Quarterly Reviews of Biophysics **32**, 1 (1999), pp. 1–56. Printed in the United Kingdom
© 1999 Cambridge University Press

Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer

Ingrid R. Vetter and Alfred Wittinghofer

MPI für Molekulare Physiologie, Abteilung Strukturelle Biologie, Otto-Hahn-Strasse 11, 44227 Dortmund, Germany

1. Introduction 2

2. Reaction catalyzed: ATP → ADP + P_i/GTP → GDP + P_i 5
- 2.1 P-loop containing nucleotide triphosphate hydrolases with Ras-like fold 6
- 2.2 Ras-like fold with or without modified P-loop 11
- 2.3 Protein kinases 12
- 2.4 NDP kinases 15
- 2.5 Transport ATPases 16
- 2.6 DNA topoisomerases 17
- 2.7 Chaperones and chaperonins 18
- 2.8 Actin-like ATPase domain 20
- 2.9 Phosphoglycerate kinase 20
- 2.10 Pyruvate kinase 23
- 2.11 Phosphofructokinase 23
- 2.12 Creatine kinase and glutamine synthetase 24
- 2.13 The ATP grasp fold 25

3. Reaction catalyzed: ATP → AMP + PP_i 27
- 3.1 DNA-ligases, mRNA capping enzymes 28
- 3.2 Class II aminoacyl tRNA synthetases and biotin synthetases 29
- 3.3 ATP pyrophosphatase fold 30
- 3.4 Luciferases 32
- 3.5 Nucleic acid polymerases 33
- 3.6 MutT proteins, PP_i releasing NTPases 36
- 3.7 dUTPase 37
- 3.8 Apyrases 37

4. Reaction catalyzed: ATP → Adenosine + PP_i + P_i 38
- 4.1 SAM synthetases 38

5. Reaction catalyzed: ApppA → AMP + ADP 38
- 5.1 HIT-like proteins 38

6. General features 39

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

7. Outlook 41

8. Acknowledgements 43

9. References 43

1. Introduction

Nucleoside triphosphates are crucial mediators of life. The high energy phosphoanhydride bond of (usually) ATP is used to drive unfavorable chemical reactions, to fuel biological machines, and to regulate a vast number of processes via phosphorylation of proteins. GTP, in turn, is used almost exclusively for the regulation of signal transduction and transport processes, whereas the other nucleotides play a less important role, except in synthesis pathways involving sugars (UTP) and phospholipids (CTP) and as building blocks of polynucleotides such as RNA and DNA. Proteins that bind and use these nucleotides for enzymatic reaction and regulation are very diverse. Although some of them constitute the largest protein superfamilies known (e.g. protein kinases), others seem to be far less conserved in evolution.

The different classes of nucleotide binding proteins have been reviewed by (Schulz, 1992). At this time only four major families of nucleotide binding proteins could be indentified: the classical mononucleotide binding proteins, the dinucleotide binding proteins, the protein kinases, and the actin family. Representatives with known structures of other families of nucleotide binding proteins were not numerous enough at this time to be grouped.

The definition of the classical mononucleotide binding fold as used by Schulz (1992) was characterized by the finger print GXXGXGK which was initially mistaken for the similar finger print of dinucleotide binding proteins GXGXXG (Wierenga & Hol, 1983). The family of mononucleotide binding proteins at this time comprised adenylate kinase, guanylate kinase, Ras and EF-Tu. The similarity in folds and nucleotide binding made it tempting to assume a divergent evolution. Furthermore, the family was thought to include myosin and F_1-ATPase which turned out to be correct after their structures were solved (Rayment *et al*. 1993; Abrahams *et al*. 1994). Other candidates were the cystic fibrosis gene product, heat-shock proteins (e.g. Hsp104 (Parsell *et al*. 1991), which is the yeast homologue of Clp), protease La (Chin *et al*. 1988), Clp ATPase (Gottesman *et al*. 1990), a large group of DNA recombination, repair and replication proteins (Gorbalenya & Koonin, 1990) and NTRC, a regulatory protein of nitrogen metabolism (Austin *et al*. 1991). The structures of these proteins (except a homologue of the cystic fibrosis gene product (Hung *et al*. 1998), see paragraph on ABC-transporters) are still unknown to date. Additionally, it was observed that CheY (Volz & Matsumura, 1991), $\gamma\delta$-resolvase (Yang & Steitz, 1995) and flavodoxin (Watenpaugh *et al*. 1972) did fit the scheme with their chain folds but lack the finger print (Schulz, 1992).

The classical dinucleotide-binding chain fold including proteins that bind Nicotinamide Adenine Dinucleotide (NAD) was described by Schulz (1992) to be quite different from the mononucleotide binding fold. The dinucleotide binding proteins are usually based on Rossman-fold type domains, which are supersecondary structural elements comprising three parallel β-strands connected by α-helices, β-strands or irregular loops (Kutzenko *et al*. 1998). They are not subject of this review and therefore not discussed further.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Table 1 *Classification of nucleotide binding proteins*

$ATP \rightarrow ADP + P_i$/
$GTP \rightarrow GDP + P_i$

P-loop-containing NTP hydrolases, definition: [AG]xxxxGK[ST] (here including GxxxxGKG and GxxxxGKD)

G-proteins:	small GTPases (Ras, Rap, Ran, Arf…)	GTP binding
	EF-Tu, EF-G	
	G protein α-subunit	
	FtsY	
	Ffh	
kinases:	guanylate kinase, thymidine kinase, thymidylate kinase, adenylate kinase, uridylate kinase, UMP/CMP kinase, deoxynucleotide monophosphate kinase, 6-P-fructo-2-kinase/fructose-2,6-bisphosphatase (kinase domain)	ATP binding
motor proteins:	myosin/kinesin heavy chains	
RecA-family:	Rec A	
	F_1-ATPase	
	ATP-dependent helicases	
	AAA-type ATPases	
	ABC transporters	
nitrogenase:	nitrogenase iron protein	

Ras-fold, P-loop related
adenylosuccinate synthetase
dethiobiotin synthetase
tubulin
FtsZ

P-loop but no Ras fold
PEP-carboxykinase

Protein kinases
protein kinases

NDP kinases
NDP kinase
aspartate carbamoyltransferase allosteric domain

E1-E2 ATPases
Na^+, K^+-ATPase
H^+,K^+-ATPase
Ca^{2+}-ATPase

DNA topoisomerases
Type I
Type II
Hsp90 (N-terminal domain)

GroEL
GroEL

Actin-like ATPase domain
actin
heat-shock cognate protein (HSC70, ATPase domain)
DnaK (ATPase domain)
hexokinase
glycerol kinase

Phosphoglycerate kinase
phosphoglycerate kinase

Pyruvate kinase
Pyruvate kinase

Phosphofructokinase
phosphofructokinase

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Table 1 (*cont.*)

Glutamine synthetase
creatine kinase
glutamine synthetase

ATP-grasp fold
(a) ***Glutathione-synthetase ATP-binding domain like***
glutathione synthetase
D-Ala:D-Ala ligase
biotin carboxylase
synapsin
carbamoyl phosphate synthetase
succinyl-CoA synthetase

$ATP \rightarrow AMP + PP_i$

pyruvate phosphate dikinase (N-terminal domain)
(b) ***DNA ligases, mRNA capping enzymes***
T7 DNA ligase
mRNA capping enzyme

tRNA synthetases class II
tRNA synthetases class II
biotin-protein ligase
asparagine synthetase

ATP pyrophosphatase fold
(a) adenine nucleotide α-hydrolases
N-type ATP pyrophosphatases:
GMP synthetase
NAD synthetase
asparagine synthetase
argininosuccinate synthetase
PAPS reductase:
PAPS reductases
ATP sulfurylases

(b) tRNA synthetases and cytidylyltransferases
tRNA synthetases class I
glycerol-3-phosphate cytidylyltransferase

Luciferases
firefly luciferase

Nucleic acid polymerases
DNA-directed DNA polymerases (DNAPs)
reverse transcriptases (RTs)
RNA-directed RNA polymerases (RNAPs)
DNA-directed RNAPs
DNA polymerase β
kanamycin nucleotidyltransferase] nucleotidyltransferase superfamily
adenylylcyclase

MutT proteins and other PPi releasing NTPases
MutT

dUTPase
dUTPase

Apyrases

$ATP \rightarrow adenosine + PP_i + P_i$

SAM synthetases
SAM synthetase

$ApppA \rightarrow AMP + ADP$

HIT-like proteins
HINT

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Table 2. *Reaction mechanisms*

Transfer of γ-P, no intermediate:	PEP-carboxykinase, all kinases except NDP kinase
Transfer/Release of γ-P:	G-proteins, motor proteins, RecA family, nitrogenase, tubulin, chaperones, actin, topoisomerases, adenylosuccinate synthetase*, dethiobiotin-synthetase*, glutamine synthetase*, glutathione synthetase*, carbamoyl phosphate synthetase*
Transf./Rel. of γ-P via P-enzyme:	NDP kinase, E1-E2 ATPases, succinyl-CoA-synthetase
Release of β,γ-PP:	MutT, apyrases, dUTPase
Tansf./Rel. of β,γ-PP via PP-enz:	pyruvate phosphate dikinase
Release of α,β,γ-PPP:	SAM synthetase
Cleavage of ApppA:	HIT proteins
Transfer of AMP to acceptor:	aminoacyl-tRNA synthetases, polymerases
AMP-substrate intermediate:	GMP/NAD synthetase*, luciferase*
Transfer of NMP via NMP-enzyme:	DNA ligases, capping enzymes

* With covalent non-enzyme intermediate.

The third class of nucleotide binding proteins mentioned by Schulz were the protein kinases which have the finger print GXGXXG (Hanks *et al*. 1988). This corresponds to the dinucleotide binding finger print, but is structurally more similar to the mononucleotide-binding glycine-rich loop. Protein kinases and the mononucleotide binding proteins were thought to be related by convergent evolution. The fourth family, the actin-related proteins, is again quite different in fold and nucleotide binding motifs, as will be discussed later (Schulz, 1992).

Since 1992, considerable progress has been made towards identifying common sequence motifs and solving the structure of many of these nucleotide triphosphate-binding proteins, which in turn helped to identify additional common folds. In the following we review the available knowledge (up to 1998) thereby focusing on common structural (Table 1) and mechanistic (Table 2) relationships.

2. Reaction catalyzed: ATP → ADP + Pi/GTP → GDP + P_i

The energy stored in the phosphoanhydride bond can be used in different ways. First, it can be transferred to other substrates, for example, in biosynthetic pathways involving mostly ATP, where the energy is used to drive the reaction into an otherwise energetically unfavorable direction. Secondly, the energy of ATP or GTP can be used to induce conformational changes in the protein, where the nature of the bound nucleotide facilitates reactions such as the polymerization of actin or tubulin, signal transduction via GTP-binding proteins or pumping of ions in transport ATPases. Protein kinases use ATP for the activation/inactivation of other proteins by attaching a phosphoryl group. Other than that, the members of enzymes that cleave the β–γ bond of ATP or GTP have little in common and are therefore discussed in separate sections, each covering a group of proteins with related folds.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

2.1 *P-loop containing nucleotide triphosphate hydrolases with Ras-like fold*

One of the most frequent sequence motifs in nucleotide binding proteins is the P-loop (Saraste *et al*. 1990), which was originally identified by Walker and coworkers in the α- and β-subunits of the F_1-ATPase and correspondingly called the Walker A motif (Walker *et al*. 1982). Its consensus sequence is G/AxxxxGK(x)S/T as defined in the PROSITE database (Bairoch *et al*. 1997), and often contains additional glycines at the position of the four variable (x) residues. This definition slightly extends the definition used by G. Schulz (Schulz, 1992) to include other clearly related families, e.g. the α-subunits of heterotrimeric G-proteins. Thus, we also include GxxxxGKG and GxxxxGKD motifs to encompass the obviously related folds of adenylate kinase, uridylate kinase, UMP/CMP-kinase and T4 deoxynucleotide kinase (DNK). Another motif identified by Walker *et al*. called the Walker B motif, is defined less clearly and much more variable. Most commonly it is a conserved aspartic acid, located in different sequence motifs, but situated in a homologous position in the nucleotide binding site and involved in binding the divalent ion crucial for catalysis. The structure of the Walker A motif has first been solved in adenylate kinase, where it has been found to form a loop between the N-terminal β-strand and an α-helix (Schulz *et al*. 1974). This loop is often stabilized by the conserved lysine forming hydrogen bonds to the main chain carbonyls of the loop, and a sulfate ion from the crystallization medium was observed in the center of the loop (Dreusicke *et al*. 1988). Later it was found that the sulfate ion mimics the β-phosphate of the substrate ATP, hence the name P-loop (Saraste *et al*. 1990).

In an effort to find universal binding motifs for all nucleotide binding proteins, kinase 1, 2, 3 motifs were defined which are believed to be involved in binding of the phosphates, the metal ion or the purine base/ribose region (Traut, 1994), respectively. However, the kinase regions do not seem to be conserved between the subfamilies and too many exceptions have to be made to see common sequences. As mentioned above for the Walker B (or kinase 2) motif, it might not be very helpful to try too hard to align such very faint homologies unless structural or functional reasons can be put forward for their conservation. This is not the case if, for example, the ATP binding sites of adenylate kinase (Müller & Schulz, 1992) and actin (Kabsch *et al*. 1990) are compared.

Besides adenylate kinase (Schulz *et al*. 1974), guanylate kinase (Stehle & Schulz, 1992), Ras (Milburn *et al*. 1990; Pai *et al*. 1990) and EF-Tu (Jurnak, 1985; LaCour *et al*. 1985; Berchtold *et al*. 1993), the structure of the P-loop has been observed in many other ATP-binding proteins such as recA (Story *et al*. 1992; Story & Steitz, 1992), the F_1-ATPase (Abrahams *et al*. 1994), ATP-dependent helicases (Subramanya *et al*. 1996a; Yao *et al*. 1997; Korolev *et al*. 1997), AAA-type ATPases (Lenzen *et al*. 1998; Yu *et al*. 1998), ABC transporters (Hung *et al*. 1998), nitrogenase iron protein (Georgiadis *et a*l. 1992; Schindelin *et al*. 1997), myosin (Rayment *et al*. 1993), kinesin (Kull *et al*. 1996; Sack *et al*. 1997), the kinase domain of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (Hirokawa, 1998), T4 deoxynucleotide kinase (DNK) (Teplyakov *et al*. 1996) and in other monophosphate kinases such as uridylate kinase (Müller-Dieckmann & Schulz, 1994), UMP/CMP kinase (Scheffzek *et al*. 1996a), thymidylate kinase from *S. cerevisiae* (Lavie *et al*. 1997) and in thymidine kinase (Wild *et al*. 1995; Brown *et al*. 1995), and the conclusion can now be safely drawn that it always connects a strand of the central β-sheet with an α-helix (Fig. 1*a–f*). The family of P-loop containing GTP-binding proteins with known structures includes – besides Ras and EF-Tu – EF-G (Al-Karadaghi *et al*. 1996), the α-subunits of transducin and G_i (Noel

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

et al. 1993; Coleman *et al*. 1994) and FtsY and Ffh (Montoya *et al*. 1997; Freymann *et al*. 1997). These, together with the Ras-related proteins (e.g. Rap, Ral, Ran, Arf…), constitute the P-loop family with the largest number of members with known three-dimensional structures. In these GTP-binding proteins, the P-loop wraps around the β-phosphate of GTP in a similar way as in the complexes of adenylate kinase with substrate analogues such as AppNHp and Ap_5A (Müller & Schulz, 1992), or in the UMP kinase structures with its substrates or with the Up_5A bisubstrate analogue (Müller-Dieckmann & Schulz, 1994; Scheffzek *et al*. 1996b). The structures of GTP-binding proteins show that the hydroxyl group of the conserved Ser/Thr of the P-loop is a ligand of the metal ion, which forms a bidentate coordination with the β,γ phosphates of GTP. This conserved hydroxyl group is somewhat different in some of the ATP utilizing enzymes because it does not directly follow the lysine residue in sequence and its role in metal ion binding is not firmly established.

A subfamily of the P-loop containing nucleotide triphosphate (NTP) hydrolases with central importance for the cell is the P-loop containing motor proteins. The best-known examples are myosin (Rayment *et al*. 1993) and kinesin (Kull *et al*. 1996; Sack *et al*. 1997) as well as the world's smallest rotary motor, the F_1-ATPase (Block, 1997; Abrahams *et al*. 1994). In contrast to the rotational catalysis of the F_1-ATPase (see below), myosin, and kinesin, which have similar structural elements despite their low sequence identity, feature a more conventional motor function involving some kind of conformational change coupled to the nucleotide hydrolysis (Howard, 1997). The conformation of the P-loop in myosin (Fig. 1*d*) is very similar to the P-loop of G-proteins whereas the position of the other structural elements contacting the nucleotide is rather different. Although the motor protein dynein contains several P-loop motifs as well, it shows no sequence homology to myosin or kinesin and its domain organization seems to be very different (Hirokawa, 1998).

The F_1-ATPase is made up of nine protein subunits and uses the protonmotive force as the source of energy for ATP synthesis. The catalytic domains form a hexagon of alternating α- and β-subunits surrounding a drive shaft made up of the γ-subunit (Abrahams *et al*. 1994). In the backward reaction, hydrolysis of ATP is coupled to the position of the shaft, i.e. each nucleotide is hydrolyzed in turn, three ATPs per revolution. The nucleotide binding sites are located at the interfaces of the α- and β-subunits, but only three of the six nucleotides are turned over. The nucleotide state of the catalytic centers together with the position of the γ-subunit determines the conformation of the associated subunit. Apart from the P-loop, the nucleotide binding sites of the F_1-ATPase show no similarity to, for example, the site in adenylate kinase, and the overall structure is quite different as well (Fig. 1*b*, *e*). On the other hand, the central domain of F_1-ATPase shows similarity to the ATPase domain of RecA (Story *et al*. 1992), thereby constituting the RecA subfamily of the P-loop containing nuceotide triphosphate hydrolases (Murzin *et al*. 1995). The same topology as RecA and F_1-ATPase has been predicted for Lon Proteases, FtsH, NSF (see below), NifH and ABC transporters (Yoshida & Amano, 1995; see below).

Another group of proteins with an established relationship to the RecA family is that of the ATP-dependent DNA helicases. These – usually multimeric – proteins play essential roles in replication, repair, recombination and transcription of DNA (West, 1996). Their structure and function seems to be very different from topoisomerases. Two classes of helicases can be defined mechanistically: those that require a 3′ flanking strand of single-stranded DNA (3′-5′ helicases) and those that require a 5′ flanking strand (5′-3′ helicases) (Bird *et al*. 1998). Based on sequence homology, five groups are known. Superfamilies I and II, which comprise most

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

3′-5′ helicases, contain seven conserved motifs and are considered to be dimeric. Superfamily III contains helicases from small DNA and RNA viruses, whereas superfamilies IV and V are hexameric 5′-3′ helicases related either to *E. coli* replicative helicase DnaB (superfamily IV) or the bacterial transcription termination factor Rho (Bird *et al.* 1998).

In superfamily I and II helicases, two of the seven conserved motifs correspond to the Walker A and B motifs involved in nucleotide binding. Most helicases function as oligomers (either dimers or hexamers) and are often integral parts of larger protein complexes, providing a molecular motor function to move the complex along the DNA while unwinding it. One of the best studied helicases is *E. coli* DnaB, which is the primary replicative helicase in *E. coli* and unwinds duplex DNA ahead of the replication fork. No high resolution structure is available for hexameric helicases, but many low-resolution studies have been performed (Egelman, 1996). Electron micrographs show DnaB as a hexameric ring composed of identical subunits, each containing an NTP binding site (usually ATP). It is assumed that the DNA passes through the ring (West, 1996). In contrast, three high-resolution structures of 3′-5′ helicases are available: PcrA, which is a DEXX box DNA helicase from *Bacillus stearothermophilus* (Subramanya *et al.* 1996a), a RNA helicase domain of the hepatitis C virus NS3 (Yao *et al.* 1997) and the *E. coli* Rep helicase bound to ss-DNA and ADP (Korolev *et al.* 1997). PcrA consists of two domains each comprising two subdomains. The fold of the first subdomains of each domain shows unexpected structural homology with each other and with the ATP-binding domain of the *E. coli* RecA recombination protein (Story *et al.* 1992; Story & Steitz, 1992). The two other subdomains are primarily α-helical and have no structural homology. The nucleotide binding site is located in a pocket between the two domains. In contrast to PcrA, the RNA helicase domain of the hepatitis C virus NS3 consists of three domains, two of which are similar in fold to the RecA-domains of PcrA. However, the connectivity of these domains does not resemble RecA and is more closely related to adenylate kinase. The fold of the third domain is unique. Finally, the Rep helicase of *E. coli* has the same overall fold as PcrA, but with a different relative orientation of the subdomains resulting in a closure of the large cleft found in PcrA. The basic structural motif of DNA helicases from superfamilies I and II seem to be two RecA-like domains which can be supplied by two different protein subunits (Bird *et al.* 1998). Indeed, two separate proteins have been found which are homologous to either the N-terminal domain of subfamily I helicases (*E. coli* PhoH, which binds nucleotides) or the C-terminal domain (*E. coli* YjhR). This suggests that these two proteins might have evolved from helicases (Koonin & Rudd, 1996).

Another family of P-loop containing proteins which are related to the RecA fold are the AAA proteins (**A**TPases **a**ssociated with different cellular **a**ctivities, an expression coined by Kunau *et al.* (1993). AAA proteins constitute a large family of ATPases associated with cellular housekeeping, cell division and differentiation. They have been identified in prokaryotes and eukaryotes. The family is characterized by a 230 residue domain that contains Walker homology motifs and a specific AAA motif (Patel & Latterich, 1998); some more distantly related members of the family such as the heat shock protein ClpB contain even two P-loop/Walker A motifs (Wawrzynow *et al.* 1996; Walker *et al.* 1982). The function of this group of proteins remains elusive, but it has been proposed that these proteins might act as ATP-dependent protein clamps. Bacteria and mitochondria share three families of ATP-dependent proteases: protease La (or Lon), the two-component Clp (or Ti) and FtsH (or HflB). Lon and Clp are soluble proteins, whereas the ATP-dependent metalloprotease FtsH is anchored in the cytoplasmatic membrane, spanning it twice. Lon, Clp and FtsH mediate

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

the degradation of a variety of regulatory proteins that control the heat shock response, the SOS response to DNA damage, capsular polysaccharide synthesis, the lytic–lysogenic cycle of bacteriophage λ, and programmed bacterial cell death (Suzuki *et al*. 1997). Electron micrographs show that some AAA proteins form doughnut-shaped homohexamers while others function as monomers, dimers or trimers (Patel & Latterich, 1998). The regulatory subunits of proteasomes have at least 15 subunits, six of which are AAA proteins (Patel & Latterich, 1998). Recently, the structure of an AAA domain of one of the best-characterized AAA proteins, the N-ethylmaleimide-sensitive factor (NSF), has been solved in complex with ATP (Lenzen *et al*. 1998; Yu *et al*. 1998). NSF is involved in vesicle fusion events occuring in, for example, secretory pathways including those involving synaptic vesicles (Rothman, 1994). It is a trimeric or homohexameric cytosolic protein which can bind to membranes via NSF attachment protein. ATP hydrolysis in the AAA domain of NSF primes the membranes for fusion. The structure of the C-terminal AAA domain of NSF revealed a nucleotide-binding domain which is similar to RecA. Additionally, it contains a smaller domain which is completely α-helical and less well conserved. The crystallized NSF fragment is a hexamer which is stabilized by loops extending from the ATP-binding domain towards the neighboring monomer. AAA proteins contain variants of the DEAD-box motif, a modified Walker B motif first identified in RNA and DNA helicases (Lindner *et al*. 1989; Wassarman & Steitz, 1991) which was proposed to contain the glutamate as catalytic base (Subramanya *et al*. 1996a; Korolev *et al*. 1997). This motif has the sequence DDIE in the crystallized domain of NSF, and the second aspartate is indeed close to the active site. It is involved in coordinating two water molecules contacting the γ-phosphate and the magnesium. ATP hydrolysis has been proposed to result in a conformational change. This in turn should trigger the disassembly of the membrane-bound complex of NSF, NSF attachment protein and v/t-SNAREs, leading to the fusion of the vesicle.

