PROTEINS: Structure, Function, and Genetics 38:310–326 (2000)

When Fold Is Not Important: A Common Structural Framework for Adenine and AMP Binding in 12 Unrelated Protein Families

Konstantin A. Denessiouk[1,2] and Mark S. Johnson[1,2]*
[1]*Department of Biochemistry and Pharmacy, Åbo Akademi University, Turku, Finland*
[2]*Turku Centre for Biotechnology, Åbo Akademi University and University of Turku, Turku, Finland*

ABSTRACT **ATP is a ligand common to many proteins, yet it is unclear whether common recognition patterns do exist among the many different folds that bind ATP. Previously, it was shown that cAMP-dependent protein kinase, D-Ala:D-Ala ligase and the α-subunit of the $\alpha_2\beta_2$ ribonucleotide reductase do share extensive common structural elements for ATP recognition although their folds are different. Here, we have made a survey of structures that bind ATP and compared them with the key features seen in these three proteins. Our survey shows that 12 different fold types share a specific recognition pattern for the adenine moiety, and 8 of these folds have a common structural framework for recognition of the AMP moiety of the ligand. The common framework consists of a tripeptide segment plus three additional residues, which provides similar polar and hydrophobic interactions between the protein and mononucleotide. Consensus interactions are represented by four key hydrogen bonds present in each fold type. Two of these four hydrogen bonds, together with three aliphatic residues, form a specific recognition pattern for the adenine moiety in all 12 folds. These similarities point to a structural-functional requirement shared by these different mononucleotide-binding proteins that represent at this time 28% of the adenine mononucleotide complexes found in the Brookhaven Protein Data Bank. Proteins 2000;38:310–326.**

Key words: ATP; ATP recognition; binding site; mononucleotide; convergent structural similarities

INTRODUCTION

Adenosine 5′-triphosphate (ATP) plays a central role in energy exchange in living cells and also serves as a vital ligand of a wide variety of enzymes required to carry out their function. There is significant diversity among the folds of ATP-dependent enzymes, and mononucleotides are recognized in numerous ways in different complexes.[1] The concept of a fuzzy recognition template currently describes the recognition of adenylate by proteins.[2] It states that the general recognition pattern consists of a sandwich-like structure with hydrophobic residues above and below the adenine rings and polar residues around its rim. The identification of a key conserved set of structural elements responsible for the local interactions in different folds that bind ATP is an important step toward understanding the common principles of mononucleotide recognition by proteins.

The molecular structure of ATP consists of a purine base, a ribose ring, and three phosphate groups. Those enzymes that bind ATP and have the classic mononucleotide-binding fold contain a β-strand-loop-α-helix supersecondary structure, which is a characteristic feature of the phosphate-binding site.[3,4] The phosphate-binding loop (P-loop) sequence, Gly/Ala-X-X-X-X-Gly-Lys-Thr/Ser (X refers to any residue), which is known as the Walker A motif,[5] extends from the end of the β-strand (residue Gly/Ala) to the beginning of the α-helix (residues Lys-Thr/Ser) and forms an anion hole for the phosphates (reviewed by Smith and Rayment[6]). The Walker B motif Arg/Lys-X_{1-4}-Gly-X_{2-4}-ϕ-X-ϕ-ϕ-Asp/Glu, where "ϕ" is a hydrophobic amino acid, functions to coordinate a bound Mg^{2+} cation. The conserved Asp/Glu residue of the Walker B motif was found to bind one or several water molecules as bridging ligands to the active site Mg^{2+} in different complexes.[7,8]

It has been shown that the recA protein, elongation factor Tu, and several other proteins with the Walker A motif have similar α- and β-phosphate positions, but quite different positions of both the ribose and the base (guanine in EF-Tu and adenine in recA).[9] Thus, in these structures, binding of the phosphates by the P-loop is found to be independent of the other interactions between the protein and mononucleotide. In addition to the Walker A and B motifs, two other motifs have been identified: the *Kinase-3* sequence motif occurs in at least four different variants in 16 protein structures,[10] whereas a comparison of shikimate kinase with adenylate kinase identified an adenine-binding motif found in a small group of ATP-binding proteins.[11] In this latter group, the NH_2 group at position N6 of adenine is tightly bound via two hydrogen bonds, one directly to the backbone carbonyl atom of glutamine and

Grant sponsor: Academy of Finland; Grant sponsor: Technology Development Centre for Finland (TEKES); Grant sponsor: Graduate School of Informational and Structural Biochemistry; Grant sponsor: Erna and Victor Hasselblad Foundation.

*Correspondence to: Mark S. Johnson, Department of Biochemistry and Pharmacy, Åbo Akademi University, Tykistökatu 6, BioCity 3A, P.O. Box 66, SF-20521 Turku, Finland. E-mail: johnson@abo.fi

Received 14 May 1999; Accepted 23 September 1999

another to a water molecule that is coordinated to the backbone carbonyl of alanine and to a residue of the P-loop.

Aminoacyl-transfer ribonucleic acid (tRNA) synthetases, in the presence of Mg^{2+} and ATP, specifically attach amino acids to the 3′-adenosine of their cognate tRNAs and are found in all organisms.[12–14] On the basis of short conserved primary sequence motifs, the 20 aminoacyl-tRNA synthetases have been divided into two classes.[15–17] The active site of enzymes belonging to class I aminoacyl-tRNA synthetases is formed out of the canonical Rossmann fold,[18] whereas the fold of the catalytic domain of the class II enzymes, in contrast, is formed from an antiparallel β-sheet.[19] Two common sequence elements interact with the ATP phosphates in the class I aminoacyl-tRNA synthetases.[16] One element consists of an 11-amino acid signature sequence that ends with the tetrapeptide His-X-Gly-His (the HIGH motif, where X is a hydrophobic residue). The other consensus element is the pentapeptide Lys-Met-Ser-Lys-Ser (the KMSKS motif). In class II aminoacyl-tRNA synthetases, three specific motifs were detected, and two of the motifs (motif 2 and motif 3) involve invariant and semi-invariant residues from the active site.[16,17]

Adenine can, in principle, form five hydrogen bonds: two from the hydrogen-bond donor at position N6 and one each from the hydrogen-bond acceptors at positions N1, N3, and N7. Kobayashi and Go[20,21] have identified an adenine-binding motif common to D-Ala:D-Ala ligase (DD-ligase)[22,23] from the ATP-grasp fold family[24–34] and the catalytic subunit of cyclic adenosine monophosphate (cAMP)-dependent protein kinase (cAPK)[35] from the protein kinase fold family.[36–39] This motif consists of a four-residue segment that forms two hydrogen bonds with the adenine base of ATP. One bond is formed between the NH_2 group at position N6 of the adenine ring and the backbone carbonyl of the second residue of the four-residue segment, and the other hydrogen bond is formed between the N1 nitrogen of adenine and the backbone amide NH group of the fourth residue.

When we examined the structures surrounding the four-residue segment reported by Kobayashi and Go,[20,21] we found that cAPK and DD-ligase have more than 100 Cα-atoms in common that form similar supersecondary structures around the ATP-binding sites.[40] In addition, the cofactors, two metal cations, and several water molecules are similarly oriented in both cAPK and DD-ligase. Later, when we compared the 103 residues common to cAPK and DD-ligase against all adenine-mononucleotide containing complexes in the Brookhaven Protein Data Bank (PDB),[41] we found that only the R1 subunit of ribonucleotide reductase (RNR R1)[42] shares large regions of structural similarity about its ATP-binding site with cAPK and DD-ligase.[43] These regions include three strands from the β-sheet and a loop (the adenosine-binding loop) following one of the strands.

In the above three folds, the key features of the ATP recognition pattern are four separate residues (not four contiguous residues, as above, although two residues are part of the four-residue motif introduced by Kobayashi and Go) that form hydrogen bonds with the AMP moieties of the bound ligands. Two of the four hydrogen bonds of the adenine-binding structural motif are formed between the adenine base of the ligands and two residues that are located at the amino-terminal end of the adenosine-binding loop. A third hydrogen bond is formed between the O2′ oxygen of the ribose ring and a residue from the carboxy-terminal end of the adenosine-binding loop. An invariant lysine forms the fourth hydrogen bond of the recognition pattern with the α-phosphate group. This lysine is conserved throughout the three families[22,33,36,42,44–46] and located in the central strand of the three-stranded β-sheet.[43] When we analyzed all of the ATP-containing complexes in the PDB that have similar structural elements for ATP recognition as we have seen in cAPK, DD-ligase, and RNR R1, we found that these key features of the recognition pattern are characteristic of a wide variety—28%—of the ATP-containing protein complexes deposited with the PDB.

MATERIALS AND METHODS

Mononucleotide-Binding Data Set

All adenine mononucleotide-protein complexes were extracted from the April 1999 release of the PDB. We considered only those complexes that contained bound ATP-analogues, bound ATP, bound ADP, bound AMP and in some cases bound aminoacyl-adenylates. The initial data set included 197 PDB entries.

First, we identified all structures that form two hydrogen bonds with the adenine moiety of the ligand: one bond between the NH_2 group at position N6 of the ligand and the main-chain carbonyl oxygen atom of residue X and another bond between the N1 atom of the ligand and the main-chain amide NH group of residue Y. Residues X and Y must be, respectively, the amino- and carboxy-terminal residues of a tripeptide segment (adenine-binding structural motif), as in the cAPK, DD-ligase and RNR R1 complexes.[43] The main-chain NH and O atoms of these proteins, which interact with the N1 and N6 atoms of the ligand, may also belong to the same residue. To obtain hydrogen-bonding data we have used the Ligand-Protein Contacts (LPC) software[47] given on the WEB page for each PDB entry (http://www.pdb.bnl.gov). A generous 3.7 Å cutoff distance between the donor and acceptor of a potential hydrogen bond was used. The strength of the hydrogen bond around 3.7 Å is negligible; however, it served to identify potential hydrogen bonding atoms, most of which lie between 2.6 and 3.3 Å of each other. In cases in which the LPC results were not available, we used visual inspection and the molecular modeling software package SYBYL (Tripos Associates, Inc., St. Louis, MO) to identify hydrogen bonds between the cofactor and the protein. Next, we eliminated any identical binding sites represented more than once in multisubunit structures: 57 different mononucleotide-binding sites were obtained.

Furthermore, we compared the structure of cAPK complexed with adenylyl imidodiphosphate (AMP-PNP) and two Mn^{2+} ions (PDB code: 1CDK, A chain) with each of these 57 mononucleotide-binding sites after superposition-

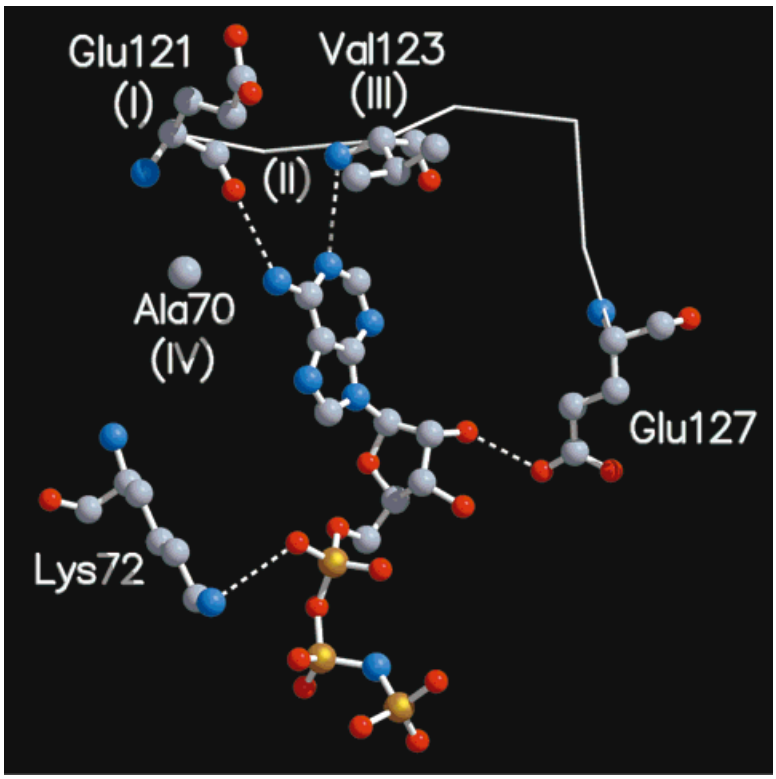

Fig. 1. The AMP-binding structural motif found in cAPK, DD-ligase, and RNR R1. Amino acid residues are shown from the structure of cAPK. Roman numerals in brackets indicate hydrophilic site I, hydrophobic site II, and aliphatic sites III and IV from Table I, respectively. Four key hydrogen bonds (dashed lines) bind the AMP moieties of the ATP ligands in these three proteins: one between the NH_2 group at position N6 of the adenine ring and the backbone carbonyl of the residue at hydrophilic site I; one between the N1 nitrogen of adenine and the backbone amide NH group of the residue at aliphatic site III; one between the O2′ hydroxyl group of the ribose ring and the side-chain oxygen from the residue at the carboxy-terminal end of the loop containing hydrophilic site I and aliphatic site III; and one between two oppositely charged groups: lysine—invariant throughout the three families—and the α-phosphate group. The side-chain of the residue is shown if the residue forms a hydrogen bond with the ligand. The large gray sphere designates the position of the Cα atom of Ala70 at aliphatic site IV. Figures 1–7 were drawn by using the program MOLSCRIPT.[94] Figures 1 and 5 were rendered by using Raster3D V2.3.[95]

ing their adenine moieties. The structure of bacteriophage T4 deoxynucleotide kinase complexed with AMP (code: 1DEL)[48] did not fit the pattern seen in cAPK and was removed. This reduced the data set to 54 entries containing 29 different proteins. Two of the 54 entries had two different mononucleotide-binding sites; thus, there were 56 mononucleotide-binding sites, having two unique hydrogen bonds between protein and the adenine moiety of the ligand.

