The PRT protein family
Sangita C Sinha* and Janet L Smith†

Members of the homologous PRT family are catalytic and regulatory proteins involved in nucleotide synthesis and salvage. New crystal structures have revealed key elements of PRT protein function, as well as glimpses of how the fold has evolved to perform both catalytic and regulatory functions.

Addresses
*Department of Biochemistry, Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA; e-mail: sangita@chop.swmed.edu
†Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA; e-mail: smithj@purdue.edu
Correspondence: Janet L Smith

Current Opinion in Structural Biology 2001, 11:733–739

0959-440X/01/$ – see front matter © 2001 Elsevier Science Ltd. All rights reserved.

Abbreviations
GPATase glutamine PRPP amidotransferase
PDB Protein Data Bank
PPi pyrophosphate
PRPP 5-phosphoribosyl-α1-pyrophosphate
PRTase phosphoribosyltransferase

Introduction

PRT proteins comprise a homologous family of enzymes and regulatory proteins of the nucleotide synthesis and salvage pathways. Members of the family are defined by the protein fold and by a short sequence motif. The name PRT comes from the phosphoribosyltransferase (PRTase) enzymes, which carry out phosphoryl transfer reactions on 5-phosphoribosyl-α1-pyrophosphate (PRPP), an activated form of ribose-5-phosphate. PRPP is a central metabolite connecting the nucleotide synthesis and salvage pathways. A broad understanding of the structure and function of the PRT family comes from the analysis of nearly 40 PRT crystal structures of 16 different PRT proteins, including representatives from all known PRT subfamilies [1–22,23,24,25,26].

Enzymes catalyzing PRTase reactions are the most numerous members of the PRT family. PRTases catalyze the displacement of pyrophosphate (PPi) from PRPP by a nitrogen-containing nucleophile, such as ammonia, adenine, guanine, hypoxanthine, xanthine, orotate or uracil, producing a β-substituted ribose-5-phosphate and PPi. Not all PRT proteins are enzymes. For example, in some bacteria, PRT proteins regulate the expression of purine and pyrimidine synthetic genes. Conversely, not all enzymes with PRTase activity are members of the PRT family. The PRT sequence motif is only found in PRTases from the nucleotide synthesis and salvage pathways. Other PRTases, from the tryptophan, histidine and nicotinamide synthetic and salvage pathways, lack the PRT sequence motif. One of these enzymes, quinolinate PRTase, was classified as a ‘type II PRTase’ because its structure is unrelated to the PRT family (‘type I PRTases’) [27,28]. The sequences of the other tryptophan, histidine and nicotinamide enzymes appear to be unrelated to each other and unrelated to the PRT family (see also Update). Here we review only members of the homologous PRT family.

Like many protein families, the PRT family was identified at the sequence level and the family relationship was confirmed later by three-dimensional structures. Nearly all PRT proteins include a 13-residue sequence motif easily recognized in database searches. The motif consists of four hydrophobic amino acids, two acidic amino acids and seven amino acids of variable character, usually including glycine and threonine. The motif was correctly predicted to be a PRPP-binding site in advance of structural information [29,30]. Apart from the PRT sequence motif, different PRT proteins have a low level of sequence identity, generally <15%.

The PRT fold binds PRPP

PRT proteins have a common core consisting of a five-stranded, parallel β sheet (β2–β1–β3–β4–β5) sandwiched between layers of α helices, α1 and α2 on one face of the β sheet and α3 on the other face (Figure 1). The core fold is expanded by between two and five additional secondary structures, which vary among PRT family members. At the C-terminal edge of the core β sheet, a second domain or subdomain, known as the ‘hood’, is formed predominantly by residues at the N terminus of the protein.

PRT proteins have a conserved PRPP-binding site, as expected for a homologous protein family with a common ligand. PRPP binds in association with a divalent metal ion, usually Mg2+. The PRPP-binding site comprises three loops: a ‘PPi loop’ between β1 and α2; a ‘flexible loop’ between β2 and β3; and a ‘PRPP loop’ between β3 and α3 (Figure 1). PRPP binds to the C-terminal edge of the central β sheet between the PPi and PRPP loops. The sequences of the three loops of the PRPP-binding site are highly conserved for each PRT protein, although there are no sites of strict conservation across the PRT family.