ABC transporters (**A**TP **B**inding **C**assette) contain a P-loop and may constitute the largest family of paralogous membrane proteins (Cleves & Kelly, 1996). The *E. coli* genome alone codes for about 80 ABC proteins (Linton & Higgins, 1998). This family is essential for the transport of ions, sugars, peptides and proteins. In higher organisms some members of the family are involved in pathological processes such as cystic fibrosis, where the responsible gene is called CFTR-cystic fibrosis transmembrane conductance regulator (Riordan *et al*. 1989), or the multi-drug resistance mdr (Germann *et al*. 1996), which is selectively overexpressed in cancer cells and responsible for resistance against anti-cancer drugs. ABC transporters consist of two membrane spanning domains, which probably constitute the channel, and two nucleotide binding domains with P-loop motifs (Linton & Higgins, 1998; Tampé, 1998). These four functional units can be encoded by one, two or four different polypeptide chains. Characteristic for these highly conserved nucleotide binding domains is a short sequence between the Walker A and B motifs ((L)SGG(Q), called the 'C-Loop') which is essential for function. Very recently, the structure of the ATP-binding subunit of the histidine permease of *Salmonella typhimurium* was solved (Hung *et al*. 1998). The overall fold is different from other known proteins. However, there are limited similarities of parts of the structure with RecA and the α- and β-subunits of the F_1-ATPase. The P-loop is at the same position relative to the nucleotide as in Ras or adenylate kinase, for example, and there is a conserved aspartate which probably interacts with the divalent cation during ATP hydrolysis (there is no magnesium ion in the structure; magnesium was omitted because it would have induced ATP hydrolysis) (Hung *et al*. 1998). In contrast to Ras or adenylate

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

kinase, the ATP-binding subunit of histidine permease is a dimer *in vivo* and needs to be a dimer to be active.

In summary, except from the P-loop itself the rest of the structure of P-loop containing proteins is rather variable. The only common feature of these proteins is the presence of an α,β-structural core of several β-strands alternating with α-helices. It has been argued by many authors that this α,β-core represents a structural element conserved in evolution for the binding of nucleotides which has later experienced many different alterations. This motif has been called the 'ancestral core structure' (Milner-White *et al.* 1991) or 'classical mononucleotide binding fold' (Schulz, 1992). Comparing all proteins that contain P-loops, and including in the comparison such diverse examples as nitrogenase and the α- and β-subunits of the F_1-ATPase, it becomes increasingly difficult to justify the statement that all the P-loop proteins share common topological elements, although the assumption that the P-loop proteins developed from a common ancestor may still hold.

A striking example for a protein with an unique fold but containing a P-loop is the PEP-carboxykinase (PCK), which exists in an ATP and a GTP-dependent variant. The ATP dependent enzymes are found in bacteria, yeast, trypanomastid parasites and plants, the GTP dependent enzymes occur in a variety of eukaryotes including mammals (Matte *et al.* 1997). The structure of the ATP dependent *E. coli* PCK (Matte *et al.* 1996, 1997; Tari *et al.* 1996) is bilobal, the C-terminal domain binds the mononucleotide. The fold of this domain seems to be unique in spite of having a classical Walker A motif that connects a β-strand with an α-helix (Fig. 1*g*). In contrast to the classical mononucleotide binding fold, which has a central parallel β-sheet flanked by α-helices, PCK has a mixed β-sheet with a $\beta2$-α-$\beta4$-α arrangement and thus features a different ordering of β-strands compared to any of the known structural families.

The variety of the architectures found in known P-loop proteins leads to the expectation that more different structures are bound to be determined.

In the same way that different topologies are observed for different P-loop proteins, the mechanism of phosphoryl transfer and the nature of the reaction catalyzed varies considerably between the different proteins. The nucleoside monophosphate kinases (NMP kinases) such as adenylate or uridylate kinase catalyze the direct transfer of phosphate from ATP to AMP (adenylate kinase) or UMP (uridylate kinase), respectively (Schlichting & Reinstein, 1997). Structures of NMP kinases with bi-substrate analogues – considered analogues mimicking the transition state of the reaction by some authors (Müller & Schulz, 1992; Scheffzek *et al.* 1996a; Abele & Schulz, 1995) – have a number of arginine residues, in addition to the lysine of the P-loop, that stabilize the large number of negative charges. It has been shown that these arginines play an important role in catalysis (Yan *et al.* 1990; Tsai & Yan, 1991). The evolution of the catalytic machinery can be shown nicely by the structure of T4 deoxynucleotide kinase (DNK) (Teplyakov *et al.* 1996) in relation to that of other NMP kinases. The core ATP binding domain of DNK resembles that of adenylate (AK), guanylate (GK) or uridylate kinase (UK), but it contains an additional GMP binding domain of equal size that has no counterpart in the others. It turns out that all the functional residues of AK/GK/UK are found in DNK and are in a position to function as catalytic residues after a large conformational change between the two domains (Teplyakov *at al.* 1996).

It is interesting to note that the kinase domain of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase, which has been expected to resemble PFK (6-phosphofructo-1-kinase), is related to the NMP kinases and the catalytic domain of G

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

proteins (Fig. 1*c*, *n*). This striking homology lead Hasemann *et al*. (1996) to suggest a potentially similar mechanism of 6-phosphofructo-2-kinase with a stabilization of the transition state as in G proteins. This would be different from the mechanism of PFK and a large group of nucleotide triphosphate phosphotransferases, in which a nucleophilic activation of the substrate is required for catalysis (in PFK, the fructose-6-phosphate hydroxyl group is activated) (Knowles, 1980; Yoshida & Amano, 1995).

No residue acting as a catalytic base to activate the nucleophilic water has been found in Ras-like GTP-binding proteins, and it has been proposed that the γ-phosphate itself acts as a base in what has been called substrate-assisted catalysis (Schweins *et al*. 1995, 1996). Most Ras-like proteins have been shown to have a slow catalytic rate because a second component (GAP, for **G**TPase-**A**ctivating **P**rotein) supplying an arginine residue into the active site has to be added in trans to complement the active site and to speed up the reaction (Ahmadian *et al*. 1997; Scheffzek *et al*. 1996b, 1997). An arginine and a crucial glutamine residue are both needed in the transition state of Ras-related and heterotrimeric GTP-binding proteins (Coleman *et al*. 1994; Scheffzek *et al*. 1996b, 1997; Sondek *et al*. 1994; Rittinger *et al*. 1997a, 1997b). In the catalytically active β-subunit of the F_1-ATPase a glutamic acid sidechain is thought to activate a water molecule next to the γ-phosphate, thus promoting an inline nucleophilic attack on this phosphate. The developing charges are stabilized by an arginine residue from the α-subunit. (Abrahams *et al*. 1994). In myosin, which has a very fast single turnover reaction with an equilibrium constant on the enzyme of 1, there is neither an arginine nor a glutamine in the active site of the transition state mimic structure (Fisher *et al*. 1995; Smith & Rayment, 1996), and it is not clear how the fast ATP hydrolysis reaction rate is achieved. In nitrogenase, a homodimeric enzyme, catalysis is believed to come about by an aspartic acid acting as a possible general base and a lysine as a stabilizer, both supplied to the active site of one subunit by the neighboring subunit (Schindelin *et al*. 1997). These few examples should suffice as an illustration that the similar mode of binding of ATP/GTP by the P-loop proteins does not imply a similar reaction mechanism.

2.2 *Ras-like fold with or without modified P-loop*

The discussion about a common evolutionary relationship between P-loop proteins gets even more complicated by the finding that the structures of some mononucleotide binding proteins such as dethiobiotin synthetase have a Ras-like fold without any sequence similarity. The 'P-loop' of this enzyme contains an extra residue between the first and second invariant glycine (Huang *et al*. 1994, 1995).

A particularly interesting example is adenylosuccinate synthetase (ASS) (Silva *et al*. 1995; Poland *et al*. 1993) which catalyzes the first step in the synthesis of AMP, using IMP, aspartate and GTP as substrates. Like other P-loop proteins, ASS has the GxxxxGK motif (although without a Ser/Thr), which forms a loop between the first β-strand and helix, and if aligned structurally with the corresponding loop in Ras, the substrate would bind in a similar way. However, the rest of the β-sheet and helices are rather different from Ras. This led Honzatko and coworkers to propose that the two proteins, despite using GTP to transfer phosphate, are examples of converging evolution (Poland *et al*. 1993). Even though the enzyme contains a sequence closely resembling a P-loop motif, structural alignment of the GppNHp complex of ASS with that of various GTP-binding proteins shows that the binding of the phosphate moiety of the nucleotide differs very much, with a small number of interactions in ASS as

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

compared to Ras, Gα and EF-Tu (Poland *et al*. 1996). However, the close relationship of ASS to the P-loop proteins is emphasized by the fact that the ASS fold is closely related to the P-loop containing nitrogenase.

Tubulin is a GTP-binding protein whose sequence has very little resemblance to that of the Ras-like GTP-binding proteins, although numerous attempts for such an alignment requiring reversal of some of the sequence motifs have been made. Recently the structures of the tubulin α,β-dimer and the prokaryotic tubulin homologue FtsZ have been determined (Nogales *et al*. 1998a; Loewe & Amos, 1998). The structures showed the molecule to consist of two domains, one of which roughly resembles the Ras-like G domain of alternating β-strands and helices. However, the nucleotide binding site of tubulin and FtsZ resembles more that of glyceraldehyde-3-phosphate dehydrogenase, a dinucleotide-binding protein (Nogales *et al*. 1998b). The GDP is coordinated differently than by conventional GTP-binding proteins. As in many other nucleotide binding proteins the phosphate residues of the nucleotide are contacted by glycine-rich sequences, one of which is the tubulin signature sequence GGGTGTG, as predicted (Schweins & Wittinghofer, 1994). The guanine base is sandwiched between hydrophobic residues and H-bonded to an aspartic acid, reminiscent of the situation in Ras-like proteins. So far nothing is known about the mechanism of GTP hydrolysis in these non-conventional GTP-binding proteins, and no catalytic base was identified.

A similarity to the Ras fold has also been proposed for a part of the γδ-resolvase structure (Schulz, 1992). However, γδ-resolvase has no P-loop and the detailed fold differs significantly (Yang & Steitz, 1995), so the resemblance might have no evolutionary significance.

2.3 *Protein kinases*

The largest class of nucleotide binding proteins, or possibly of any protein, is probably that of the protein kinases. The decade-old estimation of 1001 kinases (Hunter, 1987) constituting roughly 1% of the mammalian genome turned out to be an underestimation after the first eukaryotic genome has been sequenced and more and more human EST (**E**xpressed **S**equence **T**ags) become available. The budding yeast (*S. cerevisiae*) genome – which has been completely sequenced recently – has 113 conventional protein kinase genes, corresponding to ~ 2 % of the total number of genes (Hunter & Plowman, 1997). More than 60% of these protein kinases have either known or suspected functions (Hunter & Plowman, 1997). The *C. elegans* genome contains 435 protein kinases, again comprising more than 2% of the total number of genes (The C. elegans Sequence Consortium, 1998).

All protein kinases catalyze the phosphoryl transfer of the γ-phosphate group of ATP either onto a hydroxyl group of an aromatic or aliphatic amino acid side chain or onto a histidine side chain. They all contain a set of 12 conserved sequence elements whereby the nucleotide binding site includes a glycine-rich loop with a GxGxxG motif (Taylor *et al*. 1993).

The most obvious classification is according to the specificity, the two major classes being the Ser/Thr specific and Tyr specific kinases. A smaller, but no less important, class are the dual specificity kinases such as MAP kinase kinases (MEKs) which catalyze phosphoryltransfer to both types of hydroxyl side chains. A fourth, and structurally completely different, class is formed by the histidine kinase family. A recent classification of protein kinases lists many different families and subfamilies (Hanks & Hunter, 1995). The structures of the catalytic domains of the following protein kinases are known: the non-

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

receptor tyrosine kinases HCK (Sicheri *et al*. 1997), SRC (Sicherie & Kuriyan, 1997; Williams *et al*. 1997; W. Xu *et al*. 1997) and LCK (Yamaguchi & Hendrickson, 1996), the cyclin-dependent kinase 2 (CDK2) (DeBondt *et al*. 1993; Jeffrey *et al*. 1995; Schulze-Gahmen *et al*. 1996), cAMP-dependent protein kinase (PKA) (Knighton *et al*. 1991; Bossemeyer *et al*. 1993; Adams & Taylor, 1992; Lew *et al*. 1997), casein kinase (Xu *et al*. 1995), p38 MAP (Wilson *et al*. 1996), ERK2 (Zhang *et al*. 1994; Canangarajah *et al*. 1997), c-Jun N-terminal kinase (Xie *et al*. 1998), FGF-Receptor tyrosine-kinase (Mohammadi *et al*. 1996, 1997), phosphorylase kinase (Owen *et al*. 1995), calmodulin-dependent protein kinase (Goldberg *et al*. 1996), insuline receptor tyrosine kinase (Hubbard *et al*. 1994; Hubbard, 1997), twitchin kinase (Hu *et al*. 1991), the serine kinase domain of titin (Mayans *et al*. 1998) and protein kinase CK2 from *Zea mays* (Niefind *et al*. 1998). Thus, after 16 different three-dimensional structures of catalytic domains of protein kinases have been solved by crystallography, it is now clear that they all contain a conserved structural fold, as anticipated. The first structure solved was that of the of the cAMP regulated Ser/Thr protein kinase (PKA) (Knighton *et al*. 1991). It was solved in complex with an inhibitor, and later also with ATP and substrate analogues (Bossemeyer *et al*. 1993; Zheng *et al*. 1993). It showed a two-lobed structure where the first lobe (in sequence) is mostly β-sheet and the large C-terminal lobe is mostly α-helical (Fig. 1*b*). The small lobe contains the sequence elements involved in ATP binding such as the glycine rich GxGxxG motif. As in the P-loop there is an invariant lysine, which is situated approximately 20 amino acids C-terminal of the glycine-rich 'phosphate anchor', but points towards the phosphate binding site and seems to be involved in neutralizing negative charges on the β-phosphate. The large lobe contains two conserved sequence elements, one of which, the catalytic loop, is important for the phosphoryl transfer reaction. The other is called the activation loop and is involved in the regulation of the enzymatic activity.

PKA, in terms of regulation probably one of the more simple examples, is the best understood protein kinase with regard to kinetics and enzymatic mechanism and it can be expected that the results gained from PKA are applicable to other serine/threonine protein kinases. Using single turnover kinetics it has been shown that the phosphoryl transfer reaction of PKA is very fast, 200 s^{-1}, and that the release of product is the rate limiting step (Adams & Taylor, 1992; Lew *et al*. 1997). As in the other phosphoryl transfer enzymes, a divalent cation is absolutely required for enzymatic activity. There is an invariant aspartic acid which binds Mg^{2+} with both of its carboxylate oxygen atoms and mutation of this Asp has the most dramatic effect on the enzymatic activity compared to mutations of the other invariant residues (Taylor *et al*. 1993). In addition, there is an invariant aspartic acid in the 'catalytic loop' that is in a position to serve as a general base for activation of the attacking nucleophilic hydroxyl group.

Protein kinases (and protein phosphatases as their counterpart) play the most decisive role in the regulation of almost all important biological pathways such as differentiation, signal transduction, metabolism etc., and since all of these processes are tightly regulated it is mandatory for the proper functioning of complex organisms that the activity of the regulators is itself tightly regulated. Therefore the major difference between the many different kinases is the way their activity is regulated. Very often, the major regulatory element is phosphorylation, but either instead or in addition, many diverse mechanisms such as pseudosubstrate inhibition, autoinhibition, and regulation by second messengers are used in the regulation of protein kinase activity (Hunter, 1995; Johnson *et al*. 1996). The major impact of the three-dimensional structures of protein kinases is not so much to show that their

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

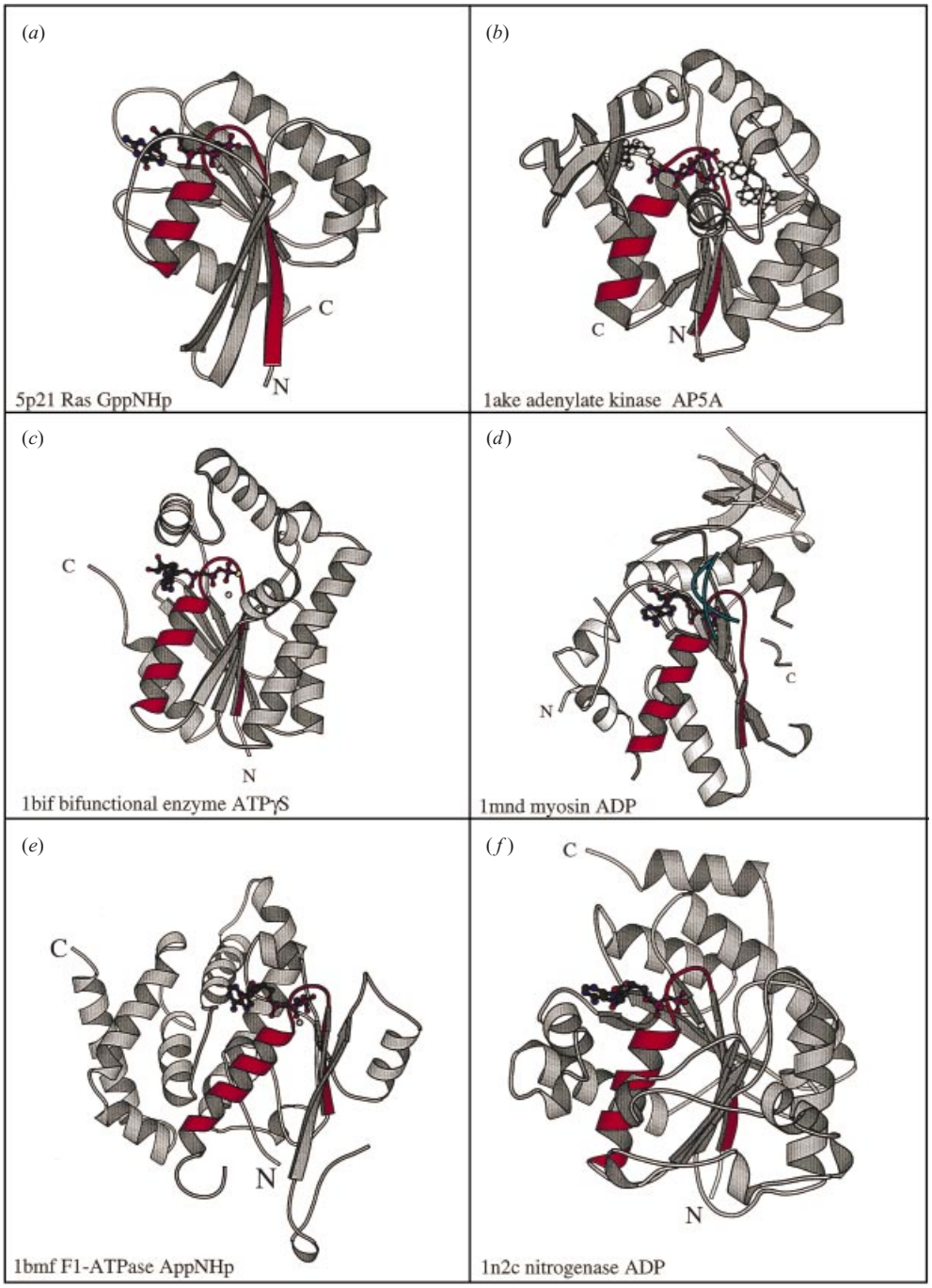


Fig. 1. MOLSCRIPT plots of representative examples of nucleotide binding proteins. Protein Data Bank codes and bond nucleotides are indicated. The secondary structure around the P-loop is shown in red. (*a*) The human small GTPase Ras (residues 1–166). (*b*) Adenylate kinase from *E. coli* with the bisubstrate inhibitor Ap_5A bound (residues 1–214). (*c*) The kinase domain of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase from rat, E.C. 2.7.1.105 and 3.1.3.46 (residues 37–248, the C-terminal domain was deleted for clarity). (*d*) Truncated head of myosin from *Dictyostelium discoideum* with MgADP-AlF_4^- bound (residues 2–690). Segments 202–225 and 239–544 have been omitted to allow a view on the P-loop. The segment 226–238, which is close to the nucleotide, is

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

topology is similar or that they use the same set of invariant residues to bind ATP but rather how the enzyme is regulated. In the most simple case of PKA and others it was shown how the addition of a phosphoryl group in the 'activation loop' creates a catalytically efficient conformation in the active site (Taylor & Radzio-Anselm, 1994), or how a dual phosphorylation of MAP kinases on both Tyr and Thr in the 'phosphorylation lip' affects catalysis (Canagarajah *et al*. 1997). The complex regulation of cyclin dependent kinases by phosphorylation of residues in the ATP binding sites, by cyclin and inhibitors and how phosphorylation and ligand binding affects the enzymatic activity was unraveled by structural studies (DeBondt *et al*. 1993; Jeffrey *et al*. 1995; Russo *et al*. 1996). Similarly, it was shown by the three-dimensional structures of two Src-like kinases, how the intramolecular interaction between the catalytic domain and the two non-kinase domains SH2 and SH3 (for src homology domains type 2 and 3, respectively) inhibits the enzyme (W. Xu *et al*. 1997; Sicheri *et al*. 1997; Williams *et al*. 1997). Therefore, even though we know the basic structure and the basic mode of how protein kinases work, many more structures (probably not 1001) will be needed to understand the full repertoire of activation mechanisms that these key regulators show.

Recently, the structure of the histidine kinase domain of the *E. coli* osmosensor EnvZ has revealed a new type of protein kinase fold which is unrelated to the Ser/Thr/Tyr kinases (Tanaka *et al*. 1998). This has been expected since the histidine kinase family does not qualify on the sequence level as member of this superfamily. Instead, the α/β-sandwich fold of the histidine kinase domain with a five-stranded β-sheet and an additional layer of three α-helices shows unanticipated similarities to the heat shock protein Hsp90 (Prodromou *et al*. 1997a, b) and DNA gyrase B (Wigley *et al*. 1991; Tanaka *et al*. 1998).

2.4 *NDP kinases*

NDP kinases are ubiquituous enzymes which reversibly exchange γ-phosphates between nucleoside di- and triphosphates via a phosphohistidine intermediate. The alternate binding of NDP and NTP to the same site is a classic example of a pingpong mechanism and involves a catalytic mechanism different from most kinases (Traut, 1994). The enzymes show little specificity towards the base and accept nucleotides and deoxynucleotides as substrates since the nucleotide is bound in a shallow groove (Fig. 1*i*) where neither base nor phosphates are buried (Dumas *et al*. 1992; Moréra *et al*. 1994). There are no polar interactions with main chain groups, only with side chains (Moréra *et al*. 1994). NDP kinase does not contain a P-loop, and there is only a single glycine in the nucleotide binding site, which is in contact with the base rather than the phosphates (Moréra *et al*. 1994). At physiological pH values the active site histidine carries one proton, leaving the other imidazole nitrogen free for the reaction with the γ-phosphate of the NTP. NDP kinases show almost no induced fit movements, as is typical for the other, bilobal kinases (Moréra *et al*. 1995).

