Sundrud, M. S., Hendrick, M., Edenius, M., Yum, J., Kim, Y.-J., Lee, H.-K., Cortese, J. F., Wirth, D. F., Dignam, J. D., Rao, A., Yeo, C.-Y., Mazitschek, R. & Whitman, M. (2012). *Nature Chem. Biol.* **8**, 311–317.
Kim, S., You, S. & Hwang, D. (2011). *Nature Rev. Cancer*, **11**, 708–718.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Larson, E. T., Kim, J. E., Napuli, A. J., Verlinde, C. L. M. J., Fan, E., Zucker, F. H., Van Voorhis, W. C., Buckner, F. S., Hol, W. G. J. & Merritt, E. A. (2012). *Acta Cryst.* D**68**, 1194–1200.
Logan, D. T., Mazauric, M. H., Kern, D. & Moras, D. (1995). *EMBO J.* **14**, 4156–4167.
Martinis, S. A., Plateau, P., Cavarelli, J. & Florentz, C. (1999). *Biochimie*, **81**, 683–700.
Mukhopadhyay, R., Jia, J., Arif, A., Ray, P. S. & Fox, P. L. (2009). *Trends Biochem. Sci.* **34**, 324–331.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Park, S. G., Ewalt, K. L. & Kim, S. (2005). *Trends Biochem. Sci.* **30**, 569–574.
Park, S. G., Schimmel, P. & Kim, S. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 11043–11049.
Ray, P. S., Arif, A. & Fox, P. L. (2007). *Trends Biochem. Sci.* **32**, 158–164.
Sampath, P., Mazumder, B., Seshadri, V., Gerber, C. A., Chavatte, L., Kinter, M., Ting, S. M., Dignam, J. D., Kim, S., Driscoll, D. M. & Fox, P. L. (2004). *Cell*, **119**, 195–208.
Sankaranarayanan, R., Dock-Bregeon, A.-C., Romby, P., Caillet, J., Springer, M., Rees, B., Ehresmann, C., Ehresmann, B. & Moras, D. (1999). *Cell*, **97**, 371–381.
Schimmel, P. (1987). *Annu. Rev. Biochem.* **56**, 125–158.
Sundrud, M. S., Koralov, S. B., Feuerer, M., Calado, D. P., Kozhaya, A. E., Rhule-Smith, A., Lefebvre, R. E., Unutmaz, D., Mazitschek, R., Waldner, H., Whitman, M., Keller, T. & Rao, A. (2009). *Science*, **324**, 1334–1338.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yaremchuk, A., Tukalo, M., Grøtli, M. & Cusack, S. (2001). *J. Mol. Biol.* **309**, 989–1002.
Zhou, H., Sun, L., Yang, X.-L. & Schimmel, P. (2013). *Nature (London)*, **494**, 121–124.