The PRPP loop constitutes the last nine residues of the 13-residue PRT sequence motif and binds the ribose 5-phosphate group of PRPP (Figure 2). The length, sequence and conformation of the PRPP loop are highly conserved in PRPP-binding sites throughout the PRT family. Two acidic residues at the start of the loop are conserved in the PRPP-binding sites of all PRTs, except uracil PRTase, which has only one acidic residue. These characteristic acidic sidechains form hydrogen bonds with the ribose hydroxyl groups. The ribose 5-phosphate group binds in a positively charged pocket that is formed by backbone amides, polar sidechains and the dipole of helix α3, and
includes at least one glycine and one threonine. This interaction is apparently essential to ligand binding. The ribose ring is disordered in nucleoside complexes that lack 5-phosphate [15,24]. Anions from the crystallization solution, such as sulfate, phosphate or sulfonate, nearly always fill the 5-phosphate site of PRT structures lacking specific ligands [5,11–13,15,19,23,26].

The four-residue PPi loop binds the PPi group of PRPP. The conformation of the PPi loop is conserved across the PRT family, but sequence similarity is restricted to the fourth residue (generally glycine, alanine or serine), which contacts the adjacent PRPP loop. A key element of Mg•PRPP recognition is a rare, nonproline cis peptide between the first and second residues of the PPi loop (Figure 2). The N–H and C=O groups of the cis peptide point into the PRPP-binding site and form hydrogen bonds with the metal•PPi group of bound PRPP. The cis peptide may occur only in PRT proteins for which PRPP recognition is essential to in vivo function. For example, the cis peptide is not present in PRPP synthetase [23], for which PRPP is the product of the reaction, or in the regulatory protein PyrR [12], whose effector is UMP. Anions from the crystallization solution bind in the β-phosphate position in a few PRT crystal structures lacking PRPP or PPi [8,11,19,26].

The flexible loop closes over bound Mg•PRPP in the PRT enzymes (Figure 3a,b). This long crossover connection between strands β2 and β3 has strikingly different lengths in different PRT proteins. The flexible loop is ordered and closed only in PRPP-bound forms of PRT proteins [10,16–18,20,25] and, generally, is unfolded or highly mobile in the absence of PRPP (Figure 3a,b). Closed
flexible loops contact PRPP, usually involving a few basic sidechains. The sequences of flexible loops are highly conserved for each PRT family member but, similar to the PPi loop sequences, are poorly conserved among different PRT proteins. The closed state of the flexible loop appears to be transient, as not all structures containing appropriate ligands have closed flexible loops [4,14,22,24].

**Metal binding**

Mg2+ is associated with the PPi of PRPP and appears to be essential to PPi binding in the PRPP site. In all closed-loop PRT structures, Mg2+ in a primary site is coordinated by a β-phosphate oxygen, by both PRPP ribose hydroxyls, by the ribose–PPi bridging oxygen and by two water molecules. One of the water molecules is hydrogen bonded to the C=O of the cis peptide in the PPi loop (Figure 2). Remarkably, the Mg2+ ion in this site has no direct protein ligands.

In some open-loop PRT structures, Mg2+ occupies an alternative position within the primary site. The main and alternative positions are closer than 2 Å and thus cannot be occupied simultaneously. In the alternative position, Mg2+ is coordinated by the two ribose hydroxyls and by the two conserved carboxyl sidechains in the PRPP loop, with polar sidechains or water molecules completing the octahedral coordination sphere. Nucleotide feedback inhibitor complexes of glutamine PRPP amidotransferase (GPATase) bind Mg2+ in the alternative position [9,10] (Figure 3a). Mg2+ occupies a similar position in the ternary complex of a guanine PRTase with a nucleotide product analog and PPi [24], and in a xanthine PRTase structure with sulfate in the β-phosphate site [8]. These structures may represent the pre-catalytic or post-catalytic arrangement of substrates [24].

Purine PRTases might all have a second metal site, fully compatible with the primary site. In the second site, Mg2+ is coordinated by α and β PPi oxygens, three water molecules and a conserved aspartate sidechain from the hood subdomain ([16,18,20,22,24,25••]; Figure 2). The second Mg2+ is linked to the purine substrate through a water ligand and might orient the purine base for catalysis. The hydrogen bond between water and purine N3 is independent of the protonation state of N3. Thus, the second Mg2+ also contributes to purine binding and recognition [31].