The highly homologous family of NDP kinases does not contain the classical mononucleotide binding fold. Instead, its topology of a four-stranded antiparallel β-sheet reveals a ferredoxin fold which is also found in the very similar allosteric domain of *E. coli*

---

contributed by the neighboring domain. (*e*) F_1-ATPase α-subunit (residues 24–510) from bovine mitochondria, residues 24–142 and 191–313 are omitted to show the P-loop. (*f*) Nitrogenase complex from *Azotobacter vinelandii* with ADP-AlF_4 bound. Only the nitrogenase iron protein is shown (residues 1–274), the molybdenum-iron protein is omitted for clarity.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

aspartate transcarbamoylase (ATCase) (Moréra *et al*. 1994; Honzatko *et al*. 1982; Steves *et al*. 1990). ATCase is the enzyme that catalyses the first step in the synthesis of UMP and all other pyrimidine nucleotides. It is a classical allosteric enzyme that is negatively regulated by the end product of this metabolic pathway, CTP. In contrast to CTP, ATP acts as an activator, as high ATP concentrations are indicative of a high energy state of the cell and abundance of purine nucleotides available for nucleic acid synthesis (Traut, 1994). ATP and CTP bind to the same site on the regulatory subunit of this enzyme. In spite of ATP and CTP being only regulators, the regulatory domain shows structural similarities to the enzymatically active NDP kinases, suggesting an active enzyme as an ancestor. However, the mode of nucleotide binding is slightly different. After superposition of the β-sheets, the phosphates are only 6 Å apart, yet the riboses and bases point to opposite directions (Moréra *et al*. 1994).

2.5 *Transport ATPases*

The transport of ions and larger substrates across cell membranes is an essential feature of all cells. Several classes of membrane proteins are involved in transport, and most of them seem to be multi-subunit proteins containing a large number of membrane-spanning helices, combined with a catalytic, ATP-binding domain. One of the largest groups is the ABC transporter family which contains a P-loop motif and therefore has been dealt with in the section 'P-loop containing nucleotide hydrolases'. Another group is the cation pumping membrane proteins. They are essential not only for transporting cations like Na^+, K^+, Ca^{2+}, H^+, and heavy metals, but also for energizing all cells and organelles. Since these proteins use ATP to transport cations against the electrochemical gradient across the cell membrane (or synthesize ATP in the reverse way), they are called transport ATPases. Transport ATPases can be divided into three groups, the P-type, V-type and F-type ATPases. V-type ATPases are found solely in eukaryotic **v**acuoles (hence the name) where they are involved in pH regulation, whereas the well-known F-type ATPases are involved ATP synthesis in eubacteria, mitochondria and chloroplasts. V- and F-type ATPases are related, and both contain P-loop motifs. They differ from P-type and ABC cassette membrane pumps by consisting of two multi-subunit domains, a transmembrane and a detachable catalytic sector each containing different subunits (Finbow & Harrison, 1997).

P-type ATPases are single- or multi-subunit cation pumping ATPases (e.g. Na^+,K^+-ATPase, K^+,H^+-ATPase, Ca^{2+}-ATPase) which form an aspartyl**p**hosphate (hence the name) intermediate in the course of ATP hydrolysis. It seems likely that the original P-type ATPases were involved in heavy metal transport and later acquired the ability to transport cations of lower atomic masses (Møller *et al*. 1996). The consensus pattern identifying this protein class is DKTGT[LIVM] [TI] (Bairoch *et al*. 1997), where the first residue (aspartate) is phosphorylated during the transport process. P-type ATPases are also called E1-E2 ATPases since they oscillate between two conformational states, the high-energy E1 form which is able to phosphorylate ADP, and the low-energy E2 form. Common features for all P-type ATPases are a core structure comprising a cytosolic, hydrophilic domain, which contains the phosphorylation- and ATP binding site, and a smaller N-terminal cytosolic region. These two cytosolic regions are anchored in the membrane by a number of transmembrane segments (Møller *et al*. 1996). The structural change caused by the phosphorylation of the aspartate residue in the protruding cytosolic domain is transmitted via a smaller cytosolic loop to the

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

cation binding sites close to the membrane. Kinetic and electrophysiological evidence suggests a translocation mechanism with two gates regulating the translocation rate and transiently occluding the cations during transport through a tight channel. The detailed mechanism is unknown, but certain transmembrane segments have been shown to be of crucial importance, and mutagenesis studies are underway to reveal the role of specific residues in these segments (Møller *et al*. 1996). Recently, the structures of the neurospora plasma membrane H^+-ATPase and the sarcoplasmatic reticulum Ca^{2+}-ATPase were determined to 8 Å by electron microscopy (Auer *et al*. 1998; Zhang *et al*. 1998). They have similar transmembrane domains with 10 helices, but different cytoplasmatic domains. Furthermore, it has been proposed that the catalytic domain of P-type ATPases has the haloacid dehalogenase fold (Aravind *et al*. 1998).

2.6 *DNA topoisomerases*

DNA topoisomerases catalyze the passage of DNA strands through one another, i.e. they untwist supercoiled DNA. There are three subfamilies known: type I enzymes transiently cleave one DNA strand at a time to allow the passage of another strand (Champoux, 1990) and can be divided into the subfamilies IA (*E. coli* enyzmes) and IB (mouse enzymes) which have different structures and reaction characteristics. Type II enzymes cleave a double stranded 'gate' DNA and pass a second duplex through the break, whereupon the cleaved strands are religated (Wang, 1996). The prokaryotic topoisomerase II is also called gyrase, and it is unique in that it negatively supercoils DNA, whereas the others relax positively and negatively supercoiled DNA (Berger, 1998).

In contrast to type I topoisomerases, the reactions catalyzed by type II enyzmes are ATP dependent. However, all topoisomerases seem to have a highly conserved tyrosine, which is covalently linked to a phosphate of the DNA during the reaction. Since the new bond has comparable energy as the original phosphodiester bond, the covalent enzyme–DNA intermediate can be attacked subsequently by the previously liberated end of the DNA to restore the DNA and the free enzyme. Within the past years, several structures of topoisomerase fragments have been solved (Wang, 1996; Lima *et al*. 1994; Sharma *et al*. 1994; Redinbo *et al*. 1998; Stewart *et al*. 1998; Lue *et al*. 1995; Wigley *et al*. 1991; Holdgate *et al*. 1997; Berger *et al*. 1996; Tsai *et al*. 1997). We will focus on type II enzymes in the following description. *E. coli* DNA gyrase, which is a type II topoisomerase, consists of two subunits, A and B, in an A_2B_2 tetrameric arrangement. The B fragment contains two subdomains, of which the N-terminal binds ATP whereas the C-terminal domain interacts with the A subunit and probably with DNA. The structure of the N-terminal ATPase domain of the *E. coli* gyrase B subunit was solved with the ligand AppNHp (Wigley *et al*. 1991). This fragment of gyrase B forms a dimer, creating a 20 Å hole between the monomers which is supposed to play a role in DNA strand passage during the supercoiling reaction (Wigley *et al*. 1991). Later, the structure of a larger fragment of yeast DNA topoisomerase II was solved (Berger *et al*. 1996) that additionally contains some of the domains missing in the *E. coli* gyrase B fragment, including domains which are homologous to the *E. coli* GyrA (Wigley *et al*. 1991). Again this fragment is dimeric and forms a structure with a large hole of approximately 50–55 Å in diameter (Berger *et al*. 1996). In 1997, the structure of a fragment of *E. coli* gyrase

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

A subunit termed the 'DNA breakage/reunion domain' was determined (Cabral *et al*. 1997). Its sequence is homologous to the C-terminal two-thirds of the crystallized yeast enzyme and its overall fold is identical, but the arrangement of the monomers is different (Berger, 1998).

The available three-dimensional data lead to the assumption that ATP-dependent topoisomerases seem to act as an ATP-modulated clamp that closes when DNA binds and opens after ATP hydrolysis and release. These ATP-modulated movements are tightly coupled to the transport of one DNA double helix through another (Berger *et al*. 1996). The detailed mechanism involves a transesterification between a pair of tyrosyl residues, one in each half of the dimeric enzyme, and a pair of DNA phosphodiester bonds. The hydroxyl groups of the tyrosines become covalently linked to the phosphoryl groups at the 5′ ends of the transiently cleaved DNA. The tyrosine-linked DNA ends now separate from each other, thus opening a gate in the DNA for passage of the other DNA segment. After passage, the gate closes again, and another transesterification restores the continuity of the DNA strands. This mechanism requires that the enzyme undergoes large conformational changes (Berger *et al*. 1996). A low-resolution view of the intact enzyme was provided by single particle electron microscopy (Schultz *et al*. 1996; Benedetti *et al*. 1997). According to these results, topoisomerase II is a 'V'-shaped molecule that changes shape upon addition of AppNHp. A model was proposed in which the dimerization of the ATPase domains opens and closes a gate through which one of the DNA strands can enter. It then passes through the other, cleaved strand which is bound to the enzyme (Roca & Wang, 1994). Nucleotide binding seems sufficient to allow a single cycle of DNA transport (Roca & Wang, 1992, 1994; Bates *et al*. 1996). There is no explanation yet of how topoisomerases recognize the topology of the DNA (Berger, 1998), and more research needs to be done to reveal all properties of these enzymes.

Hsp90 is a molecular chaperone which is essential for the folding of several proteins involved in cell cycle regulation, steroid hormone responsiveness, and signal transduction. (For an introduction to chaperones see the next section.) The structure of its N-terminal domain was solved in complex with ADP, ATP and an inhibitor, geldanamycin, respectively (Prodromou *et al*. 1997a, b). The structure is a dimer with a highly twisted 16-stranded β sheet (Fig. 1*f*) with the opposing faces of the β-sheets in the dimer forming a potential peptide-binding cleft. The ATP binding site is homologous to the corresponding site in DNA gyrase B, and the overall fold shows extensive similarities as well, which was completely unexpected (Prodromou *et al*. 1997a,b; Tanaka *et al*. 1998). A fold relationship has also been detected between Hsp90 and the catalytic domain of the histidine kinase (Tanaka *et al*. 1998). ATP binding induces conformational changes in Hsp90, but the mechanism and function of these changes are unknown (Prodromou *et al*. 1997b).

2.7 *Chaperones and chaperonins*

Chaperones are proteins that promote protein folding, assembly, degradation and membrane translocation by binding to unfolded polypeptide segments. They are not only expressed in response to metabolic stress, but also under normal conditions, and their general function seems to be the binding to aggregation-prone exposed hydrophobic surfaces (Zhu *et al*. 1996). A well-known family are the ubiquitous 70 kDa heat shock proteins (Hsp70), whose

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

members have similar functions and a common domain structure: a substrate-binding domain binds the polypeptide and an ATPase domain regulates peptide exchange. Hsp70s have relatively high binding affinities for ATP – 100 fold higher than for ADP – but are slow ATPases (Theyssen *et al*. 1996; Bukau & Horwich, 1998). The nature of the bound nucleotide determines the affinity and kinetics of peptide binding: the affinity is low in the ATP-bound state with concomitant fast binding and release kinetics, and vice versa in the ADP-bound state.

One of the best studied chaperones is DnaK from *E. coli*, whose substrate binding and ATPase domains have known structures (Zhu *et al*. 1996; Harrison *et al*. 1997). The substrate binding domain consists of two subdomains, an N-terminal β-sandwich domain with an unique fold, and an α-helical domain. The catalytic ATP-binding domain of DnaK has a fold similar to the heat shock cognate Hsc70 and therefore is discussed in the following section (actin-like ATPase domains). Hsp90 was discussed in the previous section since it has a fold similar to DNA gyrase B.

Chaperonins are a well-studied family of molecular chaperones, the prototype being the GroEL/GroES system of *E. coli*. Chaperonins have been shown to be large multi-subunit assemblies which are essential for the folding of many proteins at all temperatures and require ATP for efficient folding. They function to prevent misfolding by stabilization of aggregation-prone intermediates inside a folding chamber which cycles in an ATP-dependent manner between a binding-active and a folding-active conformation (Bukau & Horwich, 1998; Braig *et al*. 1994; Smith & Fisher, 1995).

GroEL belongs to the Hsp60 family found in bacteria, mitochondria and chloroplasts, whereas the related TF55/TCP-1 family can be found in archaebacteria and eukaryotes. The bacterial chaperonins are built as two stacked seven-membered rings, whereas the archaebacterial and eukaryotic homologues contain rings of eight or nine subunits each (Hartl, 1996). The three-dimensional structure of GroEL has been widely studied by electron microscopy (Chen *et al*. 1994; Roseman *et al*. 1996; Llorca *et al*. 1997; Schmidt *et al*. 1994) and X-ray crystallography (Braig *et al*. 1994; Boisvert *et al*. 1996; Z. Xu *et al*. 1997). The GroEL subunit consists of three domains, apical, intermediate and equatorial, the latter one being the ATPase domain. It is highly α-helical – one of the few nucleotide binding domains which is not an α,β-protein – and composed of two 4-helical subdomains which are related by a pseudodyad passing through the ATP binding site (Fig. 1*k*). It is most likely that the two subdomains were created by gene duplication. The nucleotide is bound to a pocket in the equatorial domain, one ATP per subunit, 14 ATP per molecule. The pocket faces the interior of the GroEL cylinder and is formed mainly by residues from the equatorial domain. Residues from the neighboring intermediate domain also contribute residues critical for catalysis to the pocket. The strictly conserved sequence motif DGTTT interacts with the β- and γ-phosphates of ATP. In contrast to the P-loop arrangement this phosphate binding loop is located between two helices, not between a helix and a β-strand (Boisvert *et al*. 1996). The effect of nucleotide binding on the observed conformational changes in GroEL must be allosteric since the nucleotide contacts mainly one domain, whereas the observed changes affect the other domains as well (Boisvert *et al*. 1996).

A relative of GroEL is the thermosome, the archeal group II chaperonin from *T. acidophilum*. The hexameric protein has domain folds similar to GroEL, but a different kind of inter-ring contacts. The structure has been solved in complex with ADP-ALF_3 and corresponds to the closed conformation (Ditzel *et al*. 1998).

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

2.8 *Actin-like ATPase domain*

The structure of the actin-like ATPase domain seems to be widespread among proteins with quite different biological functions (Bork *et al*. 1992). No overall sequence similarity between the different families can be detected, nor is there any common biochemical reaction involved, but the structures of actin (Kabsch *et al*. 1990), the heat shock cognate Hsc70 (Flaherty *et al*. 1990), hexokinase (Anderson *et al*. 1978; Bennett & Steitz, 1980), and glycerol kinase (Hurley *et al*. 1993) show remarkable resemblance. Actin uses ATP hydrolysis to control its polymerization state and the stability of the F-actin filament, hexokinase transfers the γ-phosphate group of ATP to the C-6-OH group of glucose to produce 6-phosphorylated glucose and ADP. Glycerol kinase catalyzes the equivalent reaction with glycerol, and the chaperone Hsc70 is involved in stabilizing and folding of other proteins as described above. The ATPase domain of DnaK shares 50% identical residues with Hsc70 and has a similar structure (Harrison *et al*. 1997).

The common feature of these actin-like domains is a large cleft between subdomains containing several conserved motifs with the ATP binding site at the bottom (Fig. 1*l*). Probably the proteins have not only a common structure but also a common mechanism since the active sites have similar residues and the large cleft indicates the potential for a large conformational change during nucleotide binding. Such a domain movement has already been shown for hexokinase. Actually, hexokinase was the first example for which an induced fit mechanism was shown (Steitz *et al*. 1981). Five conserved motifs have been identified (called 'phosphate 1 and 2', 'adenosine', and 'connect 1 and 2') whose function is either binding of the ATP phosphates and the base moiety or the formation of the hinge-bending interface (Bork *et al*. 1992). The phosphates of ATP are mainly bound by residues on two β-hairpins ('phosphate' motifs), one from each side of the cleft. The loops as well as the corresponding domains are arranged symmetrically, and both domains contain a five-stranded β-sheet of identical topology, which suggests an ancient gene duplication event. The two phosphate binding motifs (sequences SGDGVT and NGSGLVK in actin) have no homology to the P-loop described above. A unique feature of the Hsc70 ATPase is the requirement for monovalent cations in the enzymatic reaction (Wilbanks & McKay, 1995) which have been shown by structural investigations to possess specific binding sites in the active site of the enzyme (Wilbanks & McKay, 1995). Usually, a divalent ion is necessary for the reaction.

A database search for all five described motifs has revealed several new members of this class, which are either sugar kinases (fucokinase, gluconokinase, xylulokinase and ribulokinase) or prokaryotic cell cycle proteins (FtsA, MreB and StbA) (Bork *et al*. 1992). These data lead to the assumption of a common ancestral protein that underwent gene duplication, thus forming an ATP binding site between the two domains, whereupon the divergent evolution created today's diverse members of the group (Bork *et al*. 1992). The diversity of biological function is implemented by inserted domains that show high variability as well as by the polymorphism of the loops in the core region. There are also similarities between the actin-like domain and RNaseH as well as the HIV RT RNAse H/connection domain (Artymiuk *et al*. 1993). This indicates a possible evolutionary relationship between these three proteins which have no detectable sequence homologies.

2.9 *Phosphoglycerate kinase*

A common feature of phosphoryl transfer enzymes is the necessity to shield their substrates

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

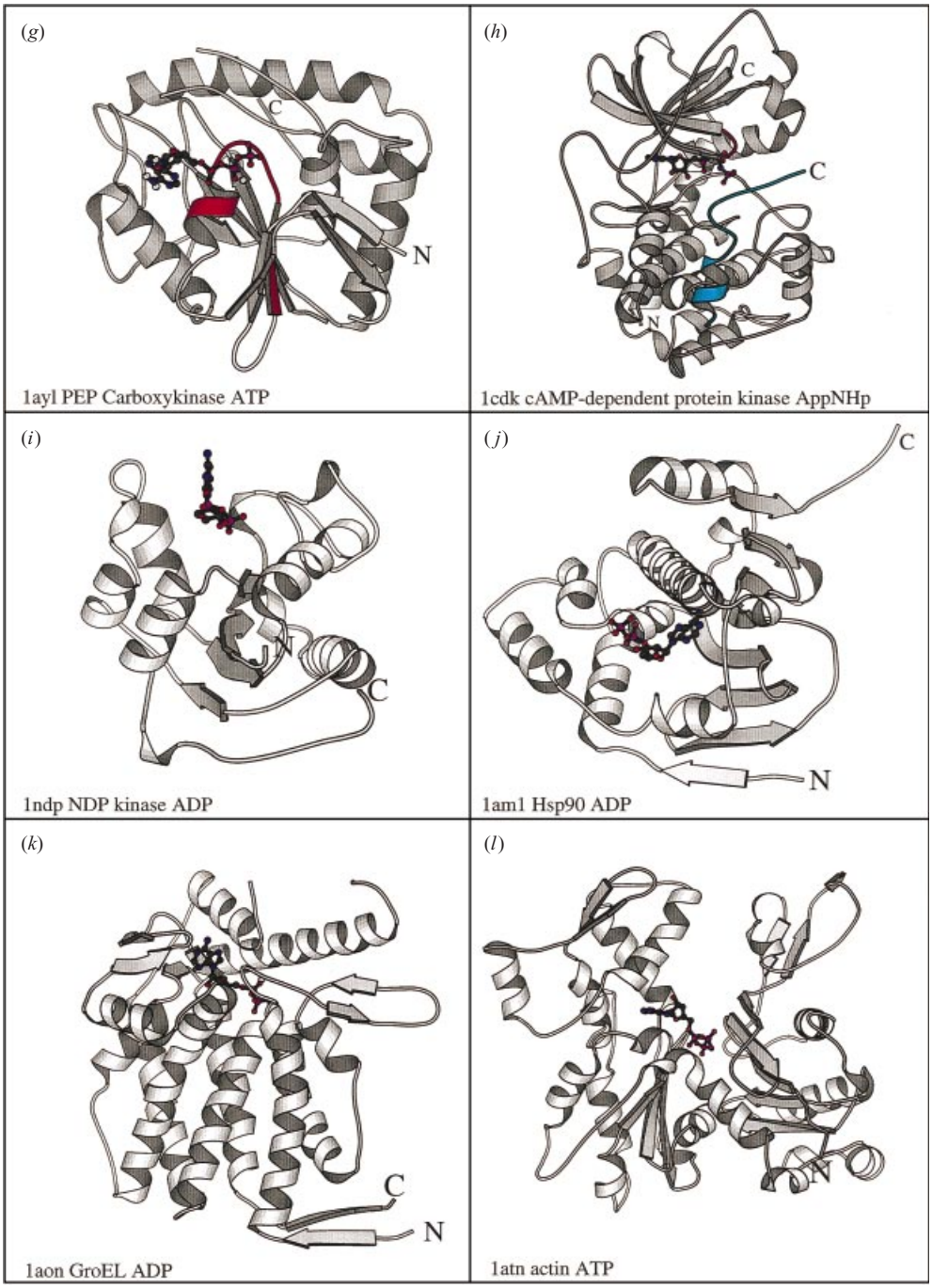


Fig. 1 (*cont.*). MOLSCRIPT plots of representative examples of nucleotide binding proteins. (*g*) Phosphoenolpyruvate-carboxykinase from *E. coli*, residues 228–540. The N-terminal domain (1–227) is omitted to show the nucleotide binding site. (The secondary structure around the P-loop is shown in red.) (*h*) cAMP-dependent protein kinase catalytic subunit from *Sus scrofa*, residues 8–350, with an inhibitory peptide from a protein kinase inhibitor (residues 5–24) bound. The glycine-rich loop has the motif GxGxxGxV (residues 50–57). (*i*) NDP-kinase from *Dictyostelium discoideum*, residues 6–155. (*j*) Heat shock protein Hsp90 from *Saccharomyces cerevisiae*, residues 2–214. (*k*) GroEL from *E. coli*, residues 2–525. Residues 152–385 have been omitted to show the nucleotide binding site with roughly twofold symmetry. (*l*) Actin from rabbit muscle, residues 1–372.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

from the solvent to prevent water from attacking the γ-phosphate and thus uncontrolled ATP hydrolysis (Anderson *et al*. 1979). As demonstrated most conclusively for the nucleoside monophosphate kinases such as adenylate kinase (Schulz *et al*. 1990; Vonrhein *et al*. 1995), this usually occurs via a large domain movement that closes the substrate binding site upon binding of the substrates. Such movements between lobes or domains have also been observed in hexokinase (Steitz *et al*. 1981), phosphofructokinase (Shirakihara & Evans, 1988) and phosphoglycerate kinase (PGK) (Auerbach *et al*. 1997).

PGK synthesizes ATP using the energy stored in 1,3-bisphosphoglycerate, which is an important step in the glycolytic pathway. PGK consists of two globular domains of almost equal size (Fig. 1*m*) which are connected by an α-helix (Bryant *et al*. 1974; Watson *et al*. 1977). The C-terminal domain features a six-stranded β-sheet and binds the nucleotide, whereas the N-terminal domain binds the phosphoglycerate.

PGK is activated by high concentrations of substrates, which results in biphasic kinetics. It has been proposed, based on different studies, that there are two binding sites for each substrate (Larsson-Raznikiewicz, 1973; Rao *et al*. 1978; Ray & Rao, 1988; Schierbeck & Larsson-Raznikiewicz, 1979; Scopes, 1978), the secondary sites having low affinity.