Next, we identified those complexes that also have a similar pattern of hydrogen bonding to both the O2′ atom of the ribose and the α-phosphate group as was seen with cAPK, DD-ligase, and RNR R1. Forty-eight complexes match the pattern seen in cAPK. In the eight remaining complexes, the hydrogen-bonding patterns did not match that seen in cAPK: four failed to match the interactions with the α-phosphate group, two failed to match the interactions with ribose, and in two cases to both the α-phosphate group and the ribose. Therefore, the first eight families include 46 entries (23 different proteins) with 48 AMP-binding sites that have a common hydrogen-bonding pattern between the protein and the ligand. Each of these 48 AMP-binding sites contains a highly conserved hydrophobic residue structurally equivalent Ala70 in cAPK and known to be highly conserved among protein kinases.[36,37]

RESULTS AND DISCUSSION

We have extracted all structures from the PDB (197 PDB entries) that are liganded with ATP, ADP, AMP, or their analogues. Among them, 29 different proteins (54 PDB entries) were identified that have a similar mode of binding for their mononucleotide cofactor as is seen in cAPK, DD-ligase and RNR R1. These 29 protein structures represent 13 different protein folds, but as we shall show below two of these families can be combined together into one.

Protein Kinase Family Fold

In cAPK, a representative of the protein kinase family fold, adenine is bound to the enzyme with two hydrogen bonds. One is formed between the backbone carbonyl of Glu121 (designated site I, Fig. 1 and Table I) and the NH_2 group at position N6 of adenine, and a second is formed between the backbone amide NH group of Val123 (aliphatic site III) and the N1 nitrogen of the purine. This specific hydrogen-bonding pattern plays an essential role in distinguishing between adenine and other bases. These same positions, N1 and N6 on adenine in the mononucleotides, participate in Watson-Crick base pairing in double-stranded DNA. In cAPK, the adenine moiety is flanked and stabilized by the nonpolar side-chains of Tyr122 (hydrophobic site II) and Val123 (aliphatic site III). In addition, Ala70 (aliphatic site IV) directly interacts with one face of the adenine ring. The ribose O2′ hydroxyl group is tightly bound via a hydrogen bond with the side-chain of Glu127, and the side-chain of Lys72 forms an electrostatic link with the oxygen atom of the α-phosphoryl group.

At this time, there are 12 other enzyme-mononucleotide complexes in the PDB (seven different proteins) that belong to the protein kinase family and have the same mode of cofactor binding as is seen in the cAPK-AMP-PNP

complex (Table I). These proteins are as follows: cAPK in complex with ATP,[49] the Src family tyrosine kinase (Hck),[50] the insulin receptor tyrosine kinase (IRK),[51] casein kinase-1 (CK1),[39] cyclin-dependent kinase 2 (CDK2) complexed with (PDB codes: 1FIN[52] and 1JST[53]) and without cyclin (PDB codes: 1HCK;[54] and 1B38 and 1B39), phosphorylase kinase (Phk)[55,56] and extracellular signal-regulated protein kinase 2 (ERK2).[57]

In some cases, the distance between the donor and the acceptor of a potential hydrogen bond exceeds 4.0 Å (Table I). For example, in CK1 the O2′ hydroxyl group of the ribose ring of ATP interacts with the side-chain of Ser91 via an intervening water molecule. In the CDK2-cyclinA complex (PDB code: 1FIN), Lys33 is distant from the α-phosphate group (4.7Å, data not shown), and this is due to conformational changes that occur on binding cyclinA.[52] In apo-CDK2 (1HCK), the Lys33-α-phosphate distance is 2.9 Å.

In the protein kinase family, the key lysine residue that interacts with the α-phosphate group of the cofactor is invariant, and an alanine residue at aliphatic site IV is highly conserved.[36,37] Hydrophilic site I is occupied by a negatively charged glutamic or aspartic acid, whereas bulky hydrophobic residues Met, Leu, Val, Phe, or Tyr are found at hydrophobic site II and aliphatic site III (Table I).

ATP-Grasp Fold Proteins

Our comparisons defined a data set of seven different PDB entries (four different proteins) that contains nine adenine-binding sites in their ATP-grasp domains (Table I). Three proteins are ligases: DD-ligase,[22,23] glutathione synthetase (GSHase; PDB code: 1GSH),[58] and carbamoyl phosphate synthetase (CPS).[32,59] The fourth protein is a fragment of synapsin I, whose structure is similar to this group of ATP-grasp-fold enzymes (SynC).[33] The large subunit of CPS (PDB codes: 1A9X and 1JDB in Table I) has two independent nonidentical ATP-binding sites of the ATP-grasp type[32,59]; both are listed in Table I.

Although GSHase and DD-ligase have the same fold and similar ATP-binding sites,[25,26] Asp208 in GSHase (column 5 in Table I) forms a hydrogen bond with the O3′ hydroxyl group of ribose (distance is 2.6 Å) instead of the O2′ hydroxyl group (4.7 Å) as is seen in all of the other ATP-grasp folds. Whether this difference is due to the effects of binding glutathione and sulfate or the other ligands present in the complex[58] is not known at this time. However, residues Arg210 and Arg225 in the vicinity of Asp208 are rearranged when the *apo*-form of the enzyme *sans* ATP is compared with the structure of the complex.[58] In the unliganded enzyme, atom OD1 of Asp208 interacts with the NH_2 position of Arg210 (3.6 Å), whereas in the complex a hydrogen bond is formed between the OD1 atom of Asp208 and the side-chain NH_1 position of Arg225 (2.9 Å).

The lysine residue whose side-chain interacts electrostatically with the α-phosphate group of the ATP-ligand is conserved over the entire superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity that contain ATP-grasp domains.[45,46] This lysine is substituted by arginine only in CPS. Sites II–IV, which are occupied mostly by large nonpolar residues and make side-chain interactions with the adenine rings, are also highly conserved. Only CPS, which has two separate mononucleotide-binding sites, has a polar serine at hydrophobic site II in one of these two sites (Table I). Site IV in the vicinity of ATP-grasp fold proteins has hydrophobic residues much larger than the conserved alanine residue found in the protein kinases, and these large hydrophobic residues, like alanine, have their side-chains pointed directly toward the adenine ring and perpendicular to its plane. Hydrophilic site I (Lys181 in DD-ligase) is usually occupied by a charged amino acid in the ATP-grasp fold proteins, and the residue that forms a hydrogen bond to the O2′ hydroxyl group of ribose (Glu187 in DD-ligase) is usually glutamic acid, aspartic acid or histidine.[45,46]

In the protein structure of SAICAR synthetase (PDB code: 1A48),[60] which consists of three domains, it was reported that one domain is similar to cAPK and another domain is similar to DD-ligase. The crystal structure does not contain an adenylate ligand in the ATP-binding site. However, the loop His110–His112 in SAICAR synthetase, which is structurally analogous to the adenine-binding loop Glu121–Val123 in cAPK, joins the two domains that form the ATP-binding cleft, as seen for Glu121–Val123 in cAPK and Lys181–Leu183 in DD-ligase (sites I, II, and III in Table I). In addition, the hydrophobic residue Leu31 in SAICAR synthetase occupies a position structurally similar to the position of Ala70 (aliphatic site IV in cAPK, Table I).

Ribonucleotide Reductase Subunit R1

Previously, we showed that cAPK, DD-ligase, and RNR R1 have a common mode of AMP-binding.[43] The only prominent difference between RNR R1 and cAPK or DD-ligase in the organisation of the ATP-binding site is the occupation of aliphatic site III by a polar residue, Asn18, in RNR R1 (Table I). However, the side-chain oxygen and nitrogen atoms of this residue are positioned away from the adenine ring. Multiple alignment of the amino-terminal 100 residues involved in ATP binding in the available RNR R1 sequences shows that aliphatic site III is frequently occupied by methionine (see Fig. 7 in Eriksson et al.[42]). Other key positions are invariant: Val7 and Ile22, whereas Lys9 is highly conserved (Table I). Hydrophilic site I is always occupied by a polar residue, whereas hydrophobic site II contains only a bulky hydrophobic amino acid. Furthermore, Glu15 in RNR R1 and the corresponding conserved Glu180 in DD-ligase form one additional common hydrogen bond between the OE2 atom of glutamate and the NH_2 group at the N6 position of the cofactor.

Dethiobiotin Synthetase and the Nitrogenase Iron (Fe)-Protein

Dethiobiotin synthetase (DTBS) participates in biotin biosynthesis and catalyzes the formation of the ureido ring of dethiobiotin from (7R,8S)-7,8-diaminononanoic acid and carbon dioxide.[61,62] The nitrogenase iron (Fe)-protein (ni-

TABLE I. Consensus Polar and Aliphatic Interactions in Mononucleotide-Binding Sites of 29 Protein- ATP Complexes

Protein name, PDB code, chain, resolution	Source, ligand	Distance (Å) between protein atom (residue) and the corresponding ligand atom				Hydrophilic site I	Hydrophobic site II	Aliphatic site III	Aliphatic site IV
		N1	N6	O2′	O1A[b]/O2A[c]				
1	2	3	4	5	6	7	8	9	10
Protein kinases									
cAPK	Porcine	N(Val123)	O(Glu121)	OE2(Glu127)	NZ(Lys72)[b]	Glu121	Tyr122	Val123	Ala70
(1CDK, A chain, 2.0 Å)	AMP-PNP	3.3	3.0	2.6	3.1				
cAPK	Mouse	N(Val123)	O(Glu121)	OE2(Glu127)	NZ(Lys72)[b]	Glu121	Tyr122	Val123	Ala70
(1ATP, 2.2 Å)	ATP	3.3	3.0	2.6	2.5				
Hck	Human	N(Met341)	O(Glu339)	OD2(Asp348)	NZ(Lys295)[b]	Glu339	Phe340	Met341	Ala293
(1AD5, B chain, 2.6 Å)	AMP-PNP	3.2	3.2	3.4	3.2				
IRK	Human	N(Met1079)	O(Glu1077)	OD1(Asp1083)	NZ(Lys1030)[b]	Glu1077	Leu1078	Met1079	Ala1028
(1IR3, 1.9 Å)	AMP-PNP	3.0	3.0	2.9	2.8				
CK1	Yeast	N(Leu88)	O(Asp86)	OG(Ser91)	NZ(Lys41)[b]	Asp86	Leu87	Leu88	Ala39
(1CSN, 2.0 Å)	ATP	2.9	2.9	4.9	2.7				
CDK2	Human	N(Leu83)	O(Glu81)	OD2(Asp86)	NZ(Lys33)[b]	Glu81	Phe82	Leu83	Ala31
(1HCK, 1.9 Å) (1B38; 1B39; 1FIN; 1JST)[a]	ATP	3.3	2.8	3.5	2.9				
Phk	Rabbit	N(Met106)	O(Asp104)	OE2(Glu110)	NZ(Lys48)[b]	Asp104	Leu105	Met106	Ala46
(1PHK, 2.2 Å); (2PHK)[a]	ATP	2.8	2.8	3.0	2.9				
ERK2	Rat	N(Met106)	O(Asp104)	OD2(Asp109)	NE(Arg52)[b]	Asp104	Leu105	Met106	Ala50
(1GOL, 2.8 Å)	ATP	2.8	2.5	2.4	4.2				
ATP-grasp fold proteins									
DD-ligase	*E. coli*	N(Leu183)	O(Lys181)	OE1(Glu187)	NZ(Lys144)[b]	Lys181	Trp182	Leu183	Ile142
(1IOW, 1.9 Å); (1IOV, 2DLN)[a]	ADP	2.8	2.9	2.8	2.9				
GSHase	*E. coli*	N(Leu201)	O(Asn199)	OD2(Asp208)	NZ(Lys160)[b]	Asn199	Tyr200	Leu201	Ile158
(1GSA, 2.0 Å)	ADP	2.9	3.1	4.7	2.7				
CPS	*E. coli*	N(Leu210)	O(Glu208)	OE1(Glu215)	NH1(Arg169)[b]	Glu208	Ser209	Leu210	Ile167
(1A9X, A chain, 1.8 Å); (1JDB)[a]	ADP	3.0	2.9	2.7	2.8				
CPS	*E. coli*	N(Leu756)	O(His754)	OE1(Glu761)	NH1(Arg715)[b]	His754	Phe755	Leu756	Val713
(1A9X, A chain, 1.8 Å); (1JDB)[a]	ADP	2.9	3.1	2.7	2.6				
SynC	Bovine	N(Ile308)	O(Pro306)	OD2(Asp313)	NZ(Lys269)[b]	Pro306	Phe307	Ile308	Val267
(1AUX, B chain, 2.3 Å)	ATPγS	2.9	3.2	3.7	2.7				
Ribonucleotide reductase R1									
RNR R1	*E. coli*	N(Asn18)	O(Arg16)	N(Ile22)	NZ(Lys9)[c]	Arg16	Ile17	Asn18	Val7
(3R1R, A chain, 3.0 Å)	AMP-PNP	2.7	3.2	3.2	2.7				
Dethiobiotin synthetase and the nitrogenase iron (Fe)-protein									
DTBS	*E. coli*	N(Leu206)	O(Pro204)	OE1(Glu211)	N(Thr16)[c]	Pro204	Trp205	Leu206	Val177
(1DAH, 1.64 Å) (1DAD;1DAF;1DAG;1DAM; 1BS1)[a]	AMP-PCP	2.9	3.1	3.1	3.0				
Fe-protein	*A. vinelandii*	N(Asp214)	O(Pro212)	OE1(Glu221)	N(Ser16)[c]	Pro212	Arg213	Asp214	Arg187
(1N2C, E chain, 3.0 Å)	ADP	3.2	3.3	3.3	3.2				
Glycogen phosphorylase									
GP	Rabbit	N(Gly317)	O(Lys315)	OD2(Asp42*)	NH1(Arg310)[c]	Lys315	Phe316	Gly317	Tyr75
(1PYG, A chain, 2.87 Å)	AMP	2.8	3.3	2.6	3.3				
Aspartate carbamoyltransferase									
ATCase	*E. coli*	N(Ile12)	O(Ile12)	O(Val9)	NZ(Lys94)[b]	Lys13	Ile12	Ala11	Ile86
(4AT1, B chain, 2.6 Å); (7AT1)[a]	ATP	2.6	3.4	2.9	3.6				
Class I tRNA synthases									
GlnRS	*E. coli*	N(Leu261)	O(Leu261)	N(Thr230)	NZ(Lys270)[c]	Asn262	Leu261	Arg260	Gly42
(1GTR, 2.5 Å) (1GTS; 1QRS; 1QRT; 1QRU)[a]	ATP	3.0	3.2	3.5	3.0				