**PRT quaternary structure**

No monomeric PRT proteins are known. In most cases, the PRT core fold is part of a dimer interface in which a basic residue from each monomer contacts Mg•PRPP in the neighboring subunit. Regions of the structure involved in cross-subunit Mg•PRPP interactions differ among PRT proteins and include the flexible loop, the hood, the PPi loop and the bottom of the core domain. Despite these cross-subunit interactions, most PRT proteins do not display cooperative Mg•PRPP binding.

For a few PRT proteins, allosterity is important to function. GPATase catalyzes the first step of purine biosynthesis and is feedback-inhibited by the end products of the pathway. Two different feedback inhibitor-binding sites overlap the Mg•PRPP site [2,9–11]. PRPP synthetase connects the nucleotide synthesis and salvage pathways, and is activated by inorganic phosphate (P_i) and inhibited by ADP. These effector molecules bind to overlapping sites between subunits of the hexameric enzyme, remote from the active site [23•]. Uracil PRTase, which is part of the cellular scheme for balancing uracil and cytidine nucleotides, is regulated by GTP [32], although the structural basis of this allosteric behavior is unknown. Allosterity is expected to be an element of function for the PRT regulatory proteins PurR and PyrR, which bind to nucleic acid in response to a small-molecule effector.

**The amazing, multifunctional flexible loop**

The PRT flexible loop is disordered in most crystal structures, confounding structural biologists who seek an intact PRT active center. In PRTases, the loop orders only when Mg•PRPP binds, completing the active site and...
sequestering the reaction from bulk solvent. Nature has also adapted the large conformational changes of the flexible loop to a variety of functions associated with Mg•PRPP binding, including catalysis, signaling and cooperativity.

The flexible loops of several PRT proteins occur in subunit interfaces and contribute to Mg•PRPP binding, including catalysis, signaling and cooperativity. For example, arginine sidechains interact with PP, in the active sites of adjacent subunits of both GPATase [9,10] and orotate PRTase [4]. Similar functions are inferred for the flexible loops of adenine PRTase [19] and PurR [26•], in which an arginine sidechain from the flexible loop is hydrogen bonded to an anion in the β-phosphate site of an adjacent subunit. In orotate PRTase, a lysine sidechain at the tip of the flexible loop is essential to catalysis in the neighboring subunit of the dimeric protein [33,34], raising the possibility that the flexible loop may close over Mg•PRPP in the neighboring subunit.

The GPATase flexible loop is remarkable in its multiple functions. When it closes over Mg•PRPP, the flexible loop signals PRPP binding by contacting a second catalytic domain and activating production of the nucleophile NH₃ by glutamine hydrolysis. The closed flexible loop also

---

Figure 3

Variable conformations of the PRTase flexible loop. The proteins are oriented identically, with flexible loops highlighted in blue. (a) Group 1: the E. coli GPATase open form (left) with Mg•GMP bound (PDB code 1ECB) and closed form (right) with Mg•cPRPP bound (PDB code 1ECC) [10]. (b) Group 2: T. cruzi hypoxanthine/guanine PRTase, open form (left), with the purine nucleoside formycin B and the buffer N-morpholinoethanesulfonic acid bound (PDB code 1TC1) [15] and closed form (right) with Mg•PRPP and 3H-pyrazolo[4,3-d]pyrimidin-7-ol, a hypoxanthine analog, bound (PDB code 1TC2) [16]. (c) Group 3: open form of S. typhimurium orotate PRTase with Mg•PRPP and orotate bound (PDB code 1OPR) [4]. (d) Group 4: open form of B. subtilis PurR with bound sulphate [26•].
forms a tunnel for NH$_3$ to travel between the glutaminase and PRTase active sites. Finally, the flexible loop contacts feedback inhibitors in two subunits, contributing to the cooperative allosteric transition [2,9,10,35–38].

The PRTs can be classified into four groups based on flexible loop conformations (Figure 3). All of the flexible loops form additional β strands, which protrude from or add to the PRT core.

Group I includes the GPATases. The flexible loop consists of two short antiparallel β strands (Figure 3a). The first β strand is a continuation of β2 in the PRT core [2,9,11]. The long, 20-residue connection between these β strands is disordered in open-loop structures and undergoes a coil→helix transition as it folds and closes over bound Mg•PRPP [10].

The open flexible loops of hypoxanthine/guanine/xyanthine) PRTase [3,6,7,15,21,22], xanthine PRTase [8,14], guanine PRTase [24] and uracil PRTase [13], which comprise group 2, form a β hairpin that is nearly perpendicular to the core β sheet (Figure 3b). The hairpin loop, of ten or fewer residues, is disordered in open-loop structures. In closed-loop structures, the β hairpin is closer to the PRT core and the connecting loop is ordered and kinked over bound Mg•PRPP [16–18,20,25*].