For a long time, the catalytically competent conformation of PGK was unknown despite extensive structural investigations. Several structures of the unliganded enzyme (horse PGK) (Blake & Evans, 1974) and of binary complexes (porcine PGK with 3PGA (Harlos *et al*. 1992), *B. stearothermophilus* PGK with ADP (G. J. Davies *et al*. 1994)) have been solved, but only recently the structures of PGK from *Trypanosoma brucei* (Bernstein *et al*. 1997) and *Thermotoga maritima* (Auerbach *et al*. 1997) were published which both contain 3-phosphoglycerate and either ADP or an ATP analogue and represent the closed form of the enzyme. In 1996, the structure of a mutant of yeast PGK (R65Q) was solved in complex with Mg–AMP–PNP and 3-phospho-D-glycerate (McPhillips *et al*. 1996), but although the two substrates were present, the structure was not fully closed.

Phosphoglycerate kinases are identified by a glycine rich consensus motif ([AV]xxGG[AS]KVxxK (Bairoch *et al*. 1997)) which is positioned in a loop connecting a β-sheet with an α-helix and thus remotely resembles a P-loop (Saraste *et al*. 1990). Indeed, there is recent evidence that the first lysine in this sequence of PGK is involved in the transfer of the phosphoryl group from the C1 position of 1,3-bisphosphoglycerate to ADP (Auerbach *et al*. 1997). Surprisingly, the article of Bernstein *et al*. (1997) does not mention this lysine at all, in spite of its position close to the catalytic site and its possible role in phosphoryl transfer. However, it was shown that the catalytic mechanism is associated with a large hinge-bending motion of the two domains, thus closing the active-site cleft and bringing the substrates in close proximity. The formation of the 'closed' form of the enzyme occurs after binding of both substrates and enables the nucleophilic attack of the oxygen atom at the 1-phosphate group of 1,3-biphosphoglycerate by the β-phosphate oxygen atom of ADP. The reaction leads to an inversion of stereochemistry at the newly formed γ-phosphorous atom and follows a single displacement mechanism in which the phosphoryl group is transferred directly via an associative (in-line) mechanism involving a pentavalent transition state (Auerbach *et al*. 1997). This is the most commonly discussed mechanism of enzymatic phosphoryl transfer, although it has not been strictly proven in most cases (for discussions see: Admiraal & Herschlag, 1995; Steitz & Steitz, 1993; Knowles, 1980; Sträter *et al*. 1996). During the transfer, the charged penta-coordinate transition state is probably stabilized by the sidechain of the conserved lysine as was proposed for Ras (Pai *et al*. 1990), adenylate kinase

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

(Schulz, 1992) or cAMP-dependent protein kinase (Bossemeyer *et al*. 1993). A similar mechanism can be found in hexokinase or nucleoside monophosphate kinases, whereas the reaction of NDP kinases or phosphatases proceeds via a phosphorylated enzyme intermediate (Auerbach *et al*. 1997).

2.10 *Pyruvate kinase*

The two major regulatory enzymes in glycolysis are phosphofructokinase and pyruvate kinase (PK). PK synthesizes the second ATP molecule of glycolysis by transferring a phosphate from phosphoenolpyruvate (PEP) to ADP, thus creating pyruvate and ATP. This reaction proceeds in two steps. The first step is the transfer of a phosphate from PEP to ADP to give ATP and the enolate form of pyruvate. In the second step, a proton is added to the enolate to produce the keto form of pyruvate (Larsen *et al*. 1994). This energetically favorable ketonization renders the overall reaction strongly exergonic so that it is metabolically irreversible (Larsen *et al*. 1994). There is evidence that the reaction needs the presence of two divalent cations (Gupta *et al*. 1976; Baek & Nowak, 1982), similar to the two-metal-ion mechanism proposed for other phosphoryl transfer reactions such as those catalyzed by DNA polymerases (Steitz & Steitz; 1993; Sträter *et al*. 1996). PK uses ADP as substrate and is also allosterically regulated by metabolites such as ATP. Most pyruvate kinases are tetrameric proteins with a high molecular weight ($\sim$ 240000 Da) and can be divided in two groups which have a different set of allosteric effectors. Most bacteria have only one type of pyruvate kinase, whereas *E. coli* has both forms. The eukaryotic enzymes are usually more similar to the type I pyruvate kinase of *E. coli* (Stuart *et al*. 1979; Larsen *et al*. 1994; Mattevi *et al*. 1995).

Each monomer of the *E. coli* type I pyruvate kinase (Mattevi *et al*. 1995) consists of three subdomains. The active site is located in a cleft between domains A and B. The largest domain (A) folds into an eight-stranded α/β-barrel, whereas the neighboring domain (B) forms a lid covering the α/β-barrel and folds into a mixed β-barrel with only one short α-helix. Ligand binding probably leads to a closure of the active site as seen in the structure of the rabbit muscle enzyme which has a similar domain structure to the *E. coli* enzyme (Larsen *et al*. 1997). The C-terminal domain (C) displays an α/β-structure and interacts tightly with the large domain (Mattevi *et al*. 1995). The allosteric effector binding site is located between domains A and C, and the allosteric effect is attributed to large domain rotations upon binding of the regulator, leading to a different shape of the active site (Mattevi *et al*. 1995). Unfortunately, there is still no structure available which has a nucleotide bound, so that the mode of binding and the catalytic mechanism remains elusive.

2.11 *Phosphofructokinase*

Phosphofructokinase (PFK), like hexokinase, is one of the enzymes involved in the initial steps of glycolysis. It catalyzes the formation of fructose-1,6-bisphosphate from fructose-6-P and ATP. It is allosterically regulated by different substrates in different organisms, e.g. by ADP, GDP, AMP or fructose-6-phosphate. Each subunit of the tetrameric enzyme consists of two domains, a larger one that binds ATP and a smaller one which participates in binding of fructose-6-phosphate (Fig. 1*n*). Both domains are $\alpha\beta\alpha$ sandwich structures with three layers (Evans *et al*. 1981; Rypniewski & Evans, 1989). In the *Bacillus stearothermophilus* enzyme, the allosteric effector site is formed by the large interface between two monomers of the enzyme which binds both activator and inhibitor. The ATP in the catalytic site is bound

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

almost entirely by residues of the large domain, whereas the fructose-6-phosphate binding site is formed by residues from both domains and another molecule of the tetramer. The nucleotide binding domain contains a glycine rich loop which is in close contact to the nucleotide like the P-loop in Ras. It has the sequence VIGGDGSYQ (4pfk), whereby structurally the second glycine (underlined) roughly corresponds to Gly12 in Ras and the third glycine (underlined) to Gly15 in Ras. The neighboring secondary structures are a β-sheet and an α-helix as in the original P-loop, but the loop in PFK is at a very different position relative to the nucleotide compared to Ras. The reaction catalyzed by PFK follows a complicated kinetic mechanism due to the cooperativity between subunits and due to the influence of allosteric effectors (Auzat *et al*. 1994). The allosteric behavior of PFK was elucidated by the structure of the enzyme from *B. stearothermophilus* PFK in an inhibited state (Schirmer & Evans, 1990). It was shown that the relative orientation of the domains of the tetrameric enzyme changes upon transition from the high-affinity active form (R state) to the low-affinity T state. An arginine residue contributes the only sidechain to the active site which changes its position upon transition between the inactive and active forms of the enzyme. It bridges the two substrates fructose-6-phosphate and ATP (although it is never highly ordered in the crystal structure), whereas in the inactive form of the enzyme it cannot interact with the substrates (Schirmer & Evans, 1990).

A second class of PFKs besides the ATP-dependent enzymes are the PP_i-dependent phosphofructo-1-kinases. They show low but significant sequence homology to each other (Ladror *et al*. 1991) so that structural similarity of the active site structures and mechanisms has been proposed (Hinds *et al*. 1998). A model of the active site based on the coordinates of the *Escherichia coli* ATP dependent PFK has been built which predicts the importance of two conserved arginine residues (Hinds *et al*. 1998), but no structure has been solved to date.

2.12 *Creatine kinase and glutamine synthetase*

Creatine kinase belongs to the family of guanidino kinases and catalyzes the regeneration of ATP from phosphocreatine in several tissues, e.g. muscle. The reaction is believed to be a direct inline transfer of the phosphoryl group accompanied by a conformational change of the enzyme (Fritz-Wolf *et al*. 1996; Kabsch & Fritz-Wolf, 1997). The octameric mitochondrial enzyme uses the intramitochondrially produced ATP for the synthesis of phosphocreatine, which is then exported to the cytosol for regeneration of cytosolic ATP by the cytosolic isoforms ('phosphocreatine-circuit' energy-shuttling system). Each subunit of creatine kinase can be divided in two subdomains, a small helical N-terminal domain and a large C-terminal domain which is similar to the C-terminal domain of glutamine synthetase (Kabsch & Fritz-Wolf, 1997; Yamashita *et al*. 1989). The common structure consists of a duplicated $\beta\alpha\beta\beta\alpha$ unit, and both enzymes bind ATP at a similar position (Fig. 1*o*), suggesting an evolutionary relationship despite the absence of significant sequence similarities (Kabsch & Fritz-Wolf, 1997).

In contrast to creatine kinase, the structure of the transition state of arginine kinase (Zhou *et al*. 1998) shows a well-ordered active site and thus allows further conclusions about the catalytic mechanism. The enzyme was crystallized in complex with ADP-NO_3 and offers direct evidence for the importance of precise substrate alignment in the catalysis of bimolecular reactions. Five conserved arginine residues and a Mg^{2+} ion are found to hold the ATP phosphoryl groups in place, four of which contact the transferred γ-phosphoryl group.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Glutamine synthetase plays an essential role in the metabolism of nitrogen by catalyzing the condensation of glutamate and ammonia to form glutamine. There seem to be three different classes: class I enzymes (prokaryotic) with 12 identical subunits, class II enzymes (eukaryotic and certain prokaryotes) which are octamers, and class III enzymes which are hexamers and have been found only in *Bacteroides fragilis* and *Butyrivibrio fibrisolvens* (Bairoch *et al*. 1997). The three classes are structurally related although the sequence similarities are not extensive. All proteins contain a glycine-rich region which is thought to be involved in ATP binding (PROSITE consensus [NPAT]G[GSTAN]GxHxxxS (Bairoch *et al*. 1997)).

Bacterial glutamine synthetase is composed of 12 identical subunits arranged in two rings of six, with the active site between each pair of subunits in a ring (Almassy *et al*. 1986; Yamashita *et al*. 1989; Liaw *et al*. 1994). The funnel-shaped active site cavity is 30 Å wide at the top, about 10 Å wide at the bottom and 45 Å deep. It is water accessible from top and bottom (Liaw *et al*. 1995). Figure 1*o* depicts the nucleotide binding domain of one of the 12 subunits. The following reaction mechanism has been proposed: ATP binds first at the top of the funnel-shaped active site cavity, then glutamate binds at the bottom of the active site near a flexible loop. The γ-carboxylate of the glutamate attacks the γ-phosphate of ATP to form γ-glutamyl-phosphate and ADP. The presence of ADP triggers changes in the vicinity of the binding site which facilitate binding of the third substrate, the ammonium ion. The ammonium ion binds and loses a proton to form the more reactive ammonia, which in turn attacks the γ-glutamyl-phosphate to form a tetrahedral intermediate. After release of phosphate one proton from the γ-amido group of the intermediate is transferred to the enzyme (Glu 327), yielding glutamine as the reaction product. It is crucial for catalysis that ammonium binds to the enzyme more strongly than water due to a negatively charged binding pocket. If water would attack instead of the ammonia, the result would be glutamate, therefore the ammonium binding site has to discriminate against water to allow this reaction in an aqueous environment (Liaw *et al*. 1995). Glutamine synthetase is inhibited by several end products of the glutamine metabolism, and X-ray studies suggested that glycine, alanine and serine, for example, inhibit the enzyme by competing with the substrate glutamate for the active site (Liaw *et al*. 1993).

2.13 *The ATP-grasp fold*

In 1993, a new nucleotide binding motif in non-Rossman fold-type structures was found in glutathione synthetase (GSHase), called the 'palmate (β-sheet) fold' (Yamaguchi *et al*. 1993). This ATP binding motif is an antiparallel β-sheet, on the edge of which a glycine-rich loop joins two β-strands (Fig. 1*p*). The glycine-rich loop differs from the P-loop since there is no conserved lysine and no serine or threonine (Yamaguchi *et al*. 1993). Later, this motif was reported not only in glutathione synthetase but also in D-alanine:D-alanine ligase (DD-ligase) (Fan *et al*. 1994, 1995), biotin carboxylase (BC) (Artymiuk *et al*. 1996), succinyl-CoA synthetase (also termed succinate-CoA ligase) (Wolodko *et al*. 1994; Matsuda *et al*. 1996), malate-CoA ligase (Galperin & Koonin, 1997), ATP-citrate lyase (Galperin & Koonin, 1997) and carbamoyl phosphate synthetase (Thoden *et al*. 1997; Raushel *et al*. 1998), and was also called the 'ATP-grasp' domain to indicate that the structure appears as an open right hand that grasps the ATP with its palm and where a glycine-rich loop constitutes the thumb (Galperin & Koonin, 1997). Sequence similarities with biotin carboxylases indicate that this

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

superfamily additionally includes the biotin-dependent carboxylase domains of pyruvate carboxylase and propionyl-CoA carboxylase (Galperin & Koonin, 1997).

The structural similarity between the enzymes GSHase and DD-ligase, which catalyze formation of a peptide bond, is extensive. Both proteins consist of three domains with similar folds, although the sequence homology is quite low. ATP binds between the two C-terminal domains. Fan and coworkers argued in favor of an evolutionary relationship between the two enzymes and postulated that they represent a general fold for a class of peptide synthetases that cleave ATP to ADP, rather than to AMP (Fan *et al*. 1995). An evolutionary relationship is also suggested by the common catalytic mechanism which involves two magnesium ions in the active sites of each enzyme.

GSHase catalyzes the synthesis of glutathione from γ-L-glutamyl-L-cysteine (γ-Glu-Cys) and glycine in the presence of ATP. The mechanism is thought to involve a phosphorylation of the carboxylate of the cysteinyl moiety of γ-Glu-Cys to give γ-L-glutamyl-L-cysteinyl phosphate. This intermediate is subsequently attacked by glycine, leading to the loss of inorganic phosphate and formation of the product (Yamaguchi *et al*. 1993; Hara *et al*. 1996). By analogy, DD-ligase ligates the ATP-activated carbonyl group of D-alanine to the amino group of another D-alanine (Fan *et al*. 1994).

The similarity between the GSHase/DD-ligase fold and biotin carboxylase (Artymiuk *et al*. 1996; Waldrop *et al*. 1994) is quite astonishing since this finding links together two families of enzymes, namely the ADP-forming peptide synthetases and the large family of homologous ADP-forming biotin-dependent carboxylases, which include carbamoyl-phosphate synthetase, propionyl-CoA carboxylase, pyruvate carboxylase and L-glutamine-amido-ligase. In *E. coli*, the biotin carboxylase catalyzes the first step of the acetyl-CoA carboxylase reaction, the ATP-dependent carboxylation of the 1′-nitrogen of biotin to form carboxybiotin. Bicarbonate is used as source of CO_2. The reactions of GSHase, DD-ligase and biotin carboxylase are similar in that all three catalyze the formation of a carbon–nitrogen bond between a carboxyl group and an amino group by utilizing ATP to activate the carboxylate. The only difference is that the biotin carboxylase reaction involves a secondary amine, whereas the other two enzymes use a primary amine. Nevertheless, the reaction mechanism is assumed to be similar (Artymiuk *et al*. 1996). The evolutionary relationship is also corroborated by sequence homologies between the three enzymes (19% overall similarity between BC and DD-ligase, 16–17% identity between GSHase and BC/DD-ligase). A relationship to GSHase, DD-ligase and biotin carboxylase has been reported for other C–N bond ligases like glutamine synthetase and adenylosuccinate synthetase (ASS) since they have a similar phosphorylated intermediate (Yamaguchi *et al*. 1993). The former enzyme phosphorylates the γ-carboxylate of glutamate which is displaced by an ammonium ion, the latter phosphorylates the carbonyl oxygen of the base moiety of IMP which is then displaced by the amino group of aspartate (Bass *et al*. 1984). However, despite certain similarities in the reaction mechanism, there is no indication that glutamine synthetases belong to the ATP-grasp superfamily (Galperin & Koonin, 1997), and ASS has a Ras-like fold as discussed above (Silva *et al*. 1995).

A relationship to the GSHase family is also discussed for an enzyme of the citric acid cycle, succinyl-CoA synthetase (SCS), which is an $\alpha 2\beta 2$ heterotetramer in *E. coli*. In mammalians, there are two isoforms of this enzyme, one of which synthesizes ATP and the other GTP, using ADP/GDP, succinyl-CoA and inorganic phosphate. In bacteria, SCS has degenerate specificity. In *E. coli*, the α-chain consists of two domains, each of which includes a classical

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

NAD-binding motif (Rossman *et al*. 1975). One molecule of CoA is bound to each α-chain. The β-chain consists of three domains, of which the C-terminal domain contains another NAD-binding motif. The two N-terminal domains show similarity to the two C-terminal ATP-binding domains of GSHase. The binding sites for the substrates succinate and ATP are not known yet, but the strong structural similarity suggests an ATP binding site similar to the one in GSHase. However, it has to be noted that the reaction mechanism of succinyl-CoA-synthetase differs from GSHase and DD-ligase. In the latter two enzymes, the γ-phosphate of ATP is transferred directly to the carboxylate group of a substrate, whereas in succinyl-CoA synthetase a phosphohistidine intermediate is formed. As a first step of the reaction, the energy of the thioester bond of succinyl-CoA is used to form an activated carboxyl-phosphoanhydride, which indicates the common intermediate between this reaction and that of the peptide ligases (Wolodko *et al*. 1994). The activated phosphoryl group is then transferred to ADP/GDP via a covalent phosphoryl–enzyme intermediate. The phosphorylated residue has been identified as histidine 246 in the α-chain, and a structure of the *E. coli* enzyme in the phosphorylated form was solved (Wolodko *et al*. 1994). If one assumes ADP/GDP binding at the same position as in GSHase, this histidine, which is located on a long loop, would be too far away from the substrate end of CoA for direct interaction. A solution for this problem could be a movement of the histidine loop upon binding of ADP that would serve as a phosphate shuttle between the CoA site and the ADP site (Matsuda *et al*. 1996; Murzin, 1996).

Although a β-sheet similar to GSHase/DD-ligase with a roughly similar glycine-rich loop can be found in protein kinases and seryl-tRNA synthetases, the chains would have to be reversed to fit the structures together, so that an evolutionary relationship is very unlikely.

Recently, another protein containing the ATP-grasp fold has been identified. Synapsin is a protein associated with synaptic vesicle membranes and is phosphorylated by multiple kinases. A vast variety of proteins interact with synapsins, for example, elements of the cytoskeleton, calmodulin, annexin and Grb2. Synapsin consists of five domains, the central of which is the most highly conserved one. The structure of this domain has been solved (Esser *et al*. 1998) and, unexpectedly, shows homology to the GSHase structure. Indeed it was then shown to bind ATP in a similar orientation, but the nature of the phosphoryl acceptor and the physiological context are still unclear (Esser *et al*. 1998).

Pyruvate phosphate dikinase (PPDK) catalyzes the conversion of ATP, inorganic phosphate and pyruvate to AMP, pyrophosphate and phosphoenolpyruvate. Since this enzyme cleaves the α-β-phosphoanhydride bond it should be treated in the next section, but since PPDK also exhibits the ATP-grasp fold it is discussed here. PPDK shows no apparent sequence similarity to DD-ligase, GSHase, BCase or SCS (Galperin & Koonin, 1997). The enzyme is a large dimeric molecule with three subdomains per monomer. Like succinyl-CoA synthetase, PPDK shuttles a phosphorylated histidine residue between the ATP- and the substrate binding domains. However, in the case of PPDK the shuttle is a whole domain rather than a loop (Murzin, 1996; Herzberg *et al*. 1996).

3. Reaction catalyzed: $\mathbf{ATP \rightarrow AMP + PP_i}$

The majority of NTP-hydrolyzing enzymes cleave the β-γ-phosphate bond of NTP. In the following we will describe enzymes which produce an adenylated intermediate during their reaction or directly cleave the α- and β-anhydride bond. The driving force for these reactions

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

may be increased by the subsequent cleavage of the released pyrophosphate to inorganic phosphate. Enzymes that cleave the α-β-phosphate bond, though less numerous, are involved in a variety of reactions and probably do not have a single phylogenetic origin since there is no detectable relationship on the sequence nor on the fold level. Therefore, the different fold families will be treated in different sections. This leads to, for example, a division of the aminoacyl-tRNA synthetases according to their different topologies, which is either an ATP pyrophosphatase fold (class I tRNA synthetases) or a completely different fold type (class II tRNA synthetases). Differences can also be found in the reaction mechanisms, e.g. DNA and RNA ligases have – in contrast to most other enzymes in this section – a covalent enzyme-AMP intermediate, and different cofactors are used in different families.

3.1 *DNA-ligases, mRNA capping enzymes*

Although the mechanism and the details of the structure are quite different from glutathione synthetase (GSHase) and its relatives, DNA ligases are also classified as ATP-grasp proteins (Murzin *et al*. 1995). Compared to GSHase, the phosphate moieties of the bound nucleotides point in opposite directions when the β-sheets and the base moieties are superimposed. DNA ligases can be divided in two classes: the bacterial DNA ligases are NAD-dependent, whereas the ligases of eukaryotes, archaebacteria, viruses and phages use ATP (Subramanya *et al*. 1996b). The two reaction mechanisms differ only in the first step, where either ATP or NAD are used; the other steps are similar. The ATP-dependent ligases catalyze the ligation of 5′-phosphate-terminated donor strands to 3′-hydroxyl-terminated acceptor strands via three sequential nucleotidyl transfer reactions. In the first step, attack on the α-phosphate of ATP by the ligase results in production of PP_i and a covalent enzyme-AMP intermediate in which AMP is linked to the enzyme via the ϵ-amino group of a conserved lysine. The nucleotide is then transferred to the 5′-phosphorylated end of the donor strand to form DNA-adenylate. Attack by the 3′ OH of the acceptor strand on the DNA-adenylate joins the two polynucleotides and liberates the AMP. The ligase reaction is similar to that of mRNA-guanylyl transferase, the enzyme that catalyzes one reaction, the 5′ capping of eukaryotic mRNAs. This reaction proceeds in two steps. First, a lysine residue of the enzyme attacks the α-phosphate of GTP, yielding a covalent enzyme–GMP intermediate. GMP is then transferred to the 5′ diphosphate terminus of RNA (Shuman, 1996).

The structure of the ATP-dependent bacteriophage T7 DNA ligase comprises two domains with a deep cleft running between them. The nucleotide binding pocket is located on the larger N-terminal domain, at the bottom of the cleft between the two domains. It was suggested that DNA also binds in this cleft (Subramanya *et al*. 1996b).

DNA ligases show a large variation in sequence and molecular weight, but it is clear from sequence alignments that the smaller enzymes (like the T7 ligase) share a common core structure (Kletzin, 1992).

The structure of a mRNA capping enzyme (Håkansson *et al*. 1997) has a topology similar to DNA ligase, and the overall three-dimensional structure is quite similar as well, particularly in the region of the NTP-binding sites. As mentioned above, the mechanism for phosphoryl transfer in this family of nucleotidyltransferases involves the formation of a GMP-protein adduct, from which the GMP moiety is transferred to the mRNA (Håkansson *et al*. 1997). The reaction involves a large conformational change between an ‘open’ and a ‘closed’ conformation of the capping enzyme (Håkansson *et al*. 1997).