Class II tRNA synthases and Asparagine synthetase									
AspRS	*P. kodakaraensis*	N(Leu224)	O(Leu224)	O(Ile362)	NH1(Arg214)[c]	Asn225	Leu224	His223	Ala227
(1B8A, A chain, 1.9 Å)	ATP	3.0	2.8	2.8	2.8				
AspRS	yeast	N(Met335)	O(Met335)	O(Glu478)	NH2(Arg325)[c]	Thr336	Met335	His334	Phe338
(1ASZ, A chain, 3.0 Å)	ATP	3.3	2.8	3.4	2.7				
HisRS	*T. thermophilus*	N(Tyr121)	O(Tyr121)	O(Leu282)	NH2(Arg112)[b]	Arg122	Tyr121	Arg120	Phe124
(1ADY, A chain, 2.8 Å)	His-AMP	3.1	3.5	2.8	3.1				
HisRS	*E. coli*	N(Tyr122)	O(Tyr122)	O(Val282)	NH2(Arg113)[c]	Arg123	Tyr122	Arg121	Phe125
(1KMN, A chain, 2.8 Å) (1HTT; 1KMM)[a]	ATP	3.0	2.6	3.3	2.9				
SerRS	*T. thermophilus*	N(Val272)	O(Val272)	O(Thr346)	NH1(Arg256)[b]	His273	Val272	Arg271	Phe275
(1SES, B chain, 2.5 Å)	SerHx-AMP	3.0	2.8	2.6	2.6				
GlyRS	*T. thermophilus*	N(Val232)	O(Val232)	O(Leu305)	NH2(Arg220)[c]	Arg233	Val232	Arg231	Phe235
(1B76, A chain, 3.4 Å); (1GGM)[a]	ATP	2.9	2.7	3.1	3.6				
AsnA	*E. coli*	N(Ser111)	O(Ser111)	O(Leu249)	NH2(Arg100)[c]	Val112	Ser111	His110	Val114
(12AS, A chain, 2.2 Å)	AMP	3.2	3.0	3.0	3.1				
Hypothetical protein MJ0577 and electron transfer flavoprotein									
MJ0577	*M. jannaschii*	N(Val41)	O(Val41)	O(Pro11)	N(Val142)[b]	Ile42	Val41	His40	Pro108
(1MJH, 1.7 Å)	ATP	3.1	2.8	2.7	2.8				
ETF	Human	N(Cys66)	O(Cys66)	O(Ala9)	—	Gly67	Cys66	Ser65	Val104
(1EFV, B chain, 2.1 Å)	AMP	3.1	3.1	2.8					
Amidotransferases									
GMP synthetase	*E. coli*	N(Val260)	O(Val260)	O(Gly233)	—	Asp261	Val260	Phe259	Phe315
(1GPM, B chain, 2.2 Å)	AMP	3.0	2.9	2.7					
NAD^+ synthetase	*B. subtilis*	N(Leu79)	O(Leu79)	O(Gly44)	—	Pro80	Leu79	Arg78	Arg139
(1NSY, A chain, 2.0 Å)	AMP	3.2	3.4	2.7					
Trimethylamine dehydrogenase	Methylotrophic bacterium	N(Met470)	O(Met470)	OG1(Thr420)	—	Thr471	Met470	Pro469	Val395
(2TMD, A chain, 2.4	ADP	3.1	3.3	3.3					
Adenylate kinase									
Aksa	*S. acidocaldarius*	N(Gly180)	O(Val178)	—	N(Ser15)[c]	Val178	Glu179	Gly180	Arg132
(1NKS, F chain, 2.57 Å)	ADP	2.8	2.7		3.6				
Exotoxin A									
PEIII	*Pseudomonas*	N(Arg456)	O(Gly454)	—	—	Gly454	Val455	Arg456	Ile450
(1AER, B chain, 2.3 Å); (1DMA)[a]	AMP	3.2	3.2						

[a] PDB code of the same protein structure from the same organism that has identical ligand and similar binding site to that of the representative structure (shown in bold).
[b,c] The atom is bound to either the oxygen O1A or to the oxygen O2A of the α-phosphate group of the AMP-moiety of the ligand.

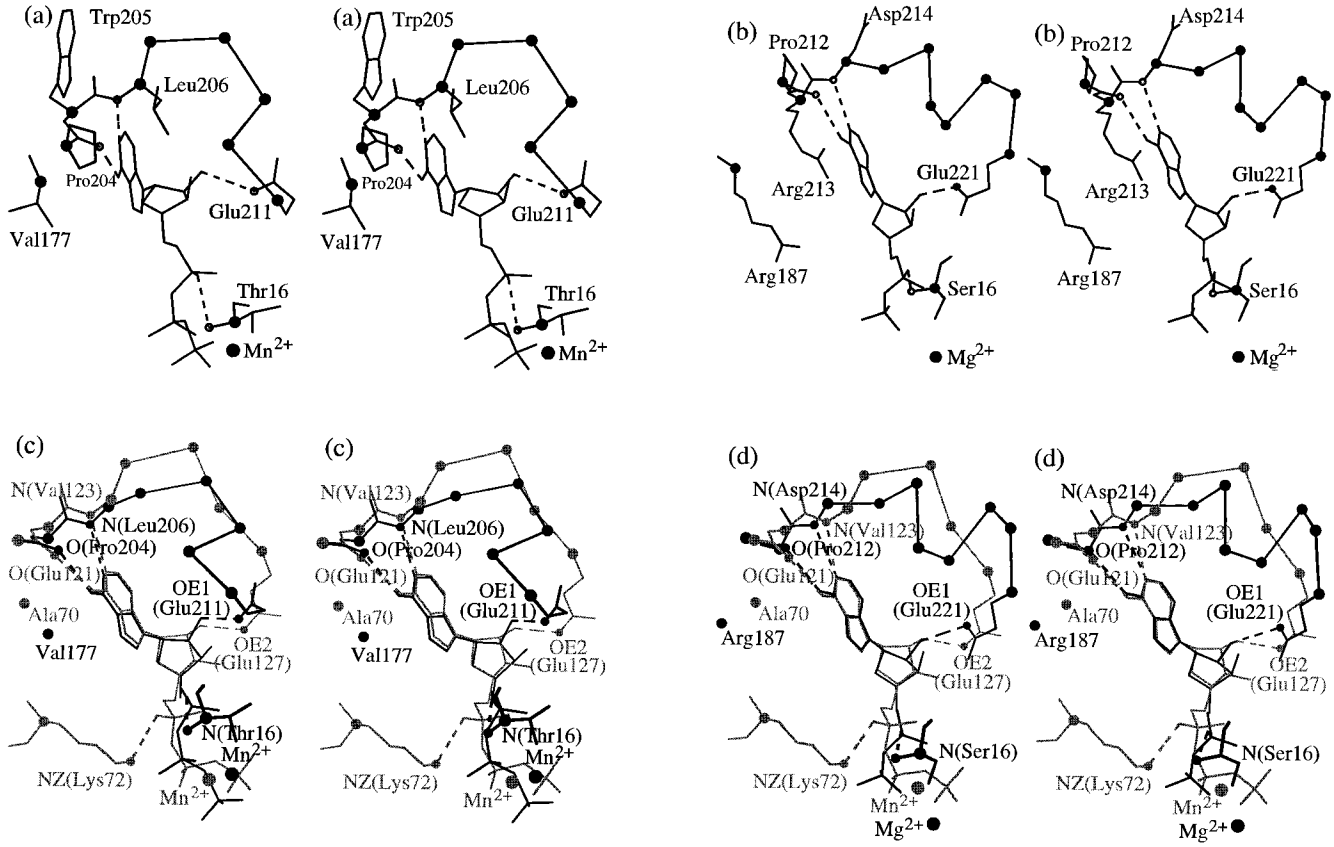

Fig. 2. Stereo views illustrating the ATP-binding motifs of (**a**) DTBS in complex with AMP-PCP and (**b**) the nitrogenase Fe-protein in complex with ADP, and their superposition with the cofactor binding site of cAPK: (**c**) DTBS (black) with cAPK (gray); and (**d**) the nitrogenase Fe-protein (black) with cAPK (gray). In (a) DTBS, the main-chain carbonyl of Pro204 (hydrophilic site I) and the amide NH group of Leu206 (aliphatic site III) form hydrogen bonds with adenine. Val177 occupies aliphatic site IV. The adenine ring is sandwiched between Trp205 (site II) and Leu206 (site III). In the nitrogenase Fe-protein (b), the main-chain carbonyl of Pro212 (site I) and the amide NH group of Asp214 (site III) form hydrogen bonds with adenine. Sites II and IV contain arginine residues, Arg213 and Arg187, whose guanidinium groups stack with the adenine ring. The side-chain of Asp214 at site III does not point toward the adenine ring as does Leu206 in DTBS. In the superpositions of (c) DTBS and cAPK and (d) the nitrogenase Fe-protein and cAPK, the side-chains of residues at sites I–IV are not shown. Small circles indicate atoms that form hydrogen bonds (dashed lines) with the ligand. Intermediate-sized circles indicate the positions of Cα-atoms. Larger circles designate the positions of the bound metal cations.

trogenase Fe-protein) is only one component of the large nitrogenase complex responsible for nitrogen fixation.[63] DTBS and the nitrogenase Fe-protein have segments with structural characteristics similar to those seen in the adenosine-binding loop from the protein kinase fold family, the ATP-grasp fold family, and RNR R1 (Fig. 2a and b). These segments are Pro204-Glu211 in DTBS and Pro212-Glu221 in the nitrogenase Fe-protein (Table I). The side-chains interacting with adenine are highly conserved in the sequences of DTBSs and the nitrogenase Fe-proteins.[64] Superpositions of the ligands AMP-PCP (in DTBS) and ADP (in the nitrogenase Fe-protein) with AMP-PNP (in cAPK) are shown in Figure 2c and d, respectively. These superpositions clearly show that the conformations of the ligands and the relative location of metal cations in DTBS and nitrogenase Fe-protein mononucleotide complexes are very similar to what is seen in the cAPK mononucleotide complex. The corresponding root-mean-square deviations (RMSD) are equal to 0.9 Å for DTBS versus cAPK (24 common ligand atoms) and 0.8 Å for the nitrogenase Fe-protein versus cAPK (22 common ligand atoms).