The flexible loops of group 3 members, including orotate PRTase [1,4,5], PyrR [12] and PRPP synthetase [23*], contribute an additional antiparallel β strand to the PRT core (Figure 3c). The 10-residue loop connecting the antiparallel β strand to β2 is disordered in open-loop structures of group 3 proteins.

The open flexible loops of PurR [26*] and adenine PRTase [19], comprising group 4, include a two-stranded, antiparallel β ribbon (Figure 3d). The least ordered part of the flexible loop is the distant end of the β ribbon and, in PurR, the connection between β2 and the β ribbon. There are no structures representing closed forms of proteins from groups 3 and 4, but the flexible loops in neither group are long enough to refold in a helical conformation.

**The PRT hood**

Above the core fold, PRT proteins have a structure known as the hood. The hood includes peptides both N- and C-terminal to the core, and can be as small as a few loops, as in PurR and xanthine PRTase, or as large as a full domain, as in GPATase and PRPP synthetase.

The structure of the hood is highly variable across the PRT family. In PRT enzymes, the hood binds or generates the nucleophilic substrate. In purine or pyrimidine PRTases and in PyrR, the hood recognizes the appropriate base through specific hydrogen bonds and ligates the second Mg$^{2+}$ in purine PRTases. In GPATase, the hood is a catalytic domain that generates the nucleophilic substrate NH$_3$ by glutamine hydrolysis [39]. In PRPP synthetase, the hood produces PP$_i$ from ATP [23*]. No base is known to bind to the hood of PurR, perhaps through the occlusion of potential binding space by two aromatic residues [26*].

**Catalysis by PRTases**

The available structures provide a wealth of information on states of PRTases relevant to catalysis, effectively providing snapshots along the reaction pathway [31,40]. The closed-loop structures, in particular, provide a remarkably consistent picture of Mg$^{2+}$PRPP binding. Mg$^{2+}$PRPP binds in a virtually identical manner in these structures, despite the absence of invariant residues among the enzyme members of the PRT family. The structural snapshots are consistent with a substrate-assisted catalytic mechanism, in which the enzyme provides a solvent-excluded, suitably charged environment and binds Mg$^{2+}$PRPP in a conformation that promotes transition-state formation. A number of details of Mg$^{2+}$PRPP binding support this hypothesis. The PP$_i$ tail is oriented to accept a proton from the incoming nucleophile and to stabilize the planar (trivalent or pentavalent) transition state through a CH–O interaction between the ribose C1 proton and an α-phosphate oxygen [10]. Mg$^{2+}$ ligation by all three exocyclic oxygens of the ribose ring provides electrostatic assistance to catalysis and influences the pucker of the ribose ring. The ribose pucker, which should assist in the formation of a partial double bond at O4–C1 in the transition state, differs among closed-loop structures, but all are within the C3-endo/C2-endo pseudorotation domain. Puckering at C5 would be optimal for double bond formation at O4–C1. The high-resolution (1.05 Å) structure of hypoxanthine/guanine PRTase has a C2-endo ribose pucker in the ternary complex with Mg$^{2+}$PRPP and 9-deazaguanine [25*].

The structural snapshots also help answer the question of whether catalysis proceeds by a more dissociative (S$_N$1) or associative (S$_N$2) mechanism. The positions of 5-phosphate, PP$_i$ and Mg$^{2+}$ are fixed in the active site. The ribose position varies, however, depending on whether Mg$^{2+}$PP$_i$ is present and whether ribose C1 is bonded to PP$_i$ or to the nucleophilic base. In this respect, the PRTase structures, taken as a group, are consistent with a dissociative (S$_N$1) mechanism.

**Regulatory PRT proteins**

Two members of the PRT family from *Bacillus subtilis* regulate the expression of purine and pyrimidine biosynthetic genes. In these proteins, the ability of a PRTase to bind its substrate PRPP or its product nucleotide has been co-opted for a regulatory function.

PurR represses transcription initiation of genes of the *de novo* purine pathway by binding to specific operator DNA sequences. PRPP induces the expression of the purine genes by reducing the affinity of PurR for operator DNA [41]. PRPP is presumed to bind to the PRT core in
the same manner as it binds other PRT proteins. The PRT domain of PurR is augmented by a helix-turn-helix domain to bind DNA, thus exploiting the PRPP-binding capability to control transcription initiation [26*].