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

The ATP-dependent ligases, GTP-dependent capping enzymes, and probably also the RNA ligases and tRNA ligases, constitute a superfamily of covalent nucleotidyl transferases containing several conserved sequence motifs (Subramanya *et al*. 1996b).

3.2 *Class II aminoacyl tRNA synthetases and biotin synthetases*

The aminoacyl tRNA synthetases, aaRS, are the crucial elements for the translation of the genetic code from a nucleic acid into a protein code (Cusack, 1995, 1997; Arnez & Moras, 1997). They are in fact the only components of the translation machinery that have to really 'know' the genetic code. They recognize tRNA specific for each amino acid and charge it with the corresponding amino acid by forming an ester between the amino acid and a ribosylic hydroxyl group of the adenine at the 3′-CCA end of the tRNA. Since this represents an endergonic reaction, it is coupled to the consumption of ATP. Thus, the amino acid is first activated by forming an aminoacyl-phosphoryl anhydride, namely the aminoacyl adenylate, with release of pyrophosphate. The second reaction is the formation of the aminoacyl-tRNA ester from the aminoacyl adenylate and tRNA, which does not require energy. The proper loading of the tRNA with the correct cognate amino acid is, besides the proper recognition between codon and anticodon, crucial for the fidelity of translation.

More than 20 different aaRS can be found in each organism, at least one for each amino acid, and they come in many different flavours, being monomers, hetero- and homodimers and higher oligomers. The sequences also show very little sequence conservation. In fact, it took the elucidation of several three-dimensional structures before it was realized that there are two general types, called the class I and class II synthetases. Each of these has a set of more or less invariant residues and a conserved core structure containing the enzymatic machinery. However, large variations of this structural core can be seen after structures covering 14 (Cusack, 1995, 1997) different amino acid specificities have been determined. Some of the aaRS structures have been determined from more than one organism, and may be homologous in structure but are nevertheless specific for the particular phylum (Delarue *et al*. 1994). Surprisingly, the active site machinery is different between the two classes, which suggests that they are evolutionary unrelated. The duality of the structure types can be correlated with the biochemistry of the system in that the type I aaRS put the activated amino acid on the 2′ OH, whereas the type II enzymes, except PheRS, put it onto the 3′ position of the 3′-terminal ribose of tRNA (Erianie *et al*. 1990). This was a surprising finding since the different loading position of different amino acids had been noted very early on, but could not be correlated with any other feature of aaRS. Interestingly, PheRS structurally belongs to the class II synthetases but functionally to class I (Goldgur *et al*. 1997). PheRS is one of the most complex enzymes in the aaRS family; it has a domain structure $(\alpha\beta)_2$, where both α and β comprise the characteristic fold of the class II catalytic domains, but only the α subunit is enzymatically active.

The core structure of class II aaRS, as first seen with the structure of SerRS (Cusack *et al*. 1990) and later also with AspRS (Delarue *et al*. 1990; Ruff *et al*. 1991), LysRS (Onesti *et al*. 1995), HisRS (Arnez *et al*. 1995; Aberg *et al*. 1997), GlyRS (Logan *et al*. 1995), PheRS (Mosyak *et al*. 1995; Goldgur *et al*. 1997), AsnRS (Berthet-Colominas *et al*. 1998;) and ProRS (Cusack *et al*. 1998) forms an antiparallel β-sheet surrounded on both sides by helices (Fig. 1*q*). In contrast to this, the evolutionary unrelated class I enzymes are classified as having an

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

ATP pyrophosphatase fold and are therefore discussed in a different section. The structures of the enzymes are modular in that they contain in addition to the enzymatic core additional structural modules which are specific for certain subclasses. Some aaRS (SerRS, ThrRS, ProRS, HisRS and some GlyRS) form a IIa subclass and contain (except SerRS) an additional C-terminal subclass-specific domain, whereas members of subclass IIb (AsnRS, AspRS and LysRS) contain a homologous N-terminal domain (Cusack, 1995). Members of subclass IIc (PheRS, AlaRS and some GlyRS) are not homodimers like IIa and IIb but have a different domain composition.

Structures of different enzymes with substrates, substrate analogues and with the aminoacyl-adenylate intermediate have been determined; thus a large amount of information on binding mode, specificity and catalytic mechanism has become available. The mode of ATP binding is conserved in different type II aaRS (Cavarelli *et al*. 1994; Belrhali *et al*. 1994, 1995). The nucleotide binds along the antiparallel sheet, its major contact site being formed by two strands with conserved residues. Unlike for class I enzymes and for many other nucleotide binding proteins covered in this review, the phosphatases are not in an extended but a rather compact conformation, where the γ-phosphate folds back over the adenine base. The base moiety stacks against an invariant Phe of the protein. The bent conformation of ATP is stabilized by one of the invariant arginines. Biochemical studies and structural considerations have provided evidence for a direct in-line attack of the carboxylate ion on the α-phosphate with a presumed penta-coordinate intermediate. Presumably, in class II enzymes an invariant arginine is responsible for stabilization of the transition state of the phosphoryl transfer reaction, along with a metal ion found between the α- and β-phosphate. In addition, one or two metal ions have been found between the β- and γ-phosphates of Asp-RS and Ser-RS, respectively, where they are believed to be required for stabilisation of the pyrophosphate leaving group. In an interesting twist to the catalytic mechanism, HisRS has been shown to contain an additional arginine at a position corresponding to that of the catalytic metal ion found in the SerRS-ATP complex, reinforcing the notion that the bidentate binding mode of a metal ion can be replaced by the bidentate binding mode of the guanidino group of arginine (Arnez *et al*. 1997).

3.3 *ATP pyrophosphatase fold*

A widespread fold of enzymes featuring an adenylated intermediate is the so-called ATP pyrophosphate domain. This domain is defined by a highly conserved sequence motif called the 'PP-motif', which appears to be a modified version of the P-loop. This fingerprint motif with the sequence SGGXDS/T interacts with the phosphates of the bound nucleotide (Bork & Koonin, 1994) and is located between a β-strand and an α-helix like the original P-loop. The fold of this domain is a twisted five-stranded β-sheet with the order 32145 sandwiched between two layers of helices (Murzin *et al*. 1995). The two known prototypes of this fold are the central domain of the GMP synthetase (Fig. 1*r*) (Tesmer *et al*. 1996) and the NAD synthetase (Rizzi *et al*. 1996). The term 'N-Type' pyrophosphatase fold' (N = amino) has been coined to indicate a subfamily of enzymes that adenylate their substrates to catalyze amination reactions (Bork & Koonin, 1994). GMP synthetase belongs to the glutamine amidotransferases, which link cellular nitrogen metabolism to a variety of biosynthetic pathways, catalyzing the transfer of ammonia from the side chain of glutamine to at least 15

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

different acceptor substrates (Tesmer *et al*. 1996). The enzyme consists of two different modules, the amidotransferase and the synthetase domain, the latter being different for each substrate. The GMP synthetase fold resembles the classic dinucleotide-binding motif of dehydrogenases, except that the sixth and the last β-strands are missing. The reaction catalyzed is the adenylation of XMP by attack on the α-phosphate of ATP, to generate PP_i (ATP pyrophosphatase reaction) followed by the reaction of the adenyl-XMP with glutamine to GMP, glutamate and AMP. Interestingly, the reaction mechanism of glutamine synthetase (see above) also involves the activation of a carboxylate group (of glutamic acid) by ATP, but in this case, the β,γ-phosphate bond of ATP is attacked, the phosphoryl group instead of an adenylate is attached to the carboxyl group, and the incoming ammonia replaces the phosphate group to form glutamine and inorganic phosphate (Tesmer *et al*. 1996).

It was suggested that asparagine synthetase and argininosuccinate synthetase also belong to the structural class of N-type ATP pyrophosphatases (Bork & Koonin, 1994). Indeed they catalyze a similar biochemical reaction and contain the modified P-loop fingerprint motif. However, the structure of asparagine synthetase has recently been solved (Nakatsu *et al*. 1998), and no structural homology to the pyrophosphatase fold has been observed. In contrast, it turned out that the structure of asparagine synthetase is remarkably similar to the catalytic domain of yeast class II aspartyl-tRNA synthetase despite the low sequence similarity. Most of the active site residues are conserved between the two enzymes and they have a similar reaction mechanism, suggesting a common ancestor (Nakatsu *et al*. 1998).

Recently, the structure of phosphoadenyl sulfate (PAPS) reductase from *E. coli* has been published (Savage *et al*. 1997). This enzyme, together with the related enzyme ATP sulfurylase, is involved in sulfate metabolism and the cysteine biosynthesis pathway of prototrophic organisms such as bacteria, fungi, algae, and plants which assimilate inorganic sulphate. The fold of PAPS reductase is closely related to the N-type ATP pyrophosphatase fold. Therefore, a new family consisting of N-type ATP pyrophosphatases, PAPS reductases and the closely related ATP sulfurylases has been proposed, named adenine nucleotide α-hydrolases (Holm, 1998). Indeed, PAPS reductase, which catalyzes the reduction of phosphoadenylylsulphate (PAPS) to sulphite and phospho-adenosine-phosphate (PAP) using thioredoxin, can be described as a 'nucleotide binding protein'. It contains a modified version of the PP-motif of the N-type ATP pyrophosphatases (Savage *et al*. 1997) with only one glycine residue. The PAPS reductases and the ATP sulfurylases share about 25–30% sequence identity and are expected to have a similar fold. However, ATP sulfurylases, which are ATP pyrophosphatases, catalyze a reaction in which the energy of the ATP is coupled to the synthesis of a P–S bond. ATP sulfurylases are thus proposed to form a link between PAPS reductases and the 'classical' ATP pyrophosphatases (Savage *et al*. 1997).

It has been argued that both CoA ligases and class I tRNA synthetases belong to the broader class of ATP-PPases (Bork & Koonin, 1994). They also use adenylated intermediates but, in contrast to the N-type ATP-PPases, displace AMP by sulfur or oxygen nucleophiles and should correspondingly be called 'O/S-Type' pyrophosphatases. The phosphate-binding motifs of the CoA ligases and class I tRNA synthetases are different from the 'N-type' ATP-pyrophosphatases and they bind ATP in a different way (Tesmer *et al*. 1996; Perona *et al*. 1993). In spite of these differences, the fold of class I tRNA synthetases is indeed related to the ATP pyrophosphatases and PAPS reductases. Class I tRNA synthetases catalyze the adenylation of amino acids and transfer the activated amino acid onto the 2′-OH of the terminal adenosine of the tRNA as described above. Structures of GlnRS (Rould *et al*. 1989;

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Perona *et al*. 1993), TyrRS (Brick *et al*. 1989; Leatherbarrow *et al*. 1985), MetRS (Brunie *et al*. 1990) and more recently TrpRS (Doublié *et al*. 1995), GluRS (Nureki *et al*. 1995) and IleRS (Nureki *et al*. 1998) have been determined. The structure of the active site domain shows a Rossmann type fold typical for NADP dependent enzymes and is totally different from that of class II enzymes. Furthermore, the amino acids which are invariant among class I RS' and involved in catalysis, the HIGH and KMSKs motifs, are different from those of the class II enzymes. Stabilization of the transition state at the α-phosphate involves a lysine and histidine from each of the fingerprint motifs, whereas in class II, an arginine residue is involved.

Cytidylyltransferases were predicted to have a similar fold as class I aminoacyl tRNA synthetases (Bork *et al*. 1995). Very recently, the structure of CTP:glycerol-3-phosphate cytidylyltransferase (GCT) from *B. subtilis* has been solved, and the prediction turned out to be correct (Weber *et al*. 1999). GCT is most closely related to glutaminyl tRNA synthetase and shows conserved regions in the active site corresponding to the HIGH and KMSKS motifs in the tRNA synthetases so that a divergent evolution is likely.

3.4 *Luciferases*

Luciferases are a highly diversive family of enzymes, catalyzing a variety of reactions using widely different substrates (Baldwin, 1996). All have in common the involvement of oxygen and the emission of light, although by very different means and with different structures. It is therefore likely that luciferases from different species evolved independently, rather than from a common ancestral enzyme (Baldwin, 1996).

Firefly luciferase uses an oxidative reaction involving ATP, firefly luciferin and molecular oxygen, yielding an excited oxyluciferin species which emits light. In more detail, firefly luciferase uses ATP to activate a carboxylic substrate (luciferin) by adenylation and pyrophosphate release (Baldwin, 1996; Conti *et al*. 1996) like the ATP pyrophosphatases and tRNA synthetases. This is followed by a reaction of the activated carboxylate with molecular oxygen, release of AMP and production of the excited state intermediate. This excited state then decays to the ground state oxyluciferin, emitting yellow-green light.

Firefly luciferase shows extensive sequence similarity to a number of enzymes that use ATP to form adenylate intermediates. Closely related enyzmes are acyl-coenzymeA ligases and peptide synthetases. The ligases activate various substrates for their transfer to coenzymeA, whereas the peptide synthetases activate amino acids in the non-ribosomal syntheses of polypeptides (Baldwin, 1996).

Firefly luciferase consists of two domains of different size. The large N-terminal domain consists of a β-barrel and two β-sheets flanked by α-helices to form a five-layered $\alpha\beta\alpha\beta\alpha$ structure. The C-terminal domain consists of five β-strands and three α-helices that fold into a compact structure. This arrangement is known as 'anvil and hammer' motif (Baldwin, 1996). The active site probably is located between the two domains (Baldwin, 1996).

The structure of firefly luciferases bears no resemblance to the structure of class I or class II tRNA synthetases.

In contrast to firefly luciferase, the bacterial luciferases do not bind ATP but use FMN to activate molecular oxygen, resulting in an excited state of a flavin derivative whose return to the ground state involves light emission. All bacterial luciferases studied so far appear to be

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

homologous and catalyze the same reaction (Baldwin *et al*. 1995), and their structure is vastly different from the firefly luciferases, namely two domains consisting of an eight-stranded β/α TIM barrel motif ($[\beta/\alpha]_8$) each (Baldwin *et al*. 1995).

3.5 *Nucleic acid polymerases*

Nucleic acid polymerases use NTPs to transfer a mononucleotide to the growing polynucleotide chain. The reaction is a nucleophilic attack on the α-phosphate of the NTP by the 3′-OH end of the DNA or RNA. A phosphodiester bond is formed and PP_i is released. The driving force might be increased by the subsequent hydrolysis of PP_i. Different types of polymerases are DNA-directed DNA polymerases (DNAP), reverse transcriptases (RT), RNA-directed RNA polymerases (RNAP) and DNA-directed RNAPs (Delarue *et al*. 1990; Sousa, 1996). A superfamily of nucleic acid polymerases was proposed (Sousa, 1996) on the basis of the structures of the HIV-1 reverse transcriptase (Arnold *et al*. 1992; Kohlstaedt *et al*. 1992; Jacobo-Molina *et al*. 1993), DNA polymerase I from *E. coli* (Klenow fragment) (Ollis *et al*. 1985) and the T7 RNA polymerase (Sousa *et al*. 1993). This superfamily is likely to include the pol I and pol α classes of DNAPs, mitochondrial and phage DNA-directed RNA polymerases and most RNA-directed polymerases, based on the occurrence of conserved sequence motifs. Today, structures of four families of DNA polymerases have been solved: RB69gp43 from the pol α or 'B' family (Wang *et al*. 1997); Klenow fragment (KF) (Ollis *et al*. 1985), Taq polymerase (Kim *et al*. 1995; Eom *et al*. 1996), *E. coli* phage T7 polymerase (Doublié *et al*. 1998), KF analogues from *Bacillus stearothermophilus* (Kiefer *et al*. 1997) and *Thermus aquaticus* from the pol I family (Kim *et al*. 1995; Korolev *et al*. 1995); rat (J. F. Davies *et al*. 1994; Pelletier *et al*. 1994; Sawaya *et al*. 1994) and human (Pelletier *et al*. 1996) DNA pol β from the terminal transferase family (Holm & Sander, 1995); HIV-1 RT (Kohlstaedt *et al*. 1992; Jacobo-Molina *et al*. 1993; Ren *et al*. 1995; Rodgers *et al*. 1995) and a fragment of Moloney murine leukaemia virus RT from the reverse transcriptase family (Georgiadis *et al*. 1995). The similarities of shape in the structurally characterized enzymes lead to the description of a 'cupped right hand' with the three subdomains 'fingers', 'palm' and 'thumb', thus forming a DNA-binding groove that leads to the polymerase active site (Doublié *et al*. 1998). The palm forms the base of the active site by contributing two highly conserved acidic residues which are important for metal binding (Fig. 1*s*). The fingers close the binding site by making numerous contacts with the bound nucleotide and undergo a substrate-induced conformational change that may be important for catalysis (Brautigam & Steitz, 1998).

It now can be concluded that the 'palm' domains of pol I and pol α enzymes are structurally related to each other and to RT and RNAP. In contrast, the corresponding domain of the polymerase β structure, although it retains some of the modular features of the polymerase active site, is not structurally related to the others (Brautigam & Steitz, 1998). Interestingly, the fingers and thumb domains in all four families are completely unrelated; nevertheless, their function is similar in all polymerase families. The proposed universal mechanism for polymerases uses two divalent metal ions where ion A facilitates the attack of the 3′-OH on the α-phosphate and metal ion B assists leaving of pyrophosphate, with both metal ions stabilizing the penta-coordinate transition state (Steitz & Steitz, 1993; Steitz, 1998). The topology of the two types of polymerases is different, but the acidic residues,

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

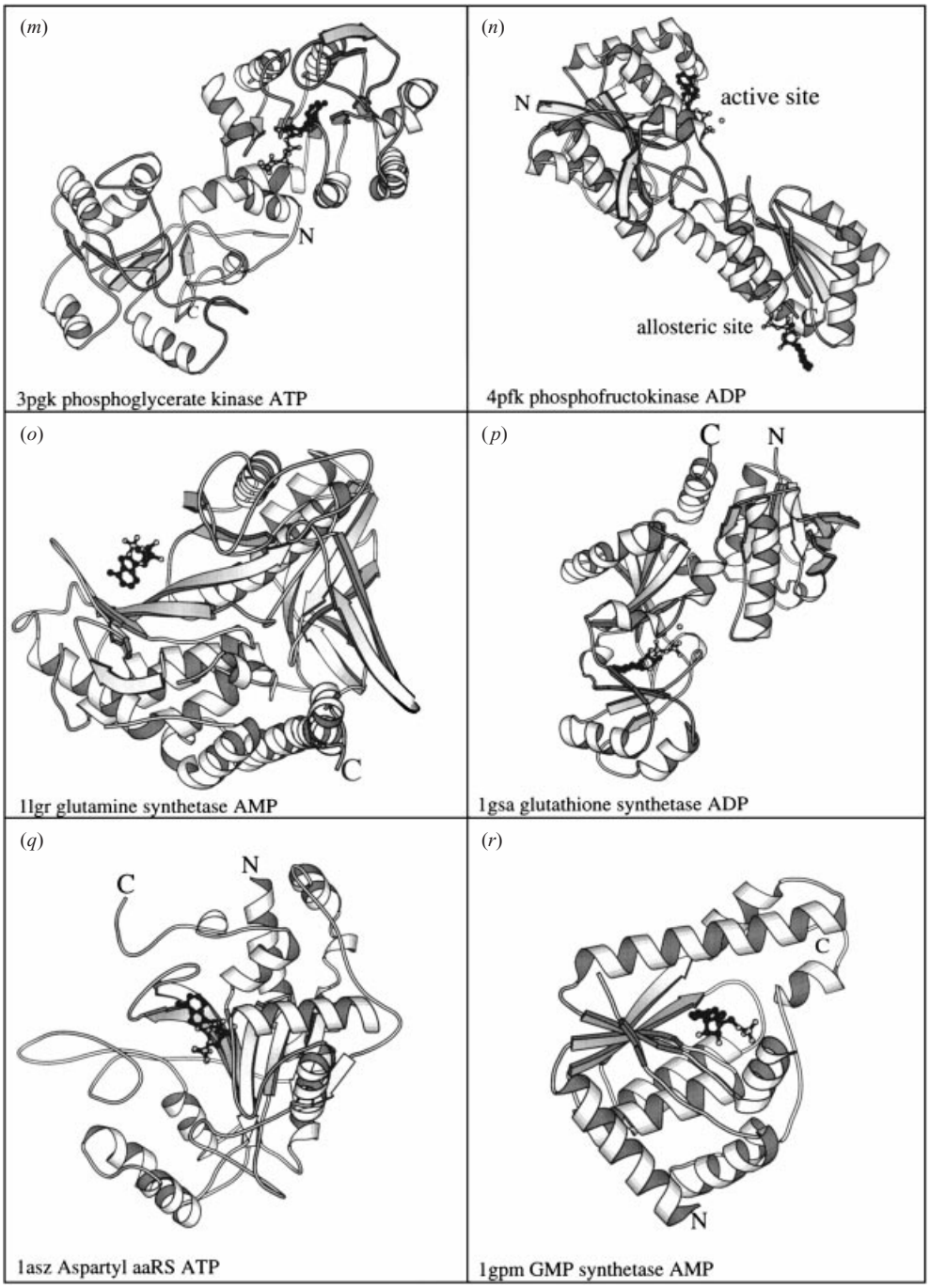


Fig. 1 (*cont.*). MOLSCRIPT plots of representative examples of nucleotide binding proteins. (*m*) Phosphoglycerate kinase from *S. cerevisiae*, residues 1–415. (*n*) Phosphofructokinase from *Bacillus stearothermophilus* with two ADP molecules bound, residues 1–320. Only one monomer is shown, the active enzyme is a tetramer. The allosteric ADP is bound by residues from two subunits (second subunit not shown). (*o*) Glutamine synthetase from *Salmonella typhimurium*, residues 1–468. Several regions were not visible in the electron density and are therefore omitted from the PDB file. (*p*) Glutathione synthetase from *E. coli*, residues 1-314. (*q*) Aspartyl-aminoacyl-tRNA synthetase (aaRS, residues 255-557).

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

mostly aspartic acids, constituting the metal ion binding sites are conserved and can be superimposed structurally (Wang *et al*. 1997; Brautigam & Steitz, 1998). Thus, we have to assume convergent evolution for the two types of polymerases. The positional conservation of the functionally important groups lead to a proposal for a conserved mechanism of the phosphoryltransfer reaction involving two metal ions (Steitz, 1998). It is interesting to note that GSHase also uses two divalent metal ions for catalysis although the details of the reaction are quite different.

Interestingly, the 'palm' domain has a ferredoxin-like fold (Murzin *et al*. 1995), and the catalytic domain of DNA polymerase I (Klenow fragment) (Ollis *et al*. 1985) has a similar topology and nucleotide binding site as the actin-like domains (Branden & Tooze, 1991). However, the similarities are merely topological, the three-dimensional structures are quite different (Branden & Tooze, 1991).

The fold of the catalytic fragment of DNA polymerase β positions it in the superfamily of nucleotidyltransferases which also comprises kanamycin nucleotidyltransferase (Holm & Sander, 1995). Kanamycin nucleotidyltransferase was shown to catalyze the transfer of a nucleoside monophosphate group from a nucleotide to the 4′-hydroxyl group of kanamycin (Sakon *et al*. 1993). This enzyme can utilize ATP, GTP or UTP and can inactivate a wide range of aminoglycosides. The molecule consists of two domains. The N-terminal domain consists of a five-stranded mixed β-sheet and four α-helices, whereas the C-terminal domain is purely α-helical. The C-terminal domain has the classical fold of up-and-down helical bundles. Other proteins with this fold include, for example, myohemerythrin, cytochrome b562, cytochrome c′ and apolipophorin-III (Sakon *et al*. 1993). The enzyme probably functions as a dimer with the active site build by the interface (Pedersen *et al*. 1995). The N-terminal catalytic domain of kanamycin nucleotidyltransferase is very similar to the palm domain of DNA polymerase β, and since several active site features are conserved, a common catalytic mechanism was suggested (Holm & Sander, 1995).