The structure of both DTBS and the nitrogenase Fe-protein contain the phosphate-binding Walker A motif (P-loop) and the metal cation-binding Walker B motif: Gly8-X-X-X-X-Gly-Lys-Thr16 and Glu115-X-X-Gly118 in DTBS,[61] and Gly9-X-X-X-X-Gly-Lys-Ser16 and Asp125-X-X-Gly128 in the nitrogenase Fe-protein.[63] Figure 3a and b shows the common network of electrostatic interactions among residues of the Walker A and B motifs, divalent metal cations and the phosphate groups in cAPK and DTBS (Fig. 3a), and in cAPK and the nitrogenase Fe-protein (Fig. 3b). The main-chain nitrogen of the last residue of the Walker A motif, Thr16 in DTBS (Ser16 in the nitrogenase Fe-protein), bridges two oxygen atoms from the α- and β-phosphates of the cofactor. Lys72 performs a similar function in cAPK.[36,37] The manganese ion Mn^{2+} coordinates oxygens of the β- and γ-phosphoryl groups in the structures of both cAPK and DTBS, and Mg^{2+} fulfils this role in the nitrogenase Fe-protein struc-

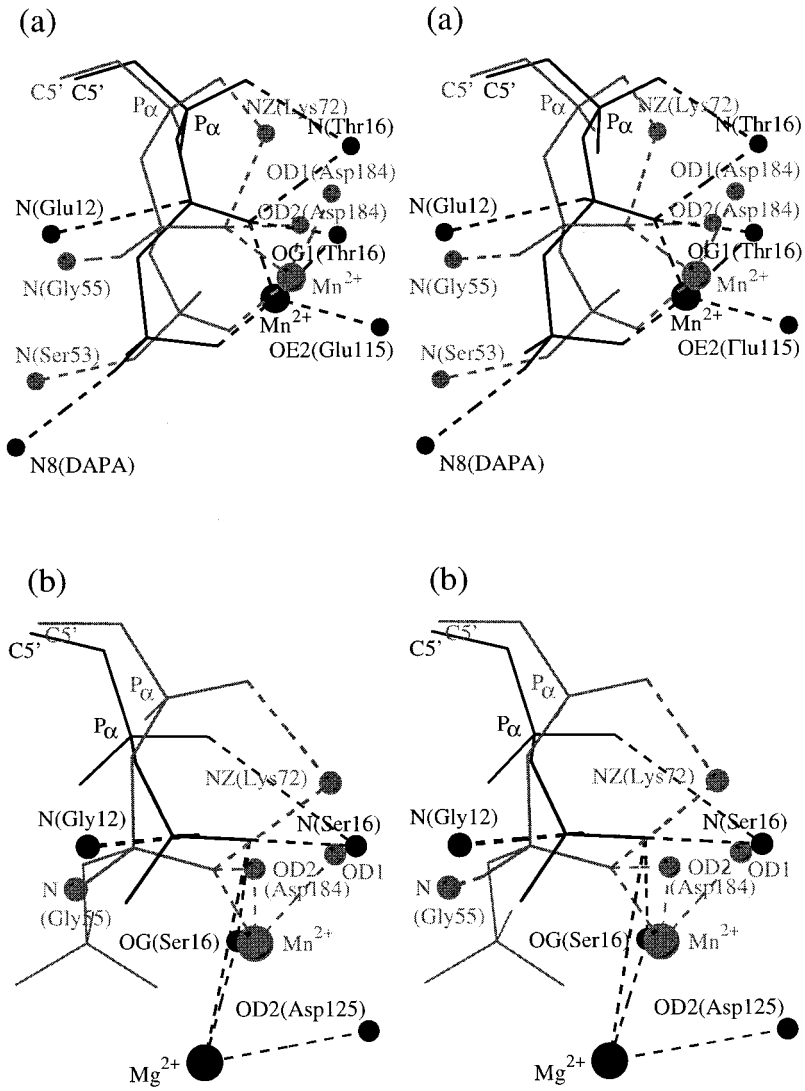

Fig. 3. Similar networks of electrostatic interactions between the α- and β-phosphate groups of the ligands, metal cations, and several conserved amino acids in (**a**) cAPK (gray) and DTBS (black) complexes and (**b**) cAPK (gray) and the nitrogenase Fe-protein (black) complexes. Bound Mn^{2+} and Mg^{2+} ions are shown as large spheres. The main-chain nitrogen of the highly conserved Gly55 (in cAPK) and the main-chain nitrogen of Glu12 and Gly12 from the P-loop of DTBS and the nitrogenase Fe-protein, respectively, interact with the β-phosphates. There is a common network of electrostatic links between atom OD2 of Asp184 (cAPK; invariant in the kinase family), atom OG1 of Thr16 (in the P-loop of DTBS), atom OG of Ser16 (in the P-loop of the nitrogenase Fe-protein), and the oxygen atoms from the β-phosphates and metal cations. In (a), the cAPK and DTBS complexes show additional equivalent hydrogen bonds between the γ-phosphate groups and the main-chain nitrogen of Ser53 (highly conserved in Ser/Thr kinases) and the eight-amino group of (7R,8S)-7,8-diaminononanoic acid (the substrate DAPA in DTBS), respectively. DTBS and the nitrogenase Fe-protein also share considerable structural similarities with adenylate kinase[96] (family 12 in Table I) about the phosphate-binding sites (data not shown).

ture. The side-chain oxygen of the first residue of the conserved Walker B motif, atom OE2 of Glu115 in DTBS (side-chain atom OD2 of Asp125 in the nitrogenase Fe-protein), interacts electrostatically with the metal cation and forms a hydrogen bond to the side-chain oxygen of Thr16 (Ser16). Atom OD1 of Asp184 is likely to play an equivalent role in cAPK.

It was previously reported that the overall folds of DTBS and the nitrogenase Fe-protein are similar to the GTP-dependent protein H-Ras p21.[7,65,66] However, when we aligned DTBS and the nitrogenase Fe-protein complexes with the H-Ras p21 complex (PDB code: 5P21), we found that the loop Pro204-Asn212 between β-strand S1 and α-helix H9 in DTBS (Pro212-Lys235 between β-strand S1 and α-helix H115 in the nitrogenase Fe-protein), which contains the adenine-binding motif (Table I), has no equivalent structure in H-Ras p21. Instead, a short loop, Ser145-Ala146-Lys147, which is important for guanine-binding specificity, connects the corresponding β-strand S6 and α-helix HE in the tertiary structure of H-Ras p21. In H-Ras p21, the highly conserved regions Asn116-Lys117-Cys118-Asp119 and Ser145-Ala146-Lys147 provide the most important electrostatic and hydrophobic interactions with guanine.[67] Asn116 from the *Kinase-3b* segment[10] is part of the highly conserved NKXD motif and binds specifically to the N7 position of the guanine moiety.[7] Asn175 (DTBS) and Asn185 (the nitrogenase Fe-protein) are structurally equivalent to Asn116 in H-Ras p21, and they also form hydrogen bonds to the N7 nitrogen atom but in adenine. Thus, we do see similarities between guanine binding to H-Ras p21 and adenine binding to DTBS and the nitrogenase Fe-protein.

Glycogen Phosphorylase

The major physiological substrate of phosphorylase kinase (Phk), a member of the protein kinase family, is glycogen phosphorylase (GP). Glycogen phosphorylase can be independently activated both by phosphorylation and by AMP binding. Binding of AMP occurs at the activation locus that partially includes the interface between the two

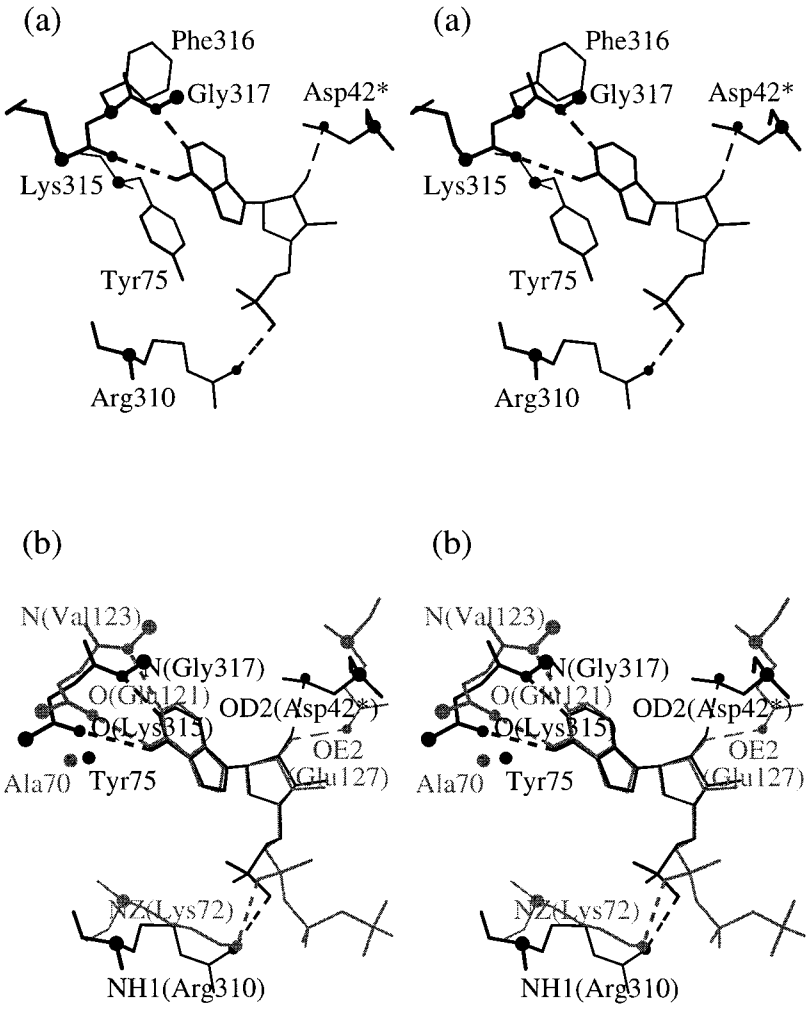

Fig. 4. Stereo view of (**a**) the AMP-binding structural motif in glycogen phosphorylase and (**b**) the superposition of the AMP-binding motifs in cAPK (gray) and GP (black). The ATP cofactor is situated at the interface between the two subunits of the GP homodimer, and Asp42* (from the other subunit of the dimer) makes a hydrogen bond to the ribose O2′ hydroxyl group.

subunits of the GP homodimer.[68] Features of the AMP-binding site in GP are similar to those in cAPK (Table I; Fig. 4a and b), and the 20 equivalent atoms of their ligands superpose closely with an RMSD of 0.3 Å. The main-chain carbonyl oxygen of Lys315 and the amide NH group of Gly317 hydrogen bond to the N6 amino group and the N1 nitrogen of adenine, respectively. Asp42 from the other subunit in the dimer hydrogen bonds to the ribose O2′ hydroxyl group. The O2A atom of the phosphate group of AMP forms a hydrogen bond to Arg310. The plane of the purine ring of AMP is parallel to the plane of the side-chain ring of Tyr75 (the hydrophobic site IV).

Aspartate Carbamoyltransferase

Aspartate carbamoyltransferase (ATCase, PDB code: 4AT1)[69] and cAPK have a similar hydrogen-bonding pattern for adenosine recognition (Table I), but the adenosine-binding loops in the two complexes have opposite chain directions. ATCase catalyzes the reaction between carbamoyl phosphate and L-aspartate to give phosphate and *N*-carbamoyl-L-aspartate. ATP stimulates the enzymatic activity of this allosteric enzyme by binding to the regulatory subunits. The short amino-terminal loop from ATCase, Val9-Lys13, is involved in ATP binding and is shown in Figure 5. The conformations of the AMP moieties of the ligands in cAPK and ATCase are similar and can be superimposed with an RMSD of 0.8 Å (24 identical atoms; Fig. 5).

A striking feature of the ATP/ATCase complex is that the hydrogen-bonding pattern for adenine recognition is formed from a single residue, Ile12, instead of from two residues as is seen above. Ile12 is located at hydrophobic site II (Table I), whose main-chain oxygen and nitrogen form two hydrogen bonds to the N6 and N1 positions of adenine in ATP (Fig. 5). At the same time, the side-chain of Ile12 is hydrophobic and interacts with the adenine ring as do the nonpolar residues found at hydrophobic site II in the other structures above (Table I). The hydrophobic side-chains of Ala11 and Ile86 are also in contact with the adenine rings. The O2′ hydroxyl group of ribose is hydrogen bonded to the main-chain carbonyl group of Val9, and the side-chain NZ atom of Lys94 forms hydrogen bonds with oxygens of both the α- and β-phosphates. However, in the crystal structure of ATCase complexed with phosphonoacetamide, malonate, and ATP (PDB code: 7AT1)[70] the hydrogen bond between Val9 and ribose is absent.