PyrR regulates the expression of pyrimidine genes by attenuation. PyrR binds a specific stem-loop structure in the regulatory region of RNA transcripts of pyrimidine genes, allowing formation of a downstream terminator [42]. UMP is a co-attenuator, enhancing the RNA affinity of PyrR, and is presumed to bind in the PRPP loop of the PRT core. A basic surface of the PyrR dimer is the proposed binding site for the RNA stem–loop [12]. PyrR also has uracil PRTase activity, although catalytic activity is unlikely to be relevant to its regulatory function. In PyrR, the PRT fold is adapted to bind RNA, thus exploiting the binding of the PRTase product to attenuate gene expression.

Sequence homologs of PurR and PyrR are not widely distributed. They are found in a few bacteria, primarily close relatives of B. subtilis. On the other hand, PRT enzymes are ubiquitous. We therefore presume that the PRT regulatory proteins arose from PRT enzymes by relatively recent gene duplications. The closest structural relatives of PurR among the PRTases is adenine PRTase. Despite its uracil PRTase catalytic activity, the closest relative of PyrR is hypoxanthine/guanine PRTase.

A distant relative—PRPP synthetase

N-terminal domain

So far, we have discussed PRT proteins that bind Mg•PRPP and are recognized at the sequence level through the PRT motif, which includes the PRPP loop. However, distant relatives in many protein families are not recognizable at the sequence level and are discovered only through three-dimensional structures. To date, only one such distant relative has been identified in the PRT family. PRPP synthetase catalyzes the formation of PRPP from ribose-5-phosphate and ATP. The C-terminal domain of PRPP synthetase includes the PRT sequence motif and, as expected, has the PRT fold, which binds the substrate, ribose-5-phosphate, and the product, PRPP [23*].

Unexpectedly, the N-terminal domain of PRPP synthetase also belongs to the PRT family. The N-terminal domain binds the substrate, ATP, and the allosteric effectors, ADP and P_i, in unique sites that are not topological analogs of the Mg•PRPP site. The topological analog of the PRT flexible loop in the N-terminal domain is found in a subunit interface of the hexameric protein and participates in base-specific recognition of ATP in another subunit. The two PRT domains of PRPP synthetase face one another as if each were the other’s ‘hood’ domain. The N-terminal domain is presumed to have arisen by duplication of the gene encoding the C-terminal domain. Other distant relatives of the PRT family remain to be identified.

Conclusions

Three-dimensional structures have fleshed out hints about a common structure and function among the homologous family of PRT proteins. Part of the characteristic 13-residue sequence motif forms the PRPP loop for binding Mg•PRPP, assisted by the PP_i loop and the flexible loop. The flexible loop has an essential role in the specific biological function of each PRT, such as catalysis, allostery, interdomain signaling and channel formation. The mode of Mg•PRPP binding is virtually identical among PRTases, strongly suggestive of a common catalytic mechanism. The lack of invariant amino acids in the active site leads to obvious proposals for substrate-assisted catalysis. PRTases of only the nucleotide synthesis and salvage pathways possess the characteristic sequence motif; other proteins with PRTase activity are unlikely to belong to the homologous PRT family. The mechanism of regulatory PRT proteins that bind DNA or RNA remains to be determined. Only one distant family member has been identified.

Update

The structure of the tryptophan biosynthetic enzyme anthranilate phosphoribosyltransferase was solved recently (O Mayans, A Ivens, K Kirschner, M Wilmanns, personal communication). As expected, this protein fits the pattern of other PRTases: it does not catalyse a reaction of nucleotide synthesis or salvage, it does not possess the PRT sequence motif, it has a fold different from the PRT proteins discussed in this review and it is not a member of the homologous PRT family. In addition, it has a different structure when compared to quinoline PRTase, the ‘type II PRTase’ mentioned in the text.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

• of special interest
•• of outstanding interest


23. Erskins TA, Kadiola A, Bentsen A, Harlow KW, Larsen S: Structural basis for the function of Bacillus subtilis phosphoribosylpyrophosphate synthetase. Nat Struct Biol 2000, 7:303-308. The structure of PRPP synthetase reveals that, in addition to the expected PRT fold for the C-terminal domain, the N-terminal domain also belongs to the PRT family. This is the only identified distant relative of the PRT family.