The adenylyl cyclase catalytic core domain shows surprising similarity to the 'palm' domain of the polymerase I family of prokaryotic DNA polymerases, including *E. coli* DNA polymerase I and *Thermus aquaticus* (Taq) polymerase (Kim *et al*. 1995; Eom *et al*. 1996; Bryant *et al*. 1997). Adenylyl cyclases catalyze the synthesis of cAMP from ATP, whereas guanylyl cyclases catalyze a similar reaction with GTP. cAMP and cGMP are second messenger molecules generated by signal transduction pathways involving G protein coupled receptors. These receptors are membrane proteins with seven transmembrane helices. The N-adenylyl cyclases, like guanylyl cyclases and DNA polymerases, catalyze the Mg^{2+}-dependent in-line attack of a ribose or deoxyribose 3′-hydroxyl on the α-phosphate of a nucleoside triphosphate, cleaving off PP_i as a result (Bryant *et al*. 1997). Nine mammalian isoforms of adenylyl cyclases have been described all of which are stimulated by the α-subunit of heterotrimeric G_s proteins. Each member is individually regulated by G_i protein α-subunits, other G protein subunits, forskolin, Ca-calmodulin and phosphorylation. The enzyme consists of 12 transmembrane helices and two cytosolic regions which can each be subdivided

---

The N-terminal domain (residues 68–245) was omitted for clarity. In Class IIb enzymes (Lys,Asp,Asn aaRS), this anticodon-binding N-terminal domain has a β-barrel shape. In contrast, most of Class IIa enzymes (Pro, Thr, His, Gly but not Ser aaRS) have a domain with a similar function at the C-terminus. (*r*) GMP synthetase from *E. coli*, central domain (residues 210–416). The glutamine amidotransferase domain (N-terminal residues 3–209) and the C-terminal dimerization domain (residues 417–525) are omitted for clarity.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

into two domains, an N-terminal catalytic and a C-terminal regulatory domain, respectively. The recombinant catalytic domains can be expressed individually and can form dimers, or they can be tethered together to form a highly active enzyme (Tang & Gilman, 1995). The structures of a homodimeric enzyme with low activity (Yan *et al*. 1996; Zhang *et al*. 1997) and of a tethered heterodimer with high activity have been determined (Tesmer *et al*. 1997). The structures showed the catalytic core domain to be similar to the core domain of DNA polymerases I, leading to the speculation that they share a common ancestor (Bryant *et al*. 1997). The active site of adenylyl cyclase has been elucidated using the inhibitor 2′3′dAMP and showed the presence of one metal ion coordinated by two acidic residues and an invariant arginine in the active site (Tesmer *et al*. 1997). The resemblance between the two enzyme families thus includes the presence of two (cyclase) out of three (polymerases) acidic Mg^{2+}-binding groups and the presence of at least one metal ion versus two in polymerases. It was suggested that the role of the second metal in stabilizing the transition state could be taken over by the invariant arginine in adenylyl cyclases (Tesmer *et al*. 1997).

3.6 *MutT proteins, PP_i releasing NTPases*

The MutT family proteins are NTP hydrolyzing pyrophosphatases which appear to have a 'housecleaning' function in cells (Bessman *et al*. 1996). They share a small region of homology, and they all hydrolyze nucleoside diphosphates which are linked to some other moiety X, where X can be a phosphate, a sugar or another nucleotide, for example. Thus, the actual substrates are widely different including nucleoside triphosphates, coenzymes, nucleotide sugars, and dinucleotide polyphosphates (Bessman *et al*. 1996).

A common feature of all these substrates seems to be that they all can be potentially hazardous to, or endanger, important equilibria in the cell (Bessman *et al*. 1996). Several putative members of the family have been identified as ORFs in the *E. coli* genome (Bessman *et al*. 1996).

The only member of the family whose structure is known is the small antimutator protein MutT (129 residues), which has been proposed to specifically prevent AT-CG transversions by hydrolyzing a mutagenic form of dGTP, 8-oxo-dGTP (Lin *et al*. 1997). This oxidatively modified nucleotide can mispair with an adenine in the template, causing potentially deleterious mutations. The MutT protein decreases the mutation rate in *E. coli* by several orders of magnitude (Frick *et al*. 1995). The NMR structure revealed a β-grasp motif, i.e. a single helix packs against a mixed β-sheet like in ubiquitin (Fig. 1*t*) (Murzin *et al*. 1995). The nucleotide probably binds in a hydrophobic cleft near three strands of a mixed β-sheet (Lin *et al*. 1997). MutT catalyzes the hydrolysis of all eight canonical NTPs (with strong preference for dGTP) by nucleophilic substitution of the rarely attacked β-phosphorus. Substitutions at the α- and γ-phosphate are quite common, occurring in a number of enzymes as described above (Cohn, 1959). In contrast, substitutions at the β-phosphate are rare, and it has been speculated that this might be due to the higher electron density at this position (Mildvan, 1970). The only other known enzymes that catalyze such a reaction are six synthetases, namely phosphoribosyl-pyrophosphate synthetase, thiamine pyrophosphokinase, dihydropteridine pyrophosphokinase, 3-pyrophosphorylguanosine-5′-diphosphate synthetase, phosphoenol-pyruvate synthetase and pyruvate-orthophosphate dikinase (Frick *et al*. 1995; Weber *et al*. 1992). It has been shown by ^{31}P NMR spectroscopy that in the MutT-catalyzed reaction the

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

bond between the β-phosphorus atom and the α,β-bridging oxygen of dGTP is cleaved such that the site of nucleophilic substitution is at the β-phosphorus (Weber *et al*. 1992). The nucleophile has been suggested to be a water or hydroxyl ligand of the enzyme-bound divalent cation, assisted by the second, nucleotide-bound divalent cation which might neutralize charge at the β-phosphoryl group. A neighboring lysine might assist in the departure of the leaving group (Lin *et al*. 1997).

3.7 *dUTPase*

Another enzyme with a physiological function roughly similar to the MutT proteins is dUTPase. One of its functions is to keep the concentration of dUTP at a low level, which reduces the incorporation of uracil into DNA. Lower dUTPase activity leads to a higher mutation frequency. The ubiquitous enzyme dUTPase catalyzes specifically the hydrolysis of dUTP to dUMP and pyrophosphate (Larsson *et al*. 1996). It functions to regulate the cellular pool of dUMP that is an essential precursor in the synthesis of thymidine nucleotides. dUTPases from different sources are monomeric, dimeric or trimeric. The known structure from *E. coli* shows a trimeric enzyme with the three active sites composed of residues from the interfaces of the three monomers, which are predominantly β-pleated sheets with one α-helix at the interface towards an adjacent monomer (Fig. 1*u*) (Cedergren-Zeppezauer *et al*. 1992; Larsson *et al*. 1996). A conserved glycine-rich motif with the sequence RGxxGFGST in the dUTPase family resembles the P-loop. Unfortunately, in the crystal structure this motif is disordered in spite of the presence of the nucleotide, since it is located in the flexible C-terminal region. The high specificity of dUTPase for its substrate is achieved by interactions of the base moiety with the protein backbone and an asparagine sidechain. Contacts to the phosphates are less numerous than in other nucleotides, with interactions seen mainly between the α-phosphate and water molecules. The catalytic mechanism involves a magnesium ion, which is not seen in the crystal structure. Thus, the dUTPase-dUDP complex in the solved structure is probably very different from the catalytically active dUTP complex which presumably has a lass mobile C-terminal region and would then be able to bind the magnesium ion (Larsson *et al*. 1996).

3.8 *Apyrases*

ATP diphosphohydrolases, also called apyrases, are ATP-hydrolyzing enzymes of unknown metabolic function. Their vast abundance and high enzymatic activity suggest that apyrases might play an important role in cell metabolism. Apyrases are different from ATPases in several respects. ATPases are enzymes of high substrate specificity, hydrolyzing ATP to ADP and P_i. In contrast, apyrases can hydrolyze different di- and triphosphonucleotides. In the case of ATP hydrolysis, the main products are AMP and PP_i. There is some evidence that apyrases participate in neurotransmission, blood platelet aggregation and pressure regulation, protein glycosylation and sugar level control, as well as regulation of membrane integrity (Komoszynski & Wojtczak, 1996). Apyrases can be divided into groups depending on their localization and solubility. Ecto-apyrases have their catalytic site outside the cell. They are either secreted (e.g. by the salivary glands of most blood-feeding arthropods, where the

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

apyrase prevents blood clotting by hydrolyzing ATP (Valenzuela *et al*. 1998)) or attached to the membrane. Endo-apyrases are less well characterized; they might be involved in the sugar and glycoprotein metabolism inside the Golgi apparatus and the endoplasmic reticulum. This compartmentalization might be sufficient to prevent endo-apyrases from depleting the cellular ATP, but most probably there are additional regulatory mechanisms involved. For example, some apyrases are strictly dependent on Ca^{2+} ions, whereas others can be activated by Mg^{2+}, Mn^{2+} or Zn^{2+} as well (Valenzuela *et al*. 1998).

4. Reaction catalyzed: ATP → Adenosine + PPi + Pi

4.1 *SAM synthetases*

The most common methyl donor S-adenosylmethionine is synthesized by the AdoMet synthetase or S-adenosylmethionine synthetase. The unusual two-step reaction involves the cleavage of the complete triphosphate chain from ATP, which is further hydrolyzed to PP_i and P_i before the final product is released. The initial step is a direct attack of the sulfur of methionine on the C5′ atom of ATP. A lysine close to the active site is proposed to protonate the leaving O5′ atom of ATP. The subsequent triphosphate hydrolysis might involve the activation of a water molecule by another lysine for an in-line attack on the γ-phosphate (Takusagawa *et al*. 1996). The magnesium ion has been shown to be essential for hydrolysis (Markham *et al*. 1980). Thus, the enzyme catalyzes reactions at both ends of the triphosphate chain (Takusagawa *et al*. 1996).

S-adenosylmethionine synthetases consist of three similar intertwined domains with the motif β-α-β(2)-α-β in two layers (Murzin *et al*. 1995). Two three-domain monomers associate to form dimers with two active sites in the interface. Two dimers associate to form the tetrameric enzyme. The adenine ring of the substrate weakly interacts with the hydrophobic β-sheet of the central domain, whereas the phosphates interact strongly with two subunits of the enzyme (Fig. 1*v*). The adenine is recognized by forming a single hydrogen bond between its N6 amino group and a glutamine sidechain. A P-loop-like sequence, **G**GAFS**GK**D (with an aspartate instead of the usual Ser/Thr residue), has been found around the ATP binding region. However, the three-dimensional structure of this region is completely different from the typical P-loops. Most of the residues in the P-loop-like sequence are located relatively far away from the phosphates. The complete phosphate binding region consists of two consecutive β-turns with the sequence FSGK, including an invariant lysine, and DPSK. The first lysine is located near the γ-phosphate, whereas the second lysine is near the ribose moiety. These residues might be involved in the catalysis of the two different cleavage reactions (Takusagawa *et al*. 1996).

5. Reaction catalyzed: ApppA → AMP + ADP

5.1 *HIT-like proteins*

Histidine triad (HIT) proteins are dimeric purine-nucleotide binding proteins which are named for the HIT motif, His-O-His-O-His-OO near the C-terminus, where O is a hydrophobic amino acid. These proteins can be divided into two families, the Hint homologs,

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

which are ubiquitous, and the Fhit homologs, which are found in animals and fungi (Pace *et al*. 1998). The human FHIT gene is involved in the development of several common cancers and encodes an enzyme with diadenosine P1,P3-triphosphate (ApppA) hydrolase activity (Sozzi *et al*. 1996, 1997a,b; Negrini *et al*. 1996; Virgilio *et al*. 1996; Druck *et al*. 1997; Greenspan *et al*. 1997; Barnes *et al*. 1996). The structure of HINT, a member of the first HIT family, revealed a fold similar to the sequentially unrelated Galactose-1-phosphate uridylyltransferase (GalT), whose deficiency causes galactosemia (Brenner *et al*. 1997). HINT is very similar to protein kinase C inhibitor-1 despite not being able to inhibit PKC (Brenner *et al*. 1997). The main structural feature of the HINT dimer is an antiparallel ten-stranded β-sheet with two α-helices packed in the interior of a half-barrel formed by the β-sheets of both monomers (Fig. 1*w*). The two identical nucleotide binding sites which are occupied by either GMP or 8-Br-AMP, respectively in the published structures are located in a cleft on top of the β-sheet, with their 5′-phosphates facing each other, but with a distance of *c*. 22 Å between the phosphorous atoms, making it unlikely that one adenine of ApppA binds to each site. A structure of Fhit with a nonhydrolyzable ApppA anolog bound indeed revealed that two ApppA molecules bind per Fhit dimer, identifying two additional adenosine-binding sites on the dimer surface (Pace *et al*. 1998). It was shown that the HIT motif forms a portion of the two binding sites which were localized first in the HINT structure, and one or several histidines might be involved in the cleavage reaction (Brenner *et al*. 1997). Interestingly, superposition of the GalT and the HINT structure shows the UMP bound in GalT at the same position as one of these mononucleotide binding sites, which lead the authors to propose that a HINT protein was the common ancestor of FHIT, GalT and other nucleotide-binding proteins (Brenner *et al*. 1997). The peculiar way of ApppA binding that displays its highly charged phosphate groups on the protein surface, together with the biochemical data lead Pace *et al*. to propose a role of Fhit in signal transduction, where Fhit signals as an enzyme-substrate complex to a putative anti-proliferative or pro-apoptotic effector and is switched off by cleaving the bound ApppA (Brenner *et al*. 1997).

6. General features

All described nucleoside triphosphate binding proteins are α-β-structures with two exceptions: the ATP binding domain of GroEL is classified as all-α-domain, and dUTPase consists of a trimeric all-β-domain, each of which has one nucleotide binding site. All structures covered in this review – where the nucleotide position is known – have the nucleotide bound in the anti-conformation with three exceptions: PEP carboxykinase (Protein Data Bank identifier 1ayl) and kanamycin nucleotidyltransferase (1kny) where the base adopts the syn-conformation, and the mRNA capping enzyme (RNA guanylyltransferase, 1ckm, 1ckn) where the base moiety is skewed relative to the standard orientation.

A statistical analysis of all bound nucleotides and magnesium ions in the investigated subset of the Brookhaven Protein Database (PDB) (169 entries, database as of 20 February 1998) shows interesting preferences of sidechains in the vicinity of the magnesium ion or of certain atoms of the base moieties (Fig. 2). The selected subset is somewhat redundant; for example, entries were kept if the same protein was crystallized with a different nucleotide bound, so there is certainly some bias, but since only a small subset of nucleotide-binding proteins has been crystallized, it is not possible to get non-biased data. Thus, the numbers

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

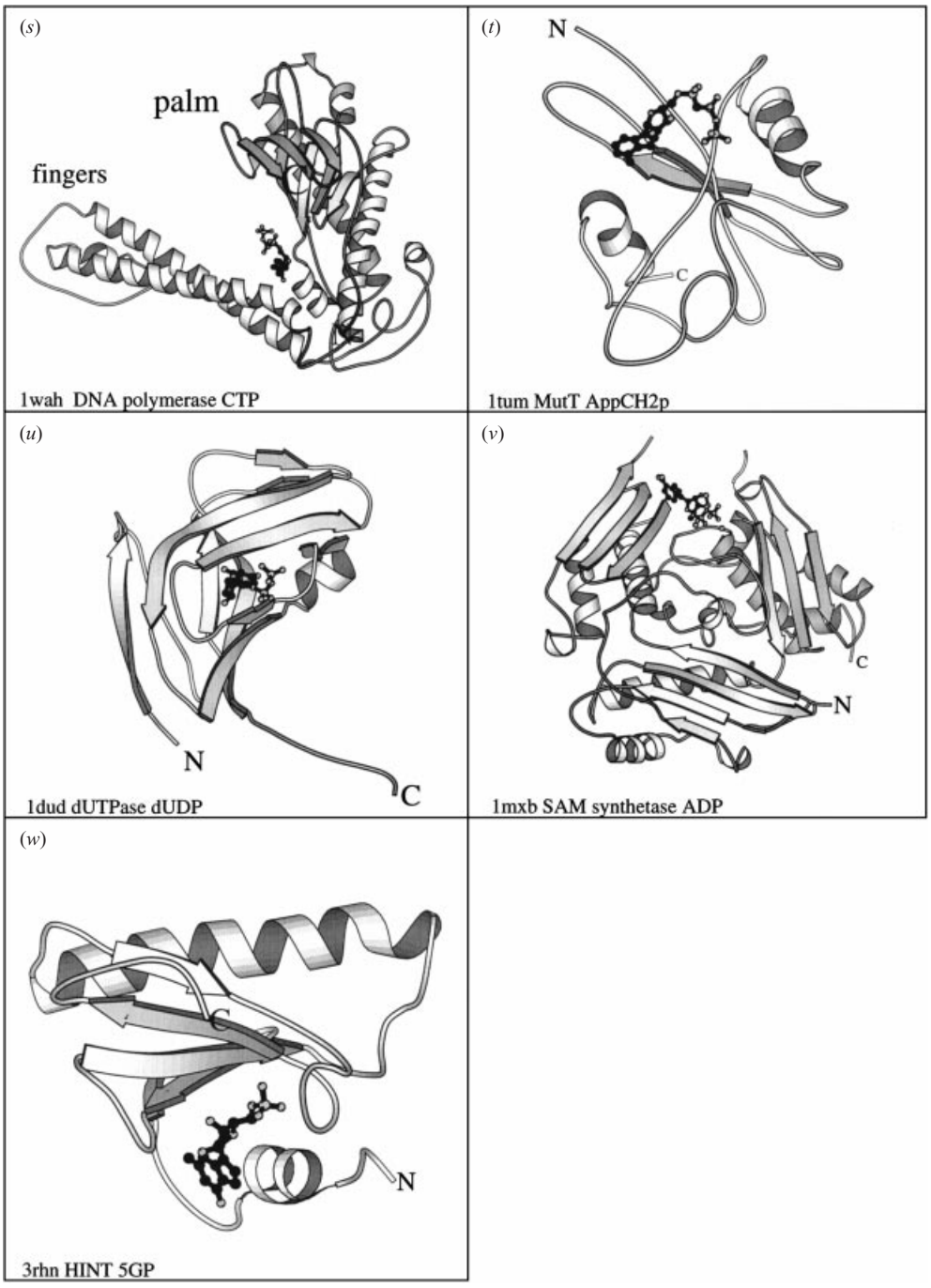


Fig. 1 (*cont.*). MOLSCRIPT plots of representative examples of nucleotide binding proteins. (*s*) DNA polymerase GP43 from bacteriophage RB69 (residues 388-727). The C-terminal 'thumb' domain (residues 728–903) and the N-terminal domains including the 3′-5′ exonuclease (residues 1–387) which is homologous to the Klenow fragment are omitted to show the nucleotide binding site. (*t*) MutT mutator protein from *E. coli* (NMR structure, residues 1–129). (*u*) deoxyuridine-5′-triphosphate nucleotide hydrolase (dUTPase, residues 1–136) from *E. coli* with dUDP bound. The active enzyme is a trimer. (*v*) S-adenosylmethionine synthetase from *E. coli*, residues 1–383. The active enzyme is a tetramer.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

given below should only be taken as showing tendecies, but there are nevertheless interesting points to observe.

The most frequent amino acid whose sidechain is contacting the magnesium ion is aspartate (42 %), followed by threonine (23 %), glutamate (17 %) and serine (10 %). The aspartate seems to be the preferred ligand compared to glutamate in spite of the slightly higher abundance of glutamate in the choosen protein sample (6.1 % compared to 5.0 %). Threonine also seems to be more suitable for magnesium binding in these proteins than serine.

One should bear in mind that magnesium ions are sometimes missing in structures where they should be present (in 58 % of the 169 coordinate sets with bound nucleotides either Mg or Mn is present, in some cases more than one ion per structure), therefore there are less contacts to a magnesium ion than expected. Only very few of these ions make contact only to the protein, not to the nucleotide.

The most frequent contacts of the nucleotide in absolute numbers are the oxygens of the phosphate group(s) (3053 contacts made by 940 oxygens), followed by the ribose oxygens and the nitrogens and oxygens of adenine and guanine, respectively. Besides the contacts of the phosphate oxygens to main chain atoms (50 % of which belong to glycine residues) which account for 43 % of the total number of interactions, the second numerous contacts are to the positively charged sidechains of arginine and lysine followed by threonine, magnesium and aspartate. The phosphates themselves are contacted most frequently by the magnesium ion (44 %), main chain atoms (25 % to 53 % of which – depending on the position α, β or γ – are glycine) and arginine (7 %). The ribose oxygens preferentially interact with main-chain atoms (39 %, 31 % of which are glycines) followed by arginine/lysine and aspartate/glutamate. Here, the preference for interactions with glycine is much less pronounced than for the phosphate oxygens. The determinants for the base specificity, the N6 and O6 atoms of the purine moiety, again show an extreme preference for contacts to main chain atoms (55 %), but in this case only 7 % of these contacts are made by glycines. The N2 atom is contacted preferentially by the side chains of aspartate (50 %) and leucine (13 %). Surprisingly, only 8 % of the contacts are to glutamate, so there seems to be a strong preference for aspartate over glutamate. When looking at the number of contacts to the base per atom, the O6 of the guanine stands out with 7 contacts per atom on average. There are on average 3.3 contacts per phosphate oxygen, 3 contacts per O2′/O3′ atoms of the ribose, 2.9 contacts per N2 of the guanine and N6 of the adenine, respectively, 1.6 per O4′, and 1.2 per phosphate atom.

Although the sample is not representative and is biased towards small soluble proteins that crystallize more easily and that contain a bound nucleotide, some interesting preferences can be seen which are unlikely to be accidental.

7. Outlook

The increasing number of structures of nucleotide triphosphate-binding proteins sheds light on the astonishing variability in the details of structure and function of this class of proteins.