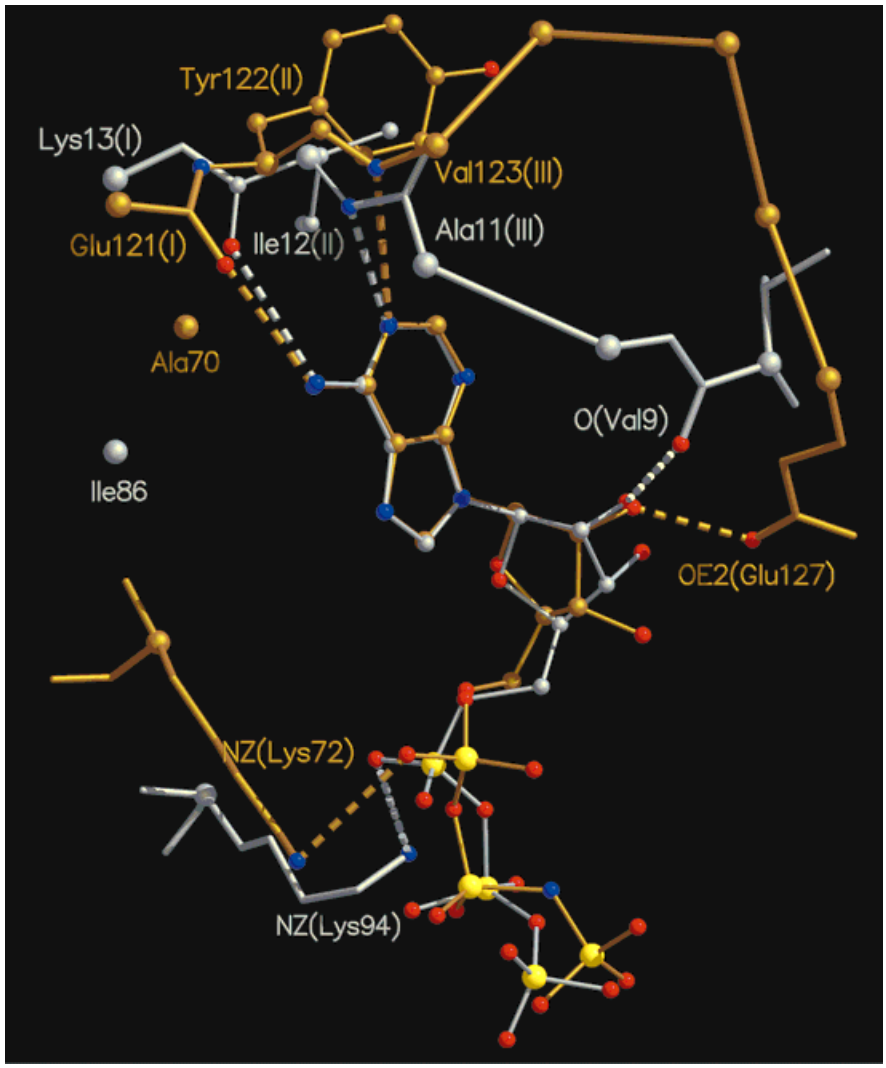

Fig. 5. Superposition of the ATP-effector binding motif in ATCase (gray) with the ATP-binding motif in cAPK (yellow). The amino-terminal tetrapeptide Val9-Ile12 (ATCase) supplies three hydrogen bonds and two hydrophobic side-chains in the construction of the ATP-binding site. Although the hydrogen bonds to the N6 and N1 positions are nearly identically placed in both ATCase and cAPK, the chain direction in ATCase is opposite to that seen in cAPK. Furthermore, in ATCase both hydrogen bonds come from the residue at site II, whereas in cAPK they come from residues at sites I and III.

Aminoacyl-tRNA Synthetases and Asparagine Synthetase

Five aminoacyl-tRNA synthetases, glutaminyl-tRNA synthetase (GlnRS; PDB code: 1GTR;[71] code: 1GTS;[72] codes: 1QRS, 1QRT, and 1QRU[73]), seryl-tRNA synthetase (SerRS; PDB code: 1SES),[74] aspartyl-tRNA synthetase (AspRS; PDB code: 1B8A;[75] code: 1ASZ[76]), histidyl-tRNA synthetase (HisRS; PDB code: 1ADY;[77] code: 1HTT;[78] codes: 1KMM and 1KMN[79]), glycyl-tRNA synthetase (GlyRS; PDB codes: 1B76 and 1GGM) and asparagine synthetase (AsnA; PDB code: 12AS),[80] which is not a tRNA synthetase, all have hydrogen-bonding patterns for ATP binding similar to that seen in cAPK and the five other families described above (Table I).

Class I Aminoacyl-tRNA Synthetases

GlnRS belongs to the class I tRNA synthetases; both of the signature motifs common to the class I synthetases, the HIGH motif (residues 40–43) and the KMSKS motif (residues 268–270), are located in the catalytic cleft of the enzyme. In GlnRS, the adenine ring of the cofactor is sandwiched between two residues (Fig. 6a): Gly42 that is invariant for class I synthetases (aliphatic site IV in Table I) and the aliphatic portion of the side-chain of Arg260 (aliphatic site III). Gly42 (aliphatic site IV) is particularly interesting. It is not an aliphatic residue and is an exception to the rule. This glycine is part of the conserved HIGH motif that forms part of an α-helix running along the face of the base with glycine positioned directly against the base, and there is no room available for any side-chain. In each of other families in this study, the corresponding position is located on a β-strand where there is ample room for the placement of a bulky side-chain. The positively charged side-chain of the invariant Lys270, the second lysine in the KMSKS motif, donates a hydrogen bond to one oxygen of the α-phosphate.

The adenine-binding loop in GlnRS, which incorporates the tripeptide segment Arg260-Asn262, has the same chain direction as is seen for ATCase and opposite to that found in cAPK. Consequently, like with ATCase, a single hydrophobic residue from aliphatic site II,

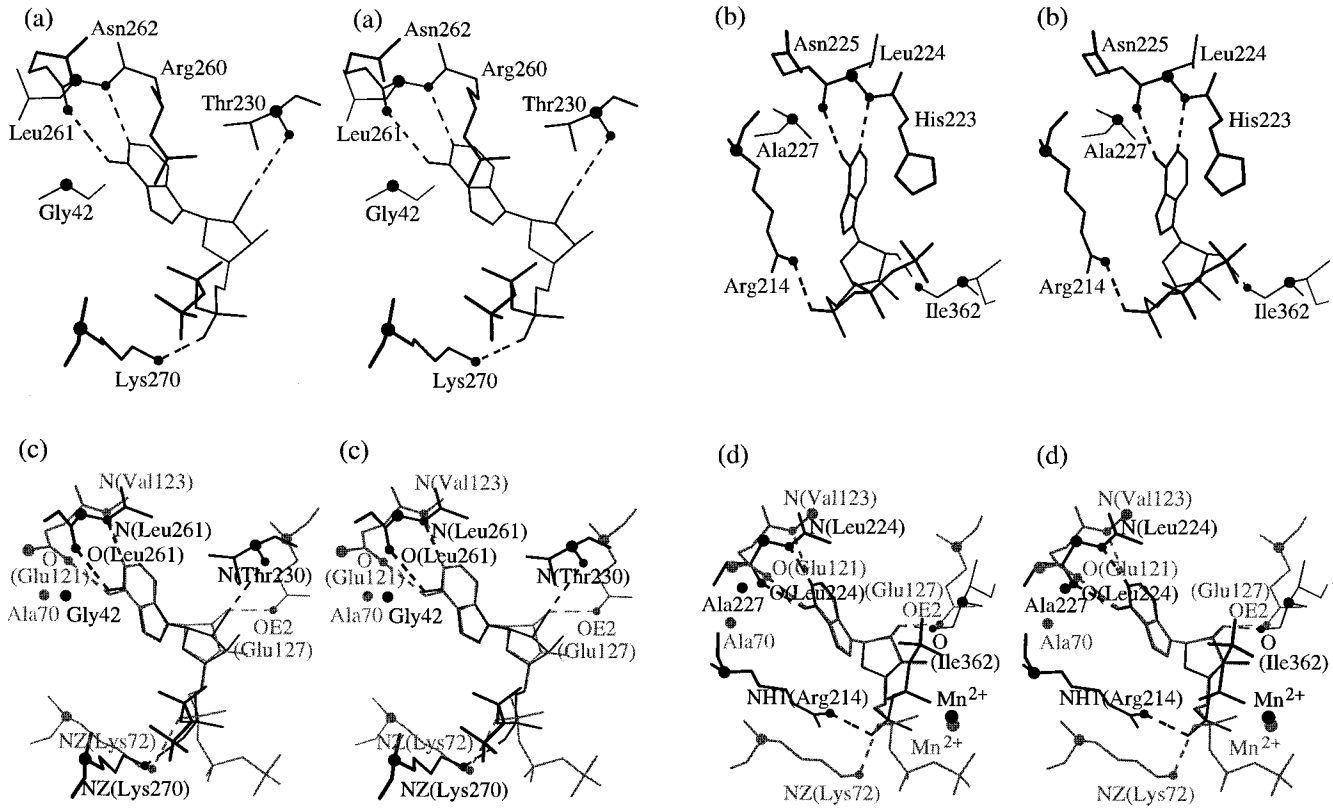

Fig. 6. The (**a**) GlnRS-ATP complex and (**b**) AspRS-ATP complex as representatives of the class I and class II aminoacyl-tRNA synthetases, respectively. Comparison of the ATP-binding sites of (**c**) GlnRS (black) and cAPK (gray) and of (**d**) AspRS (black) and cAPK (gray). In (a) and (b) Arg260 (GlnRS) and His223 (AspRS) from aliphatic site III direct their side-chains toward the negatively charged ATP phosphates in the two complexes. In the GlnRS-ATP complex, a centrally located water molecule bridges the guanidinium group of Arg260 to the β-phosphate of ATP.[72] A structurally similar network of stabilizing hydrogen-bonding interactions points to the existence of a common AMP-binding pattern in these three complexes.

Leu261 (Table I), forms two hydrogen bonds to the N1 and N6 positions of the ATP-cofactor and, thus, determines the specificity of the enzyme for ATP (Figs. 5 and 6a). The side-chain of Leu261 participates in hydrophobic interactions with the adenine ring in a similar way as Ile12 does in ATCase. Thr230 interacts through its backbone amide group with the 2′-OH of ribose, and a polar residue Asn262 is found at hydrophilic site I (Table I).

Class II Aminoacyl-tRNA Synthetases and Asparagine Synthetase

The other three aminoacyl-tRNA synthetase mononucleotide complexes, AspRS, HisRS, and SerRS, are members of the class II tRNA synthetase family. They lack the HIGH motif but contain three other motifs specific for this class.[16,19] The AMP recognition pattern is organized in a similar way among these enzymes and is shown in Figure 6b for AspRS. As seen above for both class I aminoacyl-tRNA synthetases and ATCase, hydrogen bonds to the N6 and N1 positions of the adenine ring are provided by the main-chain carbonyl oxygen and the amide NH group from a single residue, Leu224 (hydrophobic site II, Table I). Leu224 also provides stabilizing contacts to the adenine moiety of the cofactor with its side-chain hydrophobic groups and thus plays a similar role in AspRS as Leu261 does in GlnRS and Ile12 does in ATCase. In the AspRS-ATP complex, the purine ring of the cofactor is stabilized by an alanine residue, Ala227 (aliphatic site IV, Table I), which is generally a phenylalanine residue in the class II aminoacyl-tRNA synthetases (column 10, Table I). However, the change of this conserved aromatic residue from phenylalanine to alanine in AspRS (in motif 2) does not affect the relative position of the adenine moiety.[75] The side-chain NH group of Arg214 (invariant in class II tRNA synthetases) makes a hydrogen bond with the O2A oxygen of the α-phosphate. Arg214 is part of the motif 2 sequence specific for the class II tRNA synthetases.[16,19] In the families of protein kinases, ATP-grasp fold proteins, RNR R1, ATCase, and class I tRNA synthetases, a lysine usually resides in the same position as Arg214 in the class II tRNA synthetases (column 6 in Table I). The similarity and variability in the spacing of the tripeptide segment and the three amino acids that are combined to form the pocket for ATP binding in cAPK, GlnRS, and AspRS are shown in Figure 6c and d. The positions of the tripeptide segments are very similar.

Hypothetical Protein MJ0577 and Human Electron Transfer Flavoprotein

The crystal structure of the human electron transfer flavoprotein (ETF; code: 1EFV)[81] contains bound AMP and flavin adenine dinucleotide. The mononucleotide plays a structural role in ETF. The role of ATP in MJ0577 (1MJH)[82] is not known; it was initially purified and crystallized without ATP, but, surprisingly, the final structure contained ATP, possibly taken from its *Escherichia coli* host during expression.[82] Binding of the adenine-ribose moieties of their ligands is similar in both proteins. They no longer share the common AMP-binding motif with the previous eight groups of proteins but do share the common binding motif for adenine and ribose. Examination of the contacts between the α-phosphate group and the residues surrounding it in ETF and in cAPK showed that there are no apparent structural similarities in these interactions in the two complexes (Table I).

Amidotransferases

GMP synthetase and NAD^+ synthetase catalyze similar reactions and belong to the newly discovered family of N-type ATP pyrophosphatases.[83,84] The modes of adenine binding to these synthetases are identical and explain the ATP specificity of the enzymes. The backbone oxygen and nitrogen of Val260 (GMP synthetase) and Leu79 (NAD^+ synthetase) form hydrogen bonds with the N6 and N1 positions of the adenine ring of the bound AMP in each case (Table I). Moreover, the adenine rings are stabilized by hydrophobic interactions with the side-chains of Val260 and Leu79 and by ring stacking with the phenyl ring of Phe315 (GMP synthetase) and with the nonpolar part of the side-chain of Arg139 (NAD^+ synthetase) in the AMP-binding pockets of the two enzymes. Hydrophobic contacts with Phe259 (GMP synthetase) and with the aliphatic part of Arg78 (NAD^+ synthetase) on the other side of the adenine rings close the adenine-binding pockets in both enzymes. The O2′ hydroxyl group of ribose is hydrogen bonded to the carbonyl atom of Gly233 in GMP synthetase and Gly44 in NAD^+ synthetase. Visual examination of the residues that stabilize the α-phosphate groups in GMP and NAD^+ synthetases did not reveal any structural similarity to the cAPK complex.

The GMP and NAD^+ synthetase complexes include, in addition to their AMP cofactors, pyrophosphate (PP_i), Mg^{2+} and water molecules in their active sites.[83,84] Magnesium forms a bridge between the β- and γ-phosphates of PP_i (nomenclature by analogy to ATP), the phosphate of AMP, and one of the water molecules in each complex. The mode of binding of the metal cations is remarkably similar in these three enzyme complexes, where the divalent cation (Mn^{2+} or Mg^{2+}) is coordinated by four ligands: the α-phosphate group, a water molecule, and the side-chain oxygens of two invariant residues, Asn171 and Asp184 (Fig. 7a and b).