Some breaks result from residue stretches that are not defined in the PDB file. (*w*) Histidine triad nucleotide-binding protein (HINT) from rabbit, residues 12–126.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

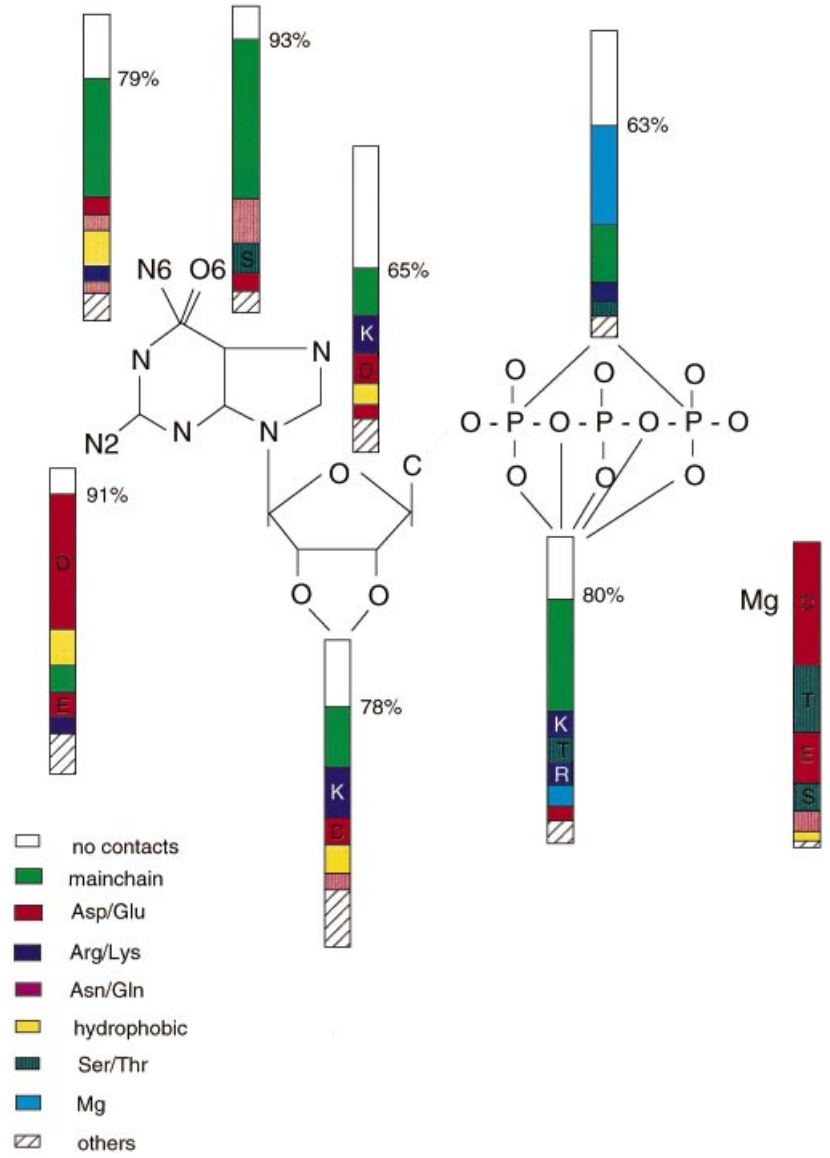


Fig. 2. Classes of amino acids providing the most frequent contacts (< 3.5 Å) to bound nucleotides. Each bar shows the contacts to specific oxygen, nitrogen or phosphate atoms of the nucleotide. The number next to each bar gives the percentage of contacts to each atom. For example, 63% of all phosphate atoms have at least one contact; the others have no contacts at all. The most frequent contacts to the phosphates are to magnesium ions, then to main chain atoms, and a much smaller number to Arg/Lys sidechains, Ser/Thr sidechains and others. In some cases the one-letter abbreviations for certain amino acids have been added to distinguish between, for example, aspartate and glutamate.

The importance of this group is stressed by the fact that some families belong to the most common fold types in the various genomes (Gerstein, 1997). The protein kinases, for example, are among the top ten folds in an eukaryotic genome (*Saccharomyces cerevisiae*). Currently, there are about 338 known protein folds (comprising all proteins, not only the

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

nucleotide binding subset) (Gerstein, 1997). The P-loop containing nucleotide hydrolases are among the top five folds in eukaryotic, eubacterial and archaeal genomes (Gerstein, 1997) and were suggested to belong to a group of five basic folds, namely the TIM-barrel fold, the Rossman fold, the Flavodoxin fold, the Thiamin-binding fold, and the P-loop-hydrolase fold. These folds have a remarkably similar architecture featuring a central sheet of β-strands with helices packed on at least one face and with, almost exclusively, $\beta\alpha\beta$ connections between adjacent strands (Gerstein, 1997). This ancient group might have been present in the last common ancestor of the three major kingdoms and evolved into the variety of proteins we have today. As the number of known structures increases, the picture will become more complete.

8. Acknowledgements

We thank Ilme Schlichting, Jochen Reinstein and Markus Rudolph for critical reading of the manuscript and for discussion, the referee for constructive comments and criticism, and Rita Schebaum for secretarial help.

Note added in proof:

The structures of four complexes of adenylyl cyclase with ATP analogs and various divalent metal ions revealed that this enzyme binds two metal ions in the active site and thus suggests an enzymatic mechanism which is very similar to the mechanism of DNA polymerases, suggesting a common ancestor for both enzymes (Tesmer *et al*. 1999). A new family of ATP-binding proteins was discovered in the course of a structural genomics study by solving the structure of a protein with unknown function from *Methanococcus jannaschii*, MJ0577 (Zarembinski *et al*. 1998). The structure is an open-twisted five-stranded parallel β-sheet with two helices on each side of the sheet. The motifs surrounding the ATP binding site have only limited similarity to the commonly known motifs. Furthermore, the sequential arrangement of these motifs is completely different.

GlnK is a protein which might be involved in the regulation of the nitrogen level in *E. coli*. It is believed that ATP has an important role in this type of proteins. The structure revealed a four stranded β-sheet packed against two helices. Three monomers form a barrel-like core where ATP binds in a cleft between monomers at the junction of two β-sheets. The binding site is formed by two loops, one of which has the highly conserved sequence TGXXGDGKIP which resembles a P-loop.

9. References

Abele, U. & Schulz, G. E. (1995). High-resolution structures of adenylate kinase from yeast ligated with inhibitor Ap_5A, showing the pathway of phosphoryl transfer. *Protein Science* **4**, 1262–1271.

Aberg, A., Yaremchuk, A., Tukalo, M., Rasmussen, B. & Cusack, S. (1997). Crystal structure analysis of the activation of histidine by Thermus thermophilus histidyl-tRNA synthetase. *Biochemistry* **36**, 3084–3094.

Abrahams, J. P., Leslie, A. G. W., Lutter, R. & Walker, J. E. (1994). Structure at 2.8 Å resolution of F_1-ATPase from bovine heart mitochondria. *Nature* **370**, 621–628.

Adams, J. A. & Taylor, S. S. (1992). Energetic limits of phosphotransfer in the catalytic subunit of cAMP-dependent protein kinase as measured by viscosity experiments. *Biochemistry* **31**, 8516–8522.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Admiraal, S. & Herschlag, D. (1995). Mapping the transition state for ATP hydrolysis: implications for enzymatic catalysis. *Chemistry & Biology* **2**, 729–739.

Ahmadian, M. R., Mittal, R., Hall, A., & Wittinghofer, A. (1997). Aluminum fluoride associates with the small guanine nucleotide binding proteins. *FEBS Lett*. **408**, 315–318.

Al-Karadaghi, S., Aevarsson, A., Garber, M., Zheltonosova, J. & Liljas, A. (1996). The structure of elongation factor G in complex with GDP – conformational flexibility and nucleotide exchange. *Structure* **4**, 555–565.

Almassy, R. J., Janson, C. A., Hamlin, R., Xuong, N. H. & Eisenberg, D. (1986). Novel subunit-subunit interactions in the structure of glutamine synthetase. *Nature* **323**, 304–309.

Anderson, C. M., Stenkamp, R. E., Steitz, T. A. (1978). Sequencing a protein by X-ray crystallography. II. Refinement of yeast hexokinase B coordinates and sequence at 2.1 ångstroms resolution. *J. Mol. Biol*. **123**, 15–33.

Anderson, C. M., Zucker, F. H. & Steitz, T. A. (1979). Space–filling models of kinase clefts and conformation changes. *Science* **204**, 375–380.

Aravind, L, Galperin, M. Y. & Koonin, E. V. (1998). The catalytic domain of the P-type ATPases has the haloacid dehalogenase fold. *TIBS* **23**, 127–129.

Arnez, J. G. & Moras, D. (1997). Structural and functional considerations of the aminoacylation reaction. *Trends Biochem. Sci*. **22**, 189–232.

Arnez, J. G., Harris, D. C., Mitschler, A., Rees, B., Francklyn, C. S. & Moras, D. (1995). Crystal structure of histidyl-tRNA synthetase from *E. coli* complexed with histidyl-adenylate. *EMBO J*. **14**, 4143–4155.

Arnez, J. G., Augustine, J. G. Moras, D. & Francklyn, C. S. (1997). The first step of aminoacylation at the atomic level in histidyl-tRNA synthetase. *Proc. Natl. Acad. Sci. USA* **94**, 7144–7149.

Arnold, E., Jacobo-Molina, A., Nanni, R. G., Williams, R. L., Lu, X., Ding, J., Clark Jr., A. D., Zhang, A., Ferris, A. L., Clark, P., Hizi, M. & Hughes, S. H. (1992). Structure of HIV-1 reverse transcriptase/DNA complex at 7 Å resolution showing active site locations. *Nature* **357**, 85–89.

Artymiuk, P. J., Grindley, H. M., Kumar, K., Rice, D. W. & Willett, P. (1993).Three-dimensional structural resemblance between the ribonuclease H and connection domains of HIV reverse transcriptase and the ATPase fold revealed by graph theoretical techniques. *FEBS* **324**, 15–21.

Artymiuk, P. J., Poirrette, A. R., Rice, D. W. & Willett, P. (1996). Biotin carboxylase comes into fold. *Nature Struct. Biol*. **3**, 128–132.

Auer, M., Scarborough, G. A. & Kühlbrandt, W. (1998). Three-dimensional map of the plasma membrane H^+ ATPase in the open conformation. *Nature* **392**, 840–843.

Auerbach, G., Huber, R. Grättinger, M., Zaiss, K., Schurig, H., Jaenicke, R. & Jacob, U. (1997). Closed structure of phosphoglycerate kinase from Thermotoga maritima reveals the catalytic mechanism and determinants of thermal stability. *Structure* **5**, 1475–1483.

Austin, S., Kundrot, C. & Dixon, R. (1991). Influence of a mutation in the putative nucleotide binding site of the nitrogen regulatory protein NTRC on its positive control function. *Nucleic Acids Res*. **19**, 2281–2287.

Auzat, I., Le Bras, G. & Garel, J.-R. (1994). The cooperativity and allosteric inhibition of *Escherichia coli* phosphofructokinase depend on the interaction between threonine-125 and ATP. *Proc. Natl. Acad. Sci. USA* **91**, 5242–5246.

Baek, Y. H. & Nowak, T. (1982). Kinetic evidence for a dual cation role for muscle pyruvate kinase. *Arch. Biochem. Biophys*. **217**, 491–497.

Bairoch, A., Buchner, P. & Hofmann, K. (1997). The PROSITE database, its status in 1997. *Nucl. Acids Res*. **25**, 217–221.

Baldwin, T. O. (1996). Firefly luciferase: the structure is known, but the mystery remains. *Structure* **15**, 223–228.

Baldwin, T. O. Christopher, J. A., Raushel, F. M., Sinclair, J. F., Ziegler, M. M., Fisher, A. J. & Rayment, I. (1995). Structure of bacterial luciferase. *Curr. Op. in Struc. Biology* **5**, 798–809.

Barnes, L. D., Garrison, P. N., Siprashvili, Z., Guranowski, A., Robinson, A. K., Ingram, S. W., Croce, C. M., Ohta, M. & Huebner, K. (1996). Fhit, a putative tumor suppressor in humans, is a dinucleoside 5′,5′″-P1,P3-triphosphate hydrolase. *Biochemistry* **35**, 11529–11535.

Bass, M. B., Fromm, H. J. & Rudolph, F. B. (1984). The mechanism of the adenylosuccinate synthetase reaction as studied by positional isotope exchange. *J. Biol. Chem*. **259**, 12330–12333.

Bates, A. D., O'Dea, M. H., & Gellert, M. (1996). Energy coupling in *E. coli* DNA gyrase: the reationship between nucleotide binding, strand passage, and DNA supercoiling. *Biochemistry* **35**, 1408–1416.

Belrhali, H., Yaremchuk, A. D., Tukalo, M. A., Larsen, K., Berthet-Colominas, C., Leberman, R., Beijer, B., Sproat, B., Als-Nielsen, J., Grubel, G., Legrand, J.-F., Lehmann, M. & Cusack, S. (1994). Crystal structures of 2.5 Å resolution of seryl-tRNA synthetase complexed with two different analogues of seryl-adenylate. *Science* **263**, 1432–1436.

Belrhali, H., Yaremchuk, A., Tukalo, M., Berthet-Colominas, C., Rasmussen, B., Bösecke, P., Diat, O. & Cusack, S. (1995). The structural basis for seryl-

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

adenylate and $A_{p4}A$ synthesis by seryl-tRNA synthetase. *Structure* **3**, 341–352.

Benedetti, P., Silvestri, A., Fiorani, P. & Wang, J. C. (1997). Study of yeast DNA topoisomerase and its truncation derivatives by transmission electron microscopy. *J. Biol. Chem.* **272**, 12132–12137.

Bennett Jr, W. S. & Steitz, T. A. (1980). Structure of a complex between yeast hexokinase A and glucose. I. Structure determination and refinement at 3.5 Ångstroms resolution. *J. Mol. Biol.* **140**, 183–209.

Berchtold, H., Reshetnikova, L., Reiser, C. P. A., Schirmer, N. K., Sprinzl, M. & Hilgenfeld, R. (1993). Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature* **365**, 126–132.

Bernstein, B. E., Michels, P. A. M. & Hol, W. G. J. (1997). Synergistic effects of the substrate-induced conformational changes in phosphoglycerate kinase activation. *Nature* **385**, 275–278.

Berthet-Colominas, C., Seignovert, L., Hartlein, M., Grotli, M., Cusack, S. & Leberman, R. (1998). The crystal structure of asparaginyl-tRNA syntetase from Thermus thermophilus and its complexes with ATP and asparaginyl-adenylate – the mechanism of discrimination between asparagine and aspartic acid. *EMBO J.* **17**, 2947–2960.

Berger, J. M. (1998). Type II DNA topoisomerases. *Curr. Op. Struc. Biol.* **8**, 26–32.

Berger, J. M., Gamblin, S. J., Harrison, S. C. & Wang, J. C. (1996). Structure and mechanism of DNA topoisomerase II. *Nature* **379**, 225–232.

Bessman, M. J., Frick, D. N. & O'Handley, S. F. (1996). The MutT proteins or 'nudix' hydrolases, a family of versatile, widely distributed, 'housecleaning' enzymes. *J. Biol. Chem.* **271**, 25059–25062.

Bird, L. E., Hosahalli, S., Subramanya, S. & Wigley, D. B. (1998). Helicases: a unifying structural theme? *Curr. Op. Struct. Biol.* **8**, 14–18 .

Blake, C. C. & Evans, P. R. (1974). Structure of horse muscle phosphoglycerate kinase, some results on the chain conformation, substrate binding and the evolution of the molecule from a 3 Ångstroms fourier map. *J. Mol. Biol.* **84**, 585–601.

Block, S. M. (1997). Real engines of creation. *Nature* **386**, 217–219

Boisvert, D. C., Wang, J., Otwinowski, Z. Horwich, A. L. & Sigler, P. B. (1996). The 2.4 Å crystal structure of the bacterial chaperonin GroEL complexed with ATPγS. *Nature Struct. Biol.* **3**, 170–177.

Bork, P. & Koonin, E. V. (1994). A P-loop-like motif in a widespread ATP pyrophosphatase domain: Implications for the evolution of sequence motifs and enzyme activity. *Proteins* 20, 347–355.

Bork, P., Sander, C. & Valencia, A. (1992). An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc. Natl. Acad. Sci.* **89**, 7290–7294.

Bork, P., Holm, L., Koonin, E. & Sander, C. (1995). The cytidylyltransferase superfamily: identification of the nucleotide binding site and fold prediction. *Proteins* **22**, 259–266.

Bossemeyer, D., Engh, R. A., Kinzel, V., Ponstingl, H. & Huber, R. (1993). 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.* **12**, 849–859.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* **371**, 578–586.

Branden, C. & Tooze, J. (1991). In *Introduction to Protein Structure*. Garland Publishing, Inc. N. Y. and London 1991 p.134

Brautigam, C. A. & Steitz, T. A. (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Op. Struc. Biol.* **8**, 54–63.

Brenner, C., Garrison, P., Gilmour, J., Peisach, D., Ringe, D., Petsko, G. A. & Lowenstein, J. M. (1997). Crystal structures of HINT demonstrate that histidine triad proteins are GalT-related nucleotide-binding proteins. *Nat. Struc. Biol.* **4**, 231–238.

Brick, P., Bhat, T. N. & Blow, D. M. (1989). Structure of tyrosyl transfer RNA synthetase refined at 2–3 Å resolution interaction of the enzyme with the tyrosyl adenylate intermediate. *J. Mol. Biol.* **208**, 83–98.

Brown, D. G., Visse, R., Sandhu, G., Davies, A., Rizkallah, P. J., Mulitz, C., Summers, W. C. & Sanderon, M. R. (1995). Crystal structure of the thymidine kinase from herpes simplex virus type-I in complex with deoxythymidine and ganciclovir. *Nature Struct. Biol.* **2**, 876–881.

Brunie, S., Zelwer, C. & Risler, J. L. (1990). Crystallographic study at 2.5 Å resolution of the interaction of methionyl transfer RNA synthetase from *Escherichia coli* with ATP. *J. Mol. Biol.* **216**, 411–424.

Bryant, T. N., Watson, H. C. & Wendell, P. L. (1974). Structure of yeast phosphoglycerate kinase. *Nature* **247**, 14–17.

Bryant, S. H., Madej, T., Janin, J., Liu, Y., Ruoho, A. E., Zhang, G. & Hurley, J. H. (1997). reply to previous letter. *Nature* **388**, 34–34.

Bukau, B. & Horwich, A. L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell* 92, 351–366.

Cabral, J. H., Jackson, A. P., Smith, C. V., Shikotra, N., Maxwell, A. & Liddington, R. C. (1997). Crystal structure of the breakage-reunion domain of DNA gyrase. *Nature* **388**, 903–906.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Canangarajah, B. J., Khokhlatchev, A., Cobb, M. H. & Goldsmith, E. J. (1997). Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. *Cell* **90**, 859–869.

Cavarelli, J., Eriani, G., Rees, B., Ruff, M., Boeglin, M., Mitschler, A., Martin, F., Gangloff, J., Thierry, J.-C. & Moras, D. (1994). The active site of yeast aspartyl-tRNA synthetase: structural and functional aspects of the aminoacylation reaction. *EMBO J.* **13**, 327–337.

Cedergren-Zeppezauer, E. S., Larsson, G., Nyman, P. O., Dauter, Z. & Wilson, K. S. (1992). Crystal structure of a dUTPase. *Nature* **355**, 740–743.

Champoux, J. J. (1990). In *DNA Topology and its Biological Effects*. Cold Spring Harbour NY: Cold Spring Harbour Lab., pp. 217–242.

Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A., Clarke, A. R. & Saibil, H. R. (1994). Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature* **37**, 261–264.

Chin, D. T., Goff, S. A., Webster, T., Smith, T. & Goldberg, A. L. (1988). Sequence of the *lon* gene in *Escherichia coli*. A heat-shock gene which encodes the ATP-dependent protease La. *J. Biol. Chem.* **263**, 11718–11728.

Cleves, A. E. & Kelly, R. B. (1996). Protein translocation: rehearsing the ABCs. *Current Biology* **6**, 276–278.

Cohn, M. (1959). Mechanisms of enzymatic cleavage of some organic phosphates. *J. Cell. Comp. Physiol.* **54**, Suppl. 1, 17–31.

Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G. & Sprang, S. R. (1994). Structures of active conformations of Gia1 and the mechanism of GTP hydrolysis. *Science* **265**, 1405–1412.

Conti, E., Franks, N. P. & Brick, P. (1996). Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* **4**, 287–298.

Cusack, S. (1995). Eleven down and nine to go. *Nat. Struct. Biol.* **2**, 824–831.

Cusack, S. (1997). Aminoacyl-tRNA synthetases. *Curr. Opinion Struct. Biol.* **7**, 881–889.

Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. & Leberman, E. (1990). A second class of synthetase structure revealed by X-ray analysis of *E. coli* seryl-tRNA synthetase at 2.5 Å resolution. *Nature* **347**, 249–255.

Cusack, S., Yaremchuk, A., Krikliviy, I. & Tukalo, M. (1998). tRNA(Pro) anticodon recognition by Thermus thermophilus prolyl-tRNA synthetase. *Structure* **6**, 101–108.

Davies, G. J., Gamblin, S. J., Littlechild, J. A., Dauter, Z. & Watson, H. C. (1994). Structure of the ADP complex of the 3–phosphoglycerate kinase from Bacillus stearothermophilus at 1.65 Ångstroms. *Acta Cryst.* **D50**, 202–209.

Davies II, J. F., Almassy, R. J., Hostomska, Z., Ferre, R. A. & Hostomsky, Z. (1994). 2.3 Ångstroms crystal structure of the catalytic domain of DNA polymerase β. *Cell* **76**, 1123–1133.

DeBondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O. & Kim, S.-H. (1993). Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595–602.

Delarue, M., Poch, O., Tordo, N., Moras, D. & Argos, P. (1990). An attempt to unify the structure of polymerases. *Protein Eng.* **3**, 461–467.

Delarue, M., Poterszman, A., Nikonov, S., Garber, M., Moras, D. & Thierry, J. C. (1994). Crystal structure of a prokaryotic aspartyl tRNA-synthetase. *EMBO J.* **13**, 3219–3229.

Ditzel, L., Lowe, J., Stock, D., Stetter, K. O., Huber, H., Huber, R. & Steinbacher, S. (1998). Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. *Cell* **93**, 125–138.

Dreusicke, D., Karplus, P. A. & Schulz, G. E. (1988). Refined structure of porcine cytosolic adenylate kinase at 2.1 Å resolution. *J. Mol. Biol.* **199**, 359–371.

Doublié, S., Bricogne, G., Gilmore, C. & Carter, C. W. Jr. (1995). Tryptophanyl-tRNA synthetase crystal structure reveals an unexpected homology to tyrosyl-tRNA synthetase. *Structure* **3**, 17–31.

Doublié, S., Tabor, S., Long, A. M., Richardson, C. C. & Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**, 251–258.

Druck, T., Hadaczek, P., Fu, T. B., Ohta, M., Siprashvili, Z., Baffa, R., Negrini, M., Kastury, K., Veronese, M. L., Rosen, D., Rothstein, J., McCue, P., Cotticelli, M. G., Inoue, H., Croce, C. M. & Huebner, K. (1997). Structure and expression of the human FHIT gene in normal and tumor cells. *Cancer Res.* **57**, 504–512.

Dumas, C., Lascu, I., Morera, S., Glaser, P. Fourme, R., Wallet, V., Lacombe, M.-L., Veron, M. & Janin, J. (1992). X-ray structure of nucleoside diphosphate kinase. *EMBO J.* **11**, 3203–3208.

Egelman, E. H. (1996). Homomorphous hexameric helicases. Tales from the ring cycle. *Structure* **4**, 759–762.

Eom, S. H., Wang, J. & Steitz, T. A. (1996). Structure of Taq polymerase with DNA at the polymerase active site. *Nature* **382**, 278–281.

Erianie, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **347**, 203–206.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Esser, L., Wang, C.-R., Hosaka, M., Smagula, S., Südhof, T. C. & Deisenhofer, J. (1998). Synapsin I is structurally similar to ATP-utilizing enzymes. *EMBO J.* **17**, 77–84.

Evans, P. R., Farrants, G. W. & Hudson, P. J. (1981). Phosphofructokinase. Structure and control. *Phil. Trans. R. Soc. London* **293**, 53–62.

Fan, C., Moews, P. C., Walsh, C. T. & Knox, J. R. (1994). Vancomycin resistance: Structure of D-Alanine:D-Alanine ligase at 2.3 Ångstroms. *Science* **266**, 439–443.

Fan, C., Moews, P. C., Shi, Y., Walsh, C. T. & Knox, J. R. (1995). 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* **92**, 1172–1176.