Trimethylamine Dehydrogenase, Adenylate Kinase and Exotoxin A (Families 11, 12, and 13)

The crystal structure of the iron-sulfur flavoprotein, trimethylamine dehydrogenase from the methylotrophic bacterium W3A1,[85] contains ADP and flavin mononucleotide. The role of ADP in the complex is unknown. Moreover, its presence in the crystal structure was unexpected.[85] However, this protein has characteristics of the adenosine-binding site in common with the previous 10 families (Table I).

The structure of adenylate kinase from the hyperthermophilic archaean *Sulfolobus acidocaldarius* (AKsa) shares the fold common to the large family of nucleoside monophosphate kinases.[86] The enzyme is trimeric, and the asymmetric unit of AKsa contains six subunits gathered into two trimers. The ADP molecule is seen in one ATP-binding site (chain F); three other sites are occupied by AMP (chains A, B and E), whereas chains C and D have their ATP-binding sites unoccupied. The α- and β-phosphates of ADP in the F subunit of AKsa (Table I) are bound via the P-loop, whose amino acid sequence, Gly8-X-X-X-X-Gly13-Lys14-Ser15 is the well-known Walker A motif.[5] The adenine moiety of ADP is bound to AKsa by hydrogen bonds between the NH_2 group at the N6 position of ADP and the carbonyl oxygen of Val178, and between the nitrogen at the N1 position and the main-chain amide of Gly180. The guanidinium group of Arg132 interacts electrostatically with the purine base. This arginine, which is strongly conserved in the nucleoside monophosphate kinase family,[86] is stabilized by a hydrogen bond to Glu179.

The structure of the catalytic domain III of *Pseudomonas aeruginosa* exotoxin A (PEIII) in the presence of nicotinamide adenine dinucleotide (NAD) revealed that NAD^+ was hydrolyzed, and only the products, AMP and nicotinamide, were bound to the enzyme.[87,88] When we compare the interactions of monomer 2 (chain B) of PEIII (complexed with AMP) and cAPK (complexed with AMP-PNP), we found that the adenine-binding motif in PEIII is similar to the adenine-binding motif in cAPK (Table I). The ribose rings and the α-phosphate groups of the two complexes do not superimpose.

Overview of Protein Recognition of Adenylate

The distribution of elements along the polypeptide chain that comprise the AMP-binding motif in the 13 families is shown in Figure 8. The adenine-binding motif, which includes two hydrogen bonds to adenine at the N1 and N6 positions and three nonpolar contacts between the protein and adenine, is observed in 54 PDB entries for 29 different ATP-binding proteins represented by these 13 families (approximately 28% of the initial data set). Twenty of the 29 proteins fall into the first eight structural families and share common features of their AMP-recognition sites (adenine-, ribose- and α-phosphate-binding motifs): (a) the fold of the protein kinase family, (b) the ATP-grasp fold, (c) the ribonucleotide reductase fold, (d) dethiobiotin synthetase and the nitrogenase Fe-protein, (e) glycogen phosphorylase, (f) the regulatory subunit of aspartate carbamoyltransferase, (g) class I aminoacyl-tRNA synthetases,

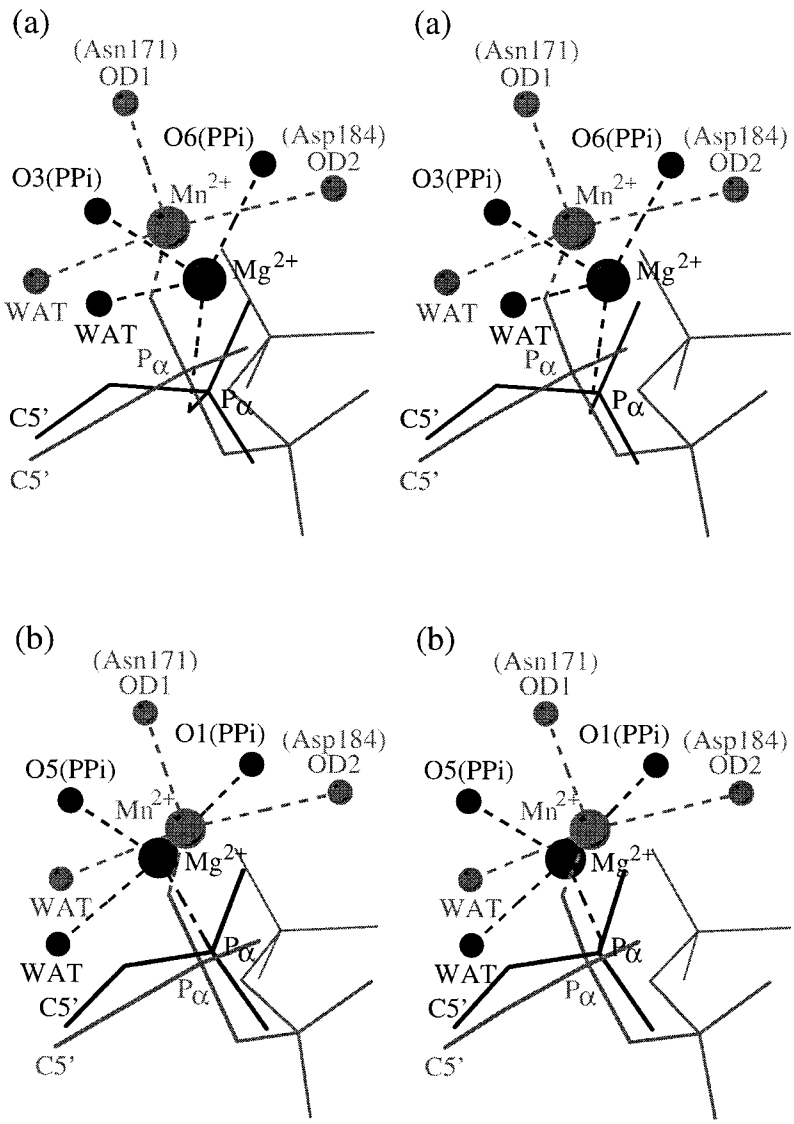

Fig. 7. Stereo pairs show the superposition of the metal-binding sites of (**a**) GMP synthetase (black) and cAPK (gray), and (**b**) NAD^+ synthetase (black) and cAPK (gray). One α-phosphate oxygen, one water molecule, and two oxygens from the surrounding residues or from pyrophosphate (PP_i) provide for the similar coordination of the cation in each complex. Note that in addition to these similarities the adenine-binding motifs in GMP synthetase and in NAD^+ synthetase are also similar (data not shown) and correspond to the motif seen in ATCase.

and (h) class II aminoacyl-tRNA synthetases plus asparagine synthetase. These common features include a tripeptide segment and three separate residues, which provide structurally similar polar and hydrophobic interactions between the protein and the ligand.

The tripeptide segment forms two specific hydrogen bonds with the N1 (aliphatic site III in Table I) and the N6 (hydrophilic site I in Table I) positions of the purine ring (N1-, N6-selective pattern). This pattern is constructed in two different ways (Fig. 5). In one case, which occurs in each of the first five families of folds (families 1–5 in Fig. 8), the main-chain carbonyl oxygen of the first residue (hydrophilic site I in Table I) and the main-chain amide NH group of the third residue (aliphatic site III in Table I) of the tripeptide segment form two hydrogen bonds to adenine (cAPK in Fig. 5). In the second case (for example, ATCase in Fig. 5), the direction of the tripeptide segment is opposite to that shown for cAPK in Figure 5, and the main-chain atoms of the central residue of the tripeptide segment make two hydrogen bonds (ATCase, site II in Fig. 5). This occurs in three families of proteins: aspartate carbamoyltransferase and the class I and class II aminoacyl-tRNA synthetases (families 6–8 in Fig. 8). The spatial locations of the N1-, N6-selective patterns, relative to adenine, are identical in both cases (Fig. 5). Site I is usually occupied by a polar residue. The side-chains of the other two consecutive residues in the tripeptide segments are hydrophobic (sites II and III) and make contacts with the two different faces of adenine (Figs. 1 and 5). Hydrophobic site II mainly contains aromatic and bulky hydrophobic residues. Bulky hydrophobic residues are also accommodated at aliphatic site III, but, sometimes, arginine is found there too and interacts with the adenine ring with the long aliphatic part of its side-chain. These findings are in agreement with the data presented by Moodie et al.[2] They found that most of the aromatic side-chains lie on one side of the bound mononucleotide. In addition, nonpolar interactions with adenine come from aliphatic site IV,

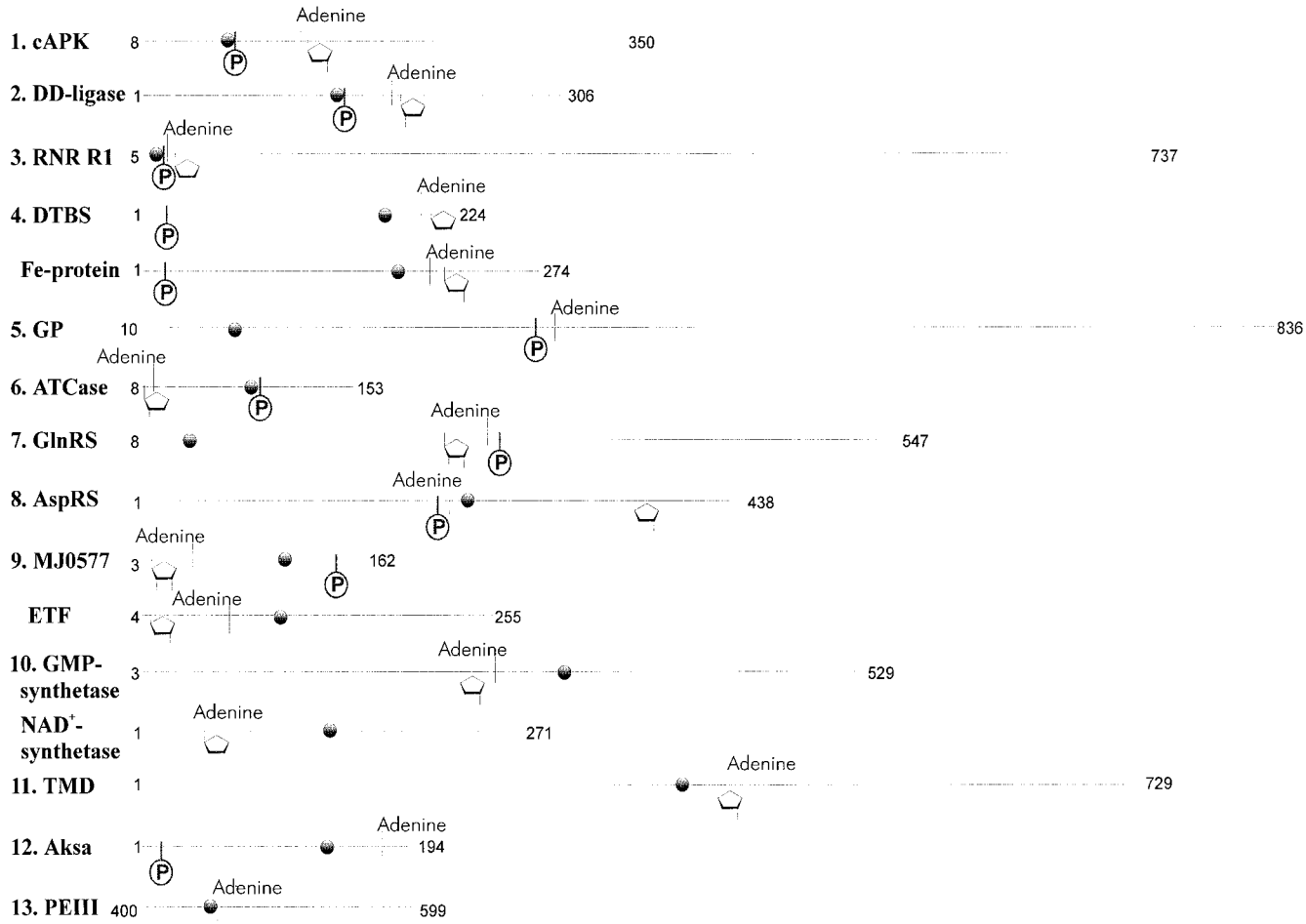

Fig. 8. Thirteen groups of structures and the distribution of the elements of their AMP-binding motifs along the polypeptide chain. Each family is represented by one protein detailed in Table I. For families 9 and 10 both structures are given. Lengths of the polypeptide chains are indicated. The position of the first residue of the adenine-binding tripeptide is marked by the word "Adenine" and points to the right if the chain direction of the adenine-binding loop is the same as seen in cAPK, DD-ligase, and RNR R1, and to the left if the direction is opposite (example: ATCase). The position of the ribose ring is also shown. The position of the residue from aliphatic site IV is shown by a gray circle, and the position of the phosphate group by a circle with the letter "P" within. Families 1–8, 12, and MJ0577 from family 9 have complete AMP-binding motifs, whereas families 10, 11, 13, and ETF from family 9 do not. The four proteins from families 9 and 10 have the same distribution along the polypeptide chain of the adenine-binding loop, ribose, and the residue at aliphatic site IV.

which is spatially close to hydrophobic site II. Sites II, III, and IV form a common nonpolar pattern for recognition of the adenine moiety of the mononucleotide.