Finbow, M. E. & Harrison, M. A. (1997). The vacuolar H^+-ATPase: a universal proton pump of eukaryotes. *Biochem. J.* **324**, 697–712.

Fisher, A. J., Smith, C. A., Thoden, J. B., Smith, R., Sutoh, K., Holden, H. M. & Rayment, I. (1995). X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP. BeFc and MgADP.AlF_4^-. *Biochemistry* **34**, 8960–8972.

Flaherty, K. M., Deluca-Flaherty, C. R. & McKay, D. B. (1990). Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* **346**, 623–628.

Freymann, D. M., Keenan, R. J., Stroud, R. M. & Walter, P. (1997). Structure of the conserved GTPase domain of the signal recognition particle. *Nature* **385**, 361–364.

Frick, D. N., Weber, D. J., Abeygunawardana, A., Gittis, A. G., Bessman, M. J. & Mildvan, A. S. (1995). NMR Studies of the conformations and location of nucleotides bound to the *Escherichia coli* MutT enzyme. *Biochemistry* **34**, 5577–5586.

Fritz-Wolf, K., Schnyder, T., Wallimann, T. & Kabsch, W. (1996). Structure of mitochondrial creatine kinase. *Nature* **381**, 341–345.

Galperin, M. Y. & Koonin, E. V. (1997). A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. *Protein Science* **6**, 2639–2643.

Georgiadis, M. M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. J. & Rees, D. C. (1992). Crystallographic structure of the nitrogenase iron protein from Azotobacter vinelandii. *Science* **257**, 1653–1659.

Georgiadis, M. M., Jessen, S. M., Ogata, C. M., Telesnitsky, A., Goff, S. P. & Hendrickson, W. A. (1995). Mechanistic implications from the structure of a catalytic fragment of Moloney murine leukaemia virus reverse transcriptase. *Structure* **3**, 879–892.

Germann, U. A., Pastan, I. & Gottesman, M. M. (1996). P-glycoproteins: mediators of multidrug resistance. *Sem. Cell Biol.* **4**, 63–76.

Gerstein, M. (1997). A structural census of genomes: comparing bacterial, eucaryotic, and archaeal genomes in terms of protein structure. *J. Mol. Biol.* **274**, 562–576.

Goldberg, J., Nairn, A. C. & Kuriyan, J. (1996). Structural basis for the autoinhibition of calcium/calmodulin-dependent protein Kinase I. *Cell* **84**, 875–887.

Goldgur, Y., Masyak, L., Reshetnikova, L., Ankilova, V., Lavrik, O., Khodyreva, S. & Safro, M. (1997). The crystal structure of phenylalanyl-tRNA synthetase from T. thermophilus complexed with cognate $tRNA^{phe}$. *Structure* **5**, 59–68.

Gorbalenya, A. E. & Koonin, E. V. (1990). Superfamily of UvrA-related NTP-binding proteins. Implications for rational classification of recombination/repair systems. *J. Mol. Biol.* **213**, 583–591.

Gottesman, J., Squires, C., Pichersky, E., Carrington, M., Hobbs, M., Mattick, J. S., Dalrymple, B., Kuramitsu, H., Shirozy, T., Foster, T. *et al.* (1990). Conservation of the Regulatory Subunit for the Clp ATP-dependent Protease in Prokaryotes and Eukaryotes. *Proc. Natl. Acad. Sci. USA* **87**, 3513–3517.

Greenspan, D. L., Connolly, D. C., Wu, R., Lei, R. Y., Vogelstein, J. T. C., Kim, Y. T., Mok, J. E., Munoz, N., Bosch, F. X., Shah, K. & Cho, K. R. (1997). Loss of FHIT expression in cervical carcinoma cell lines and primary tumors. *Cancer Research* **57**, 4692–4698.

Gupta, R. K., Oesterling, R. M. & Mildvan, A. S. (1976). Dual divalent cation requirement for activation of pyruvate kinase: essential roles of both enzyme- and nucleotide-bound metal ion. *Biochemistry* **15**, 2881–2887.

Håkansson, K., Doherty, A. J., Shuman, S. & Wigley, D. B. (1997). X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. *Cell* **89**, 545–553.

Hanks, S. K., Quinn, A. M. & Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytin domains. *Science* **241**, 42–52.

Hanks, S. K. & Hunter, T. (1995). The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–596.

Hara, T., Kato, H., Katsube, Y. & Oda, J. I. (1996). 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-Ångstrom resolution. *Biochemistry* **35**, 11967–11974.

Harlos, K., Vas, M. & Blake, C. F. (1992). Crystal structure of the binary complex of pig muscle phosphoglycerate kinase and its substrate 3-phospho-D-glycerate. *Proteins* **12**, 133–144.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Harrison, C. J., Hayerhartl, M., Di Liberto, M., Hartl, F. U. & Kuriyan, J. (1997). Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science* **276**, 431–435.

Hartl, U. (1996). Molecular chaperones in cellular protein folding. *Nature* **381**, 571–580.

Hasemann, C. A., Istvan, E. S., Uyeda, K. & Deisenhofer, J. (1996). The crystal structure of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase reveals distinct domain homologies. *Structure* **4**, 1017–1029.

Herzberg, O., Chen, C. C., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J. & Dunaway-Mariano, D. (1996). Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites. *Proc. Natl. Acad. Sci. USA* **93**, 2652–2657.

Hinds, R. M., Xu, J., Walters, D. E. & Kemp, R. G. (1998). The active site of pyrophosphate-dependent phosphofructo-1-kinase based on site-directed mutagenesis and molecular modeling. *Arch. Biochem. Biophys.* **34**, 47–52.

Hirokawa, N. (1998). Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science* **279**, 519–526.

Holdgate, G. A., Tunnicliffe, A., Ward, W. H. J., Weston, S. A., Rosenbrock, G., Barth, P. T., Taylor, I. W. F., Pauptit, R. A. & Timms, D. (1997). The entropic penalty of ordered water accounts for weaker binding of the antibiotic Novobiocin to a resistant mutant of DNA gyrase: a thermodynamic and crystallographic study. *Biochemistry* **36**, 9663–9673.

Holm, L. (1998). Unification of protein families. *Curr. Op. Struc. Biol.* **8**, 372–379.

Holm, L. & Sander, C. (1995). DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily. *TIBS* **20**, 345–347.

Honzatko, R. B., Crawford, J. L., Monaco, H., Ladner, J. E., Edwards, B. F. P., Evans, D. R., Warren, S. G., Wiley, D. C., Ladner, R. C. & Lipscomb, W. N. (1982). Crystal and molecular structures of native and CTP-liganded aspartate carbamoyltransferase from *Escherichia coli*. *J. Mol. Biol.* **160**, 219–263.

Howard, J. (1997). Molecular motors: structural adaptations to cellular functions. *Nature* 38, 561–567.

Hu, S.-H., Parker, M. W., Lei, J. Y., Wilce, M. C. J., Benian, G. M. & Kemp, B. E. (1991). Insights into autoregulation from the crystal structure of twitchin kinase. *Nature* **369**, 581–584.

Huang, W., Lindqvist, Y., Schneider, G., Gibson, K., Flint, D. & Lorimer, G. (1994). Crystal structure of an ATP dependent carboxylase, dethiobiotin synthetase, at 1.65 Ångstroms Resolution. *Structure* **2**, 407–414.

Huang, W., Jia, J., Gibson, K. J., Taylor, W. S., Rendina, A. R., Schneider, G. & Lindqvist, Y. (1995). Mechanism of an ATP-dependent carboxylase, dethiobiotin synthetase, based on crystallographic studies of complexes with substrates and a reaction intermediate. *Biochemistry* **34**, 10985–10995.

Hubbard, S. R. (1997). Crystal structure of the activated insuline receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J.* **16**, 5572–5581.

Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. A. (1994). Crystal structure of the tyrosine kinase domain of the human insuline receptor. *Nature* **372**, 746–754.

Hung, L.-W., Wang, I. X., Nikaido, K., Liu, P.-Q., Ferro-Luzzi Ames, G. & Kim, S.-H. (1998). Crystal structure of ATP-binding subunit of an ABC transporter. *Nature* 396, 703–707.

Hunter, T. (1987). A thousand and one protein kinases. *Cell* **50**, 823–829.

Hunter, T. (1995). Protein kinases and phosphatases: the Yin and Yang of protein phosphorylation and signaling. *Cell* **80**, 225–236.

Hunter, T. & Plowman, G. D. (1997). The protein kinases of budding yeast: six score and more. *TIBS* **22**, 18–22.

Hurley, J. H., Faber, H. R., Worthylake, D., Meadow, N. D., Roseman, S., Pettigrew, D. W. & Remington, S. J. (1993). Structure of the regulatory complex of *Escherichia coli* IIIGlc with glycerol kinase. *Science* **259**, 673–677.

Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark Jr., A. D., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H. & Arnold, E. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Ångstroms resolution shows bent DNA. *Proc. Natl. Acad. Sci. USA* **90**, 6230–6324.

Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massagué, J. & Pavletich, N. P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **376**, 313–320.

Johnson, L. N., Noble, M. E. M. & Owen, D. J. (1996). Active and inactive protein kinases – structural basis for regulation. *Cell* **85**, 149–158.

Jurnak, F. (1985). Structure of the GDP domain of EF-Tu and location of amino acids homologous to ras onco × gene proteins. *Science* **230**, 32–36.

Kabsch, W. & Fritz-Wolf, K. (1997). Mitochondrial creatine kinase – a square protein. *Curr. Op. Struc. Biol.* **7**, 811–818.

Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. & Holmes, K. C. (1990). Atomic structure of the Actin:DNase I complex. *Nature* 347, 37–44.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Kiefer, J. R., Mao, C., Hansen, J., Basehore, S. L., Hogrefe, H. H., Braman, J. C. & Beese, L. S. (1997). Crystal structure of a thermostable bacillus DNA polymerase I large fragment at 2.1 Å resolution. *Structure* **5**, 95–108.

Kim, Y., Eom, S. H., Wang, J., Lee, D. S., Suh, S. W. & Steitz, T. A. (1995). Crystal structure of *Thermus aquaticus* DNA polymerase. *Nature* **376**, 612–616.

Kletzin, A. (1992). Molecular characterizytion of a DNA ligase gene of the extremely thermophilic archaeon Desulfurolobus ambivalens shows close phylogenetic relationship tp eukaryotic ligases. *Nucl. Acids Res.* **20**, 5389–5396.

Knighton, D. R., Zheng, J., Ten Eyck, L. F., Ashford, V. A., Xuong, N.-H., Taylor, S. S. & Sowadski, J. M. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**, 407–414.

Knowles, J. R. (1980). Enzyme-catalyzed phosphoryl transfer reactions. *Ann. Rev. Biochem.* **49**, 877–919.

Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A. & Steitz, T. A. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**, 1783–1790.

Komoszynski, M. & Wojtczak, A. (1996). Apyrases (ATP diphosphohydrolases, EC 3.6.1.5): function and relationship to ATPases. *Biochim. Biophys. Acta* **1310**, 233–241.

Koonin, E. V. & Rudd, K. E. (1996). The two domains of superfamily I helicases may exist as separate proteins. *Protein Science* **5**, 178–180.

Korolev, S., Nayal, M., Barnes, W. M., Di Cera, E. & Waksman, G. (1995). Crystal structure of the large fragment of Thermus aquaticus DNA polymerase I at 2.5-Å resolution: structural basis for thermostability. *Proc. Natl. Acad. Sci. USA* **92**, 9264–9268.

Korolev, S., Hsieh, J., Gauss, G. H., Lohman, T. M. & Waksman, G. (1997). Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-strnded DNA and ADP. *Cell* **90**, 635–647.

Kull, F. J., Sablin, E. P., Lau, R., Fletterick, R. J. & Vale, R. D. (1996). Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. *Nature* **380**, 550–555.

Kunau, W. H., Beyer, A., Franken, T., Gotte, K., Marzioch, M., Saidowski, J., Skaletz-Rorowski, A. & Wiebel, F. (1993). Two complementary approaches to study peroxisome biogenesis in Saccharomyces cerevisiae: forward and reversed genetics. *Biochimie* **75**, 209–224.

Kutzenko, A. S., Lamzin, V. S. & Popov, V. O. (1998). Conserved supersecondary structural motif in NAD-dependent dehydrogenases. *FEBS Lett.* **423**, 105–109.

LaCour, T. F. M., Nyborg, J., Thirup, S. & Clark, B. F. C. (1985). Structural details of the binding of guanosine diphosphate to elongation factor Tu from *E. coli* as studied by X-ray crystallography. *EMBO J.* **4**, 2385–2388.

Ladror, U. S., Gollapudi, L., Tripathi, R. L., Latshaw, S. P. & Kemp, R. G. (1991). Cloning, sequencing, and expression of pyrophosphate-dependent phosphofructokinase from Propionibacterium freudenreichii. *J. Biol. Chem.* **266**, 16550–16555.

Larsen, T. M., Laughlin, T., Holden, H. M., Rayment, I. & Reed, G. H. (1994). Structure of rabbit muscle pyruvate kinase complexed with Mn^{2+}, K^+ and pyruvate. *Biochemistry* **33**, 6301–6309.

Larsen, T. M. Benning, M. M., Wesenberg, G. E., Rayment, I. & Reed, G. H. (1997). Ligand-induced domain movement in pyruvate kinase: structure of the enzyme from rabbit muscle with Mg^{++}, N^+ and L—phospholactate at 2.7 Å resolution. *Arch. Biochem. Biophys.* **345**, 199–206.

Larsson, G., Svensson, L. A. & Nyman, P. O. (1996). Crystal structure of the *Escherichia coli* dUTPase on complex with a substrate analogue (dUDP). *Nature Struc. Biol.* **3**, 532–538.

Larsson-Raznikiewicz, M. M. (1973). Graphical analyses on binding of ligands to a two-sited system. Theoretical treatments exemplified on yeast phosphoglycerate kinase. *Arch. Biochem. Biophys.* **158**, 754–762.

Lavie, A., Vetter, I. R., Konrad, M., Goody, R. S., Reinstein, J. & Schlichting, I. (1997). Structure of thymidylate kinase reveals the cause behind the limiting step in AZT activation. *Nature Structural Biology* **4**, 601–604.

Leatherbarrow, R. J., Fersht, A. R. & Winter, G. (1985). Transition-state stabilization in the mechanism of tyrosyl-tRNA synthetase revealed by protein engineering. *Proc. Natl. Acad. Sci. USA* **82**, 7840–7844

Lenzen, C. U., Steinmann, D., Whiteheart, S. W. & Weis, W. I. (1998). Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell* **94**, 525–536.

Lew, J., Taylor, S. S. & Adams, J. A. (1997). Identification of a partially rate-determining step in the catalytic mechanism of cAMP-dependent protein kinase – a transient kinetic study using stopped-flow fluorescence spectroscopy. *Biochemistry* **36**, 6717–6724.

Liaw, S. H., Pan, C. & Eisenberg, D. (1993). Feedback inhibition of fully unadenylated glutamine synthetase from Salmonella typhimurium by glycine, alanine and serine. *Proc. Natl. Acad. Sci. USA* **90**, 4996–5000.

Liaw, S.-H., Jun, G. & Eisenberg, D. (1994). Interaction of nucleotides with fully unadenylated glutamine synthetase from Salmonella typhimurium. *Biochemistry* **33**, 11184–11188.

Liaw, S. H., Kuo, I. C. & Eisenberg, D. (1995). Discovery of the ammonium substrate site on glutamine synthetase, a third cation binding site. *Protein Science* **4**, 2358–2365.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Lima, C. D., Wang, J. C. & Mondragon, A. (1994). Three-dimensional structure of the 67k N-terminal fragment of *E. coli* DNA topoisomerase I. *Nature* **367**, 138–146.

Lin, J., Abeygunawardana, C., Frick, D. N., Bessman, M. J. & Mildvan, A. S. (1997). Solution structure of the quaternary MutT-M^{2+}-AMPPCP-M^{2+}Complex and Mechanism of its Pyrophosphohydrolase Action. *Biochemistry* **36**, 1199–1211.

Lindner, P., Lasko, P. F., Leroy, P., Nielsen, P. J., Nishi, K., Schnier, J. & Slonimski, P. P. (1989). Birth of the D-E-A-D box. *Nature* **337**, 121–122.

Linton, K. J. & Higgins, C. F. (1998). The *Escherichia coli* ATP-binding cassette (ABC) proteins. *Molec. Microbiol.* **28**, 5–13.

Llorca, O., Marco, S., Carrascosa, J. L. & Valpuesta, J. M. (1997). Conformational changes in the GroEL oligomer during the functional cycle. *J. Struc. Biol.* **118**, 31–42.

Loewe, J. & Amos, L. A. (1998). Crystal structure of the bacterial cell-division protein FtsZ. *Nature* **391**, 203–206.

Logan, D. T., Mazauric, M.-H., Kern, D. & Moras, D. (1995). Crystal structure of glycyl-tRNA synthetase from T. thermophilus. *EMBO J.* **14**, 4156–4167.

Lue, N., Sharma, A., Mondragon, A & Wang, J. C. (1995). A 25kDa yeast DNA topoisomerase I fragment: crystallographic structure and mechanistic implications. *Structure* **3**, 1315–1322.

Markham, G. D., Hafner, E. W., Tabor, C. W. & Tabor, H. (1980). S-Adenosylmethionine synthetase from Escherichia coli. *J. Biol. Chem.* **255**, 9082–9092.

Matsuda, K., Mizuguchi, K., Nishioka, T., Kato, H., Go, N. & Oda, J. (1996). Crystal structure of gluthathione synthetase at optimal pH: domain architecture and structural similarity with other proteins. *Protein Engineering* **9**, 1083–1092.

Matte, A., Goldie, H., Sweet, R. M. & Delbaere, L. T. (1996). Crystal structure of *Escherichia coli* phosphoenolpyruvate carboxykinase: a new structural family with the P-loop nucleoside triphosphate hydrolase fold. *J. Mol. Biol.* **256**, 126–143.

Matte, A., Tari, L. W., Goldie, H. & Delbaere, L. T. (1997). Structure and mechanism of phosphoenolpyruvate carboxykinase, *J. Biol. Chem.* **272**, 8105–8108.

Mattevi, A., Valentini, G., Rizzi, M., Speranza, M. L., Bolognesi, M. & Coda, A. (1995). Crystal structure of *Escherichia coli* pyruvate kinase type I: molecular basis of the allosteric transition. *Structure* **3**, 729–741.

Mayans, O., van der Ven, P. F. M., Wilm, M., Mues, A., Young, P., Fürst, D. O., Wilmanns, M. & Gautel, M. (1998). Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature* **395**, 863–869.

McPhillips, T. M., Hsu, B. T., Sherman, M. A., Mas, M. T. & Rees, D. (1996). Structure of the R65Q mutant of yeast 3-phosphoglycerate kinase complexed with Mg-AMP-PNP and 3-phospho-D-glycerate. *Biochemistry* **35**, 4118–4127.

Milburn, M. V., Tong, L., DeVos, A. M., Brünger, A., Yamaizumi, Z., Nishimura, S. & Kim, S.-H. (1990). Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic *ras* proteins. *Science* **247**, 939–945.

Mildvan, A. S. (1970). In *The Enzymes* (ed. P. D. Boyer). 3rd Edn, pp. 445–536. Academic Press, New York.

Milner-White, E. J., Coggins, J. R. & Anton, I. A. (1991). Evidence for an ancestral core structure in nucleotide-binding proteins with the type A motif. *J. Mol. Biol.* **221**, 751–754.

Mohammadi, M., Schlessinger, J. & Hubbard, S. R. (1996). Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism. *Cell* 86, 577–587.

Mohammadi, M., McMahon, G., Sun, L., Tang, C., Hirth, P., Yeh, B. K., Hubbard, S. R. & Schlessinger, J. (1997). Structures Of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science* **276**, 955 -960

Møller, J. V., Juul, B. & le Maire, M. (1996). Structural organization, ion transport, and energy transduction of P-type ATPases. *Biochim. Biophys. Acta*. **1286**, 1–51.

Montoya, G., Svensson, C., Luirink, J. & Sinning, I. (1997). Crystal structure of the NG domain from the signal-recognition particle receptor FtsY. *Nature* **385**, 365–368.

Moréra, S., Lascu, I., Dumas, C., LeBras, G., Briozzo, P, Véron, M. & Janin, J. (1994). Adenosine 5′-diphosphate binding and the active site of nucleoside diphosphate kinase. *Biochemistry* **33**, 459–467.

Moréra, S., Chiadmi, M., LeBras, G., Lascu, I. & Janin, J. (1995). Mechanism of phosphate transfer by nucloeside diphosphate kinase: X-ray structures of the phosphohistidine intermediate of the enzymes from Drosophila and Dictyostelium. *Biochemistry* **34**, 11062–11070.

Mosyak, L., Reshetnikova, L., Goldgur, Y., Delarue, M. & Safro, M. G. (1995). Structure of phenylalanyl-tRNA synthetase from T. thermophilus. *Nat. Struc. Biol.* **2**, 537–547.

Müller, C. W. & Schulz, G. E. (1992). Structure of the complex between adenylate kinase from *Escherichia coli* and the inhibitor Ap_5A refined at 1.9 Å resolution. A model for a catalytic transition state. *J. Mol. Biol.* **224**, 159–177.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Müller-Dieckmann, H. J. & Schulz, G. E. (1994). The structure of uridylate kinase with its substrates: Geometry of the phosphoryl transfer transition state. *J. Mol. Biol.* **236**, 361–367.

Murzin, A. G. (1996). Structural classification of proteins: new superfamilies. *Curr. Op. Struct. Biol.* **6**, 386-394.

Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). SCOP: A structural classification of protein database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536–540.

Nakatsu, T., Kato, H. & Oda, J. (1998). Crystal structure of asparagine synthetase reveals a close evolutionary relationship to Class II tRNA synthetase. *Nature Struct. Biol.* **5**, 15–19.

Negrini, M., Monaco, C., Vorechovsky, I., Ohta, M., Druck, T., Baffa, R., Huebner, K. & Croce, C. M. (1996). The FHIT gene at 3p14.2 is abnormal in breast carcinomas. *Cancer Res.* **56**, 3173–3179.

Niefind, K., Guerra, B., Pinna, L. A., Issinger, O.-G. & Schomburg, D. (1998). Crystal structure of the catalytic subunit of protein kinase CK2 from Zea mays at 2.1 Å resolution. *EMBO J.* **17**, 2451–2462.

Noel, J. P., Hamm, H. E. & Sigler, P. B. (1993). The 2.2 Å crystal structure of transducin-α complexed with GTPγS. *Nature* **366**, 654–663.

Nogales, E., Wolf, S. G. & Downing, K. H. (1998a). Structure of the ab tubulin dimer by electron crystallography. *Nature* **391**, 199–203.

Nogales, E., Downing, K. H., Amos, L. A. & Lowe, J. (1998b). Tubulin and FtsZ form a distinct family of GTPases. *Nat. Struc. Biol.* **5**, 451–458.

Nureki, O., Vassylyev, D. G., Katayanagi, K., Shimizu, T., Sekine, S., Kigawa, T., Muyazawa, T., Yokoyama, S. & Morikawa, K. (1995). Architectures of class-defining and specific domains of glutamyl-tRNA synthetase. *Science* **267**, 1958–1965.

Nureki, O., Vassylyev, D. G., Tateno, M., Shimada, A., Nakama, T., Fukai, S., Konno, M., Hendrickson, T. L., Schimmel, P. & Yokoyama, S. (1998). Enzyme structure with two catalytic sites for double-sieve