Originally, we had grouped the folds of MJ0577 and ETF together in one structural family (family 9), and GMP synthetase and NAD^+ synthetase into another (family 10). The similarity between MJ0577 and ETF was first suggested by Zarembinski et al.,[82] where the top-scoring structural matches to MJ0577 and found by using the program DALI[89–91] include ETF, DNA photolyase (PDB code: 1QNF), and the tyrosyl-tRNA synthetase (PDB code: 2TS1). When we used DALI to compare ETF with PDB structures, we found that in addition to a high-scoring match with MJ0577, ETF (Fig. 8) also matched both GMP synthetase and NAD^+ synthetase significantly (Fig. 8).

When we compared the structures of MJ0577, ETF, GMP synthetase, and NAD^+ synthetase, we found that the adenine-binding motif, the ribose-binding motif, and the hydrophobic residue at site IV have the same relative positions along each of the polypeptide chains. In addition, these four enzymes share a four-stranded parallel β-sheet (A, C, E, and G in Fig. 9) and three α-helices (B, D, and F in Fig. 9). The seven common segments form identical β-α-β-α-β-α-β supersecondary structure from contiguous elements along the polypeptide chain. In each of the four enzymes, the residue forming the hydrogen bond to the O2′ oxygen of ribose is placed at the carboxy terminus of β-strand A. Residues from hydrophilic site I and nonpolar sites II and III (Table I) form hydrogen bonds with the nitrogens at positions N1 and N6 of adenine and are placed at the carboxy-terminal part of β-strand C. The amino terminus of α-helix F incorporates a residue from aliphatic site IV (Fig. 9).

The comparison of ETF with PDB structures using the program DALI has shown that a number of other protein structures share extensive structural similarities with

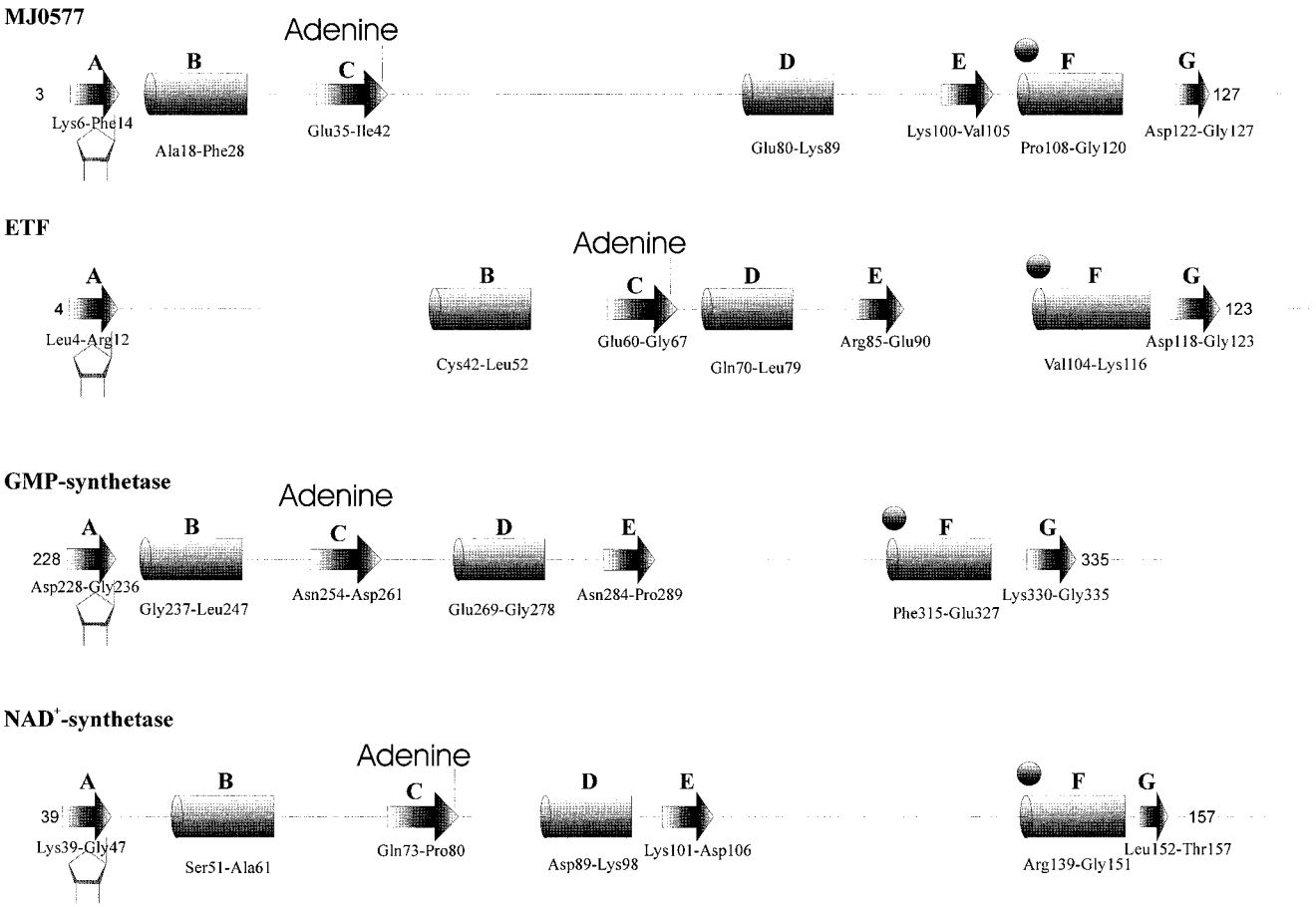

Fig. 9. A block diagram of the equivalent segments that form the cofactor-binding sites in MJ0577, ETF, GMP synthetase, and NAD^+ synthetase. The seven common segments form identical β-α-β-α-β-α-β supersecondary structures, and they are distributed contiguously along the polypeptide chains. β-strands, arrows; α-helices, cylinders. Positions of the adenine-binding loop, ribose, and the residue from aliphatic site IV are shown and designated as in Figure 8.

ETF. Among them were DNA photolyase from *Anacystis Nidulans* (PDB code: 1QNF),[92] biotin carboxylase subunit of Acetyl-CoA carboxylase (PDB code: 1BNC,[30] A chain), the flavodoxin reductase from *E. coli* (PDB code: 1FDR),[93] and others. All of these proteins share a parallel β-sheet flanked by α-helices exactly as seen in MJ0577, EFV, GMP synthetase, and NAD^+ synthetase (segments A–G in Fig. 9). This type of supersecondary structure is a dinucleotide-binding motif (Rossmann fold)[18] and is present in many other protein structures. Thus, although we originally began with 13 family groups, in the final analysis families 9 and 10 have the same folds over their ATP-binding domains and are homologous over these regions.

Finally, the proteins represented by families 1–8 (Fig. 8) do have different folds, yet there are basic structural characteristics responsible for AMP-binding that are common among all eight families. These common features constrain the AMP moiety of the ligands to the same relative orientation. The ATP-binding sites of these proteins shows a remarkable case of convergence toward the same structural motif.

ACKNOWLEDGMENT

The authors thank our colleague Jukka Lehtonen for valuable discussions.

REFERENCES

1. Schulz GE. Binding of nucleotides by proteins. Curr Opin Struct Biol 1992;2:61–67.
2. Moodie SL, Mitchell JBO, Thornton JM. Protein recognition of adenylate: an example of a fuzzy recognition template. J Mol Biol 1996;263:486–500.
3. Schulz GE, Schirmer RH. Topological comparison of adenylate kinase with other proteins. Nature 1974;250:142–144.
4. Schulz GE, Schiltz E, Tomasselli AG, Frank R, Brune M, Wittinghofer A, Schirmer RH. Structural relationships in the adenylate kinase family. Eur J Biochem 1986;161:127–132.
5. Walker JE, Saraste M, Runswick MJ, Gay NJ. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J 1982;1:945–951.
6. Smith CA, Rayment I. Active site comparisons highlight structural similarities between myosin and other P-loop proteins. Biophys J 1996;70:1590–1602.
7. Pai EF, Krengel U, Petsko GA, Goody RS, Kabsch W, Wittinghofer A. Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 angstroms resolution: implications for the mechanism of GTP hydrolysis. EMBO J 1990;9:2351–2359.

8. Abele U, Schulz GE. High-resolution structure of adenylate kinase from yeast ligated with inhibitor Ap_5A, showing the pathway of phosphoryl transfer. Protein Sci 1995;4:1262–1271.
9. Story RM, Steitz TA. Structure of the recA protein-ADP complex. Nature 1992;355:374–376.
10. Traut TW. The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. Eur J Biochem 1994;222:9–19.
11. Krell T, Coggins JR, Lapthorn AJ. The three-dimensional structure of shikimate kinase. J Mol Biol 1998;278:983–997.
12. Doolittle WF, Brown JR. Tempo, mode, the progenote, and the universal root. Proc Natl Acad Sci USA 1994;91:6721–6728.
13. Nagel GM, Doolittle RF. Phylogenetic analysis of the aminoacyl-tRNA synthetases. J Mol Evol 1995;40:487–498.
14. Brown JR, Doolittle WF. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. Proc Natl Acad Sci USA 1995;92:2441–2445.
15. Eriani G, Delarue M, Poch O, Gangloff J, Moras D. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. Nature 1990;347:203–206.
16. Moras D. Structural and functional relationships between aminoacyl-tRNA synthetase. Trends Biochem Sci 1992;17:159–164.
17. Cusack S. Aminoacyl-tRNA synthetases. Curr Opin Struct Biol 1997;7:881–889.
18. Rossmann MG, Liljas A, Branden C-I, Banaszak LJ. Evolutionary and structural relationships among dehydrogenases. In: Boyer PD, editor. The enzymes. Vol 11. New York: Academic Press; 1975. p 61–102.
19. Cusack S, Berthet-Colominas C, Hörtlein M, Nassar N, Leberman R. A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å. Nature 1990;347:249–255.
20. Kobayashi N, Go N. ATP binding proteins with different folds share a common ATP-binding structural motif. Nat Struct Biol 1997;4:6–7.
21. Kobayashi N, Go N. A method for similar protein local structures at ligand-binding sites and its application to adenine recognition. Eur Biophys J 1997;26:135–144.
22. Fan C, Moews PC, Walsh CT, Knox JR. Vancomycin resistance: structure of D-alanine: D-alanine ligase at 2.3 Å resolution. Science 1994;266:439–443.
23. Fan C, Park IS, Walsh CT, Knox JR. D-alanine: D-alanine ligase: phosphonate and phosphinate intermediates with wild type and the Y216F mutant. Biochemistry 1997;36:2531–2538.
24. Artymiuk PJ, Poirrette AR, Rice DW, Willett P. Biotin carboxylase comes into the fold. Nat Struct Biol 1996;3:128–132.
25. Fan C, Moews PC, Shi Y, Walsh CT, Knox JR. A common fold for peptide synthetases cleaving ATP to ADP: Glutathione synthetase and D-alanine:D-alanine ligase of *Escherichia coli*. Proc Natl Acad Sci USA 1995;92:1172–1176.
26. Hibi T, Nishioka T, Kato H, Tanizawa K, Fukui T, Katsube Y, Oda J. Structure of the multifunctional loops in the nonclassical ATP-binding fold of glutathione synthetase. Nat Struct Biol 1996;3:16–18.
27. Yamaguchi H, Kato H, Hata Y, Nishioka T, Kimura A, Oda J, Katsube Y. Three-dimensional structure of the glutathione synthetase from *Escherichia coli* at 2.0 Å resolution. J Mol Biol 1993;229:1083–1100.
28. Matsuda K, Mizuguchi K, Nishioka T, Kato H, Go N, Oda J. Crystal structure of glutathione synthetase at optimal pH: domain architecture and structural similarity with other proteins. Protein Eng 1996;9:1083–1092.
29. Wolodko WT, Fraser ME, James MNG, Bridger WA. The crystal structure of succinyl-CoA synthetase from *Escherichia coli* at 2.5 Å resolution. J Biol Chem 1994;269:10883–10890.
30. Waldrop GL, Rayment I, Holden HM. Three-dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase. Biochemistry 1994;33:10249–10256.
31. Herzberg O, Chen CC, Kapadia G, McGuire M, Carroll LJ, Noh SJ, Dunaway-Mariano D. Swiveling-domina mechanism for enzymatic phosphotransfer between remote reaction sites. Proc Natl Acad Sci USA 1996;93:2652–2657.
32. Thoden JB, Holden HM, Wesenberg G, Raushel FM, Rayment I. Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product. Biochemistry 1997;36:6305–6316.
33. Esser L, Wang C-R, Hosaka M, Smagula CS, Südhof TC, Deisenhofer J. Synapsin I is structurally similar to ATP-utilizing enzymes. EMBO J 1998;17:977–984.
34. Murzin AG. Structural classification of proteins: new superfamilies. Curr Opin Struct Biol 1996;6:386–394.
35. Bossemeyer D, Engh RA, Kinzel V, Ponstingl H, Huber R. Phosphotransferase and substrate binding mechanism of the cAMP-dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0 Å structure of the complex with Mn^{2+} adenylyl imidodiphosphate and inhibitor peptide PKI(5-24). EMBO J 1993;12:849–859.
36. Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science 1988;241:42–52.
37. Hanks SK, Quinn AM. Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. Methods Enzymol 1991;200:38–81.
38. Hubbard SR, Wei L, Ellis L, Hendrickson WA. Crystal structure of the tyrosine kinase domain of the human insulin receptor. Nature 1994;372:746–754.
39. Xu R-M, Carmel G, Sweet RM, Kuret J, Cheng X. Crystal structure of casein kinase-1, a phosphate-directed protein kinase. EMBO J 1995;14:1015–1023.
40. Denessiouk KA, Lehtonen JV, Korpela T, Johnson MS. Two "unrelated" families of ATP-dependent enzymes share extensive structural similarities about their cofactor binding sites. Protein Sci 1998;7:1136–1146.
41. Bernstein FC, Koetzle TF, Williams GJB, Meyer EJ Jr, Brice MD, Rodgers JK, Kennard O, Shimanouchi I, Tasumi M. The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol 1977;112:535–542.
42. Eriksson M, Uhlin U, Ramaswamy S, Ekberg M, Regnström K, Sjöberg B-M, Eklund H. Binding of allosteric effectors to ribonucleotide reductase protein R1: reduction of active-site cysteines promotes substrate binding. Structure 1997;5:1077–1092.
43. Denessiouk KA, Lehtonen JV, Johnson MS. Enzyme-mononucleotide interactions: three different folds share common structural elements for ATP recognition. Protein Sci 1998;7:1768–1771.
44. Stapleton MA, Javid-Majd F, Harmon MF, Hanks BA, Grahmann JL, Mullins LS, Raushel FM. Role of conserved residues within the carboxy phosphate domain of carbamoyl phosphate synthetase. Biochemistry 1996;35:14352–14361.
45. Galperin MY, Koonin EV. A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. Protein Sci 1997;6:2639–2643.
46. Dideberg O, Bertrand J. Tubulin tyrosine ligase: a shared fold with the glutathione synthetase ADP-forming family. Trends Biochem Sci 1998;23:57–58.
47. Sobolev V, Wade RC, Vriend G, Edelman M. Molecular docking using surface complementarity. Proteins: Struct Funct Genet 1996;25:120–129.
48. Teplyakov A, Sebastiao P, Oblomova G, Perrakis A, Brush GS, Bessman MJ, Wilson KS. Crystal structure of bacteriophage T4 deoxynucleotide kinase with its substrates dGMP and ATP. EMBO J 1996;15:3487–3497.
49. Zheng J, Knighton DR, Xuong N-H, Taylor SS, Sowadski JM, Ten Eyck LF. Crystal structures of the myristylated catalytic subunit of cAMP-dependent protein kinase reveal open and closed conformations. Protein Sci 1993;2:1559–1573.
50. Sicheri F, Moarefi I, Kuriyan J. Crystal structure of the Src family tyrosine kinase Hck. Nature 1997;385:602–609.
51. Hubbard SR. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. EMBO J 1997;16:5572–5581.
52. Jeffrey PD, Russo AA, Polyak K, Gibbs E, Hurwitz J, Massagué J, Pavletich NP. Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature 1995;376:313–320.
53. Russo AA, Jeffrey PD, Pavletich NP. Structural basis of cyclin-dependent kinase activation by phosphorylation. Nat Struct Biol 1996;3:696–700.
54. Schulze-Gahmen U, De Bondt HL, Kim SH. High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design. J Med Chem 1996;39:4540–4546.
55. Owen DJ, Noble MEM, Garman EF, Papageorgiou AC, Johnson LN. Two structures of the catalytic domain of phosphorylase

kinase: an active protein kinase complexed with substrate analogue and product. Structure 1995;3:467–482.
56. Lowe ED, Noble MEM, Skamnaki VT, Oikonomakos NG, Owen DJ, Johnson LN. The crystal structure of a phosphorylase kinase peptide substrate complex: kinase substrate recognition. EMBO J 1997;16:6646–6658.
57. Robinson MJ, Harkins PC, Zhang J, Baer R, Haycock JW, Cobb MH, Goldsmith EJ. Mutation of position 52 in ERK2 creates a nonproductive binding mode for adenosine 5′-triphosphate. Biochemistry 1996;35:5641–5646.
58. Hara T, Kato H, Katsube Y, Oda J. A pseudo-Michaelis quaternary complex in the reverse reaction of a ligase: structure of *Escherichia coli* B glutathione synthetase complexed with ADP, glutathione, and sulfate at 2.0 Å resolution. Biochemistry 1996;35: 11967–11974.
59. Thoden JB, Miran SG, Phillips JC, Howard AJ, Raushel FM, Holden HM. Carbamoyl phosphate synthetase: caught in the act of glutamine hydrolysis. Biochemistry 1998;37:8825–8831.
60. Levdikov VM, Barynin VV, Grebenko AI, Melik-Adamyan WR, Lamzin VS, Wilson KS. The structure of SAICAR synthetase: an enzyme in the de novo pathway of purine nucleotide biosynthesis. Structure 1997;6:363–376.
61. Huang W, Jia J, Gibson KJ, Taylor WS, Rendina AR, Schneider G, Lindqvist Y. Mechanism of an ATP-dependent carboxylase, dethiobiotin synthetase, based on crystallographic studies of complexes with substrates and a reaction intermediate. Biochemistry 1995; 34:10985–10995.
62. Käck H, Sandmark J, Gibson KJ, Schneider G, Lindqvist Y. Crystal structure of two quaternary complexes of dethiobiotin synthetase, enzyme-MgADP-AlF3-diaminopelargonic acid and enzyme-MgADP-dethiobiotin-phosphate; implications for catalysis. Protein Sci 1998;7:2560–2566.
63. Schindelin H, Kisker C, Schlessman JL, Howard JB, Rees DC. Structure of ADP · AlF_4^--stabilized nitrogenase complex and its implications for signal transduction. Nature 1997;387:370–376.
64. Barker WC, Garavelli JS, Haft DH, Hunt LT, Marzec CR, Orcutt BC, Srinivasarao G, Yeh L-SL, Ledley RS, Mewes H-W, Pfeiffer F, Tsugita A. The PIR-international protein sequence database. Nucleic Acids Res 1998;26:27–32.
65. Huang W, Lindqvist Y, Schneider G, Gibson KJ, Flint D, Lorimer G. Crystal structure of an ATP-dependent carboxylase, dethiobiotin synthetase, at 1.65 Å resolution. Structure 1994;2:407–414.
66. Georgiadis MM, Komiya H, Chakrabarti P, Woo D, Kornuc JJ, Rees DC. Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandi*. Science 1992;257:1653–1659.
67. Muegge I, Schweins T, Langen R, Warshel A. Electrostatic control of GTP and GDP binding in the oncoprotein $p21^{ras}$. Structure 1996;4:475–489.
68. Sprang SR, Withers SG, Goldsmith EJ, Fletterick RJ, Madsen NB. Structural basis for the activation of glycogen phosphorylase b by adenosine monophosphate. Science 1991;254:1367–1371.
69. Stevens RC, Gouaux JE, Lipscomb WN. Structural consequences of effector binding to the T state of aspartate carbamoyltransferase: crystal structures of the unligated and ATP- and CTP-complexed enzymes at 2.6 Å resolution. Biochemistry 1990;29: 7691–7701.
70. Gouaux JE, Stevens RC, Lipscomb WN. Crystal structures of aspartate carbamoyltransferase legated with phosphonoacetamide, malonate, and CTP or ATP at 2.8 Å resolution and neutral pH. Biochemistry 1990;29:7702–7715.
71. Rould MA, Perona JJ, Steitz TA. Structural basis of anticodon loop recognition by glutaminyl-tRNA synthetase. Nature 1991;352:213–218.
72. Perona JJ, Rould MA, Steitz TA. Structural basis for transfer RNA aminoacylation by *Escherichia coli* glutaminyl-tRNA synthetase. Biochemistry 1993;32:8758–8771.
73. Arnez JG, Steitz TA. Crystal structures of three misacylating mutants of *Escherichia coli* glutaminyl-tRNA synthetase complexed with tRNA (Gln) and ATP. Biochemistry 1996;35:14725–14733.
74. Belrhali H, Yaremchuk A, Tukalo M, Larsen K, Berthet-Colominas C, Leberman R, Beijer B, Sproat B, Als-Nielsen J, Grübel G, Legrand J-F, Lehmann M. Crystal structures at 2.5 Å resolution of seryl-tRNA synthetase complexed with two analogs of seryl adenylate. Science 1994;263:1432–1436.
75. Schmitt E, Moulinier L, Fujiwara S, Imanaka T, Thierry JC, Moras D. Crystal structure of aspartyl-tRNA synthetase from *Pyrococcus kodakaraensis* KOD: archaeon specificity and catalytic mechanism of adenylate formation. EMBO J 1998;17:5227–5237.
76. Cavarelli J, Eriani G, Rees B, Ruff M, Boeglin M, Mitschler A, Martin F, Gangloff J, Thierry JC, Moras D. The active site of yeast aspartyl-tRNA synthetase: structural and functional aspects of the aminoacylation reaction. EMBO J 1994;13:327–337.
77. Aberg A, Yaremchuk A, Tukalo M, Rasmussen B, Cusack S. Crystal structure analysis of the activation of histidine by *Thermus thermophilus* histidyl-tRNA synthetase. Biochemistry 1997; 36:3084–3094.
78. Arnez JG, Harris DC, Mitschler A, Rees B, Francklyn CS, Moras D. Crystal structure of histidyl-tRNA synthetase from *Escherichia coli* complexed with histidyl-adenylate. EMBO J 1995;14:4143–4155.
79. Arnez JG, Augustine JG, Moras D, Francklyn CS. The first step of aminoacylation at the atomic level in histidyl-tRNA synthetase. Proc Natl Acad Sci USA 1997;94:7144–7149.
80. Nakatsu T, Kato H, Oda J. Crystal structure of asparagine synthetase reveals a close evolutionary relationship to class II aminoacyl-tRNA synthetase. Nat Struct Biol 1998;5:15–19.
81. Roberts DL, Frerman FE, Kim J-JP. Three-dimensional structure of human electron transfer flavoprotein to 2.1 Å resolution. Proc Natl Acad Sci USA 1996;93:14355–14360.
82. Zarembinski TI, Hung LW, Mueller-Dieckmann HJ, Kim KK, Yokota H, Kim R, Kim SH. Structure-based assignment of the biochemical function of a hypothetical protein: a test case of structural genomics. Proc Natl Acad Sci USA 1998;95:15189–15193.
83. Tesmer JJG, Klem TJ, Deras ML, Davisson VJ, Smith JL. The crystal structure of GMP synthetase reveals a novel catalytic triad and is a structural paradigm for two enzyme families. Nat Struct Biol 1996;3:74–86.
84. Rizzi M, Nessi C, Mattevi A, Coda A, Bolognesi M, Galizzi A. Crystal structure of NH_3-dependent NAD^+ synthetase from *Bacillus subtilis*. EMBO J 1996;15:5125–5134.
85. Lim LW, Mathews FS, Steenkamp DJ. Identification of ADP in the iron-sulfur flavoprotein trimethylamine dehydrogenase. J Biol Chem 1988;263:3075–3078.
86. Vonrhein C, Bönisch H, Schäfer G, Schulz GE. The structure of a trimeric archaeal adenylate kinase. J Mol Biol 1998;282:167–179.
87. Li M, Dyda F, Benhar I, Pastan I, Davies DR. The crystal structure of *Pseudomonas aeruginosa* exotoxin domain III with nicotinamide and AMP: conformational differences with the intact exotoxin. Proc Natl Acad Sci USA 1995;92:9308–9312.
88. Li M, Dyda F, Benhar I, Pastan I, Davies DR. Crystal structure of the catalytic domain of *Pseudomonas* exotoxin A complexed with a nicotinamide adenine dinucleotide analog: implications for the activation process and for AMP ribosylation. Proc Natl Acad Sci USA 1996;93:6902–6906.
89. Holm L, Ouzounis C, Sander C, Tuparev G, Vriend G. A database of protein structure families with common folding motifs. Protein Sci 1992;1:1691–1698.
90. Holm L, Sander C. Protein folds and families: sequence and structure alignments. Nucleic Acids Res 1999;27:244–247.
91. Holm L, Sander C. Dictionary of recurrent domains in protein structures. Proteins 1998;33:88–96.
92. Tamada T, Kitadokoro K, Higuchi Y, Inaka K, Yasui A, de Ruiter PE, Eker AP, Miki K. Crystal structure of DNA photolyase from *Anacystis nidulans*. Nat Struct Biol 1997;4:887–891.
93. Ingelman M, Bianchi V, Eklund H. The three-dimensional structure of flavodoxin reductase from *Escherichia coli* at 1.7 Å resolution. J Mol Biol 1997;268:147–157.
94. Kraulis PJ. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr 1991;24: 945–949.
95. Merritt EA, Bacon DJ. Raster3D: photorealistic molecular graphics. Methods Enzymol 1997;277:505–524.
96. Müller CW, Schulz GE. Structure of the complex between adenylate kinase from *Escherichia coli* and the inhibitor Ap_5A refined at 1.9 Å resolution. J Mol Biol 1992;224:159–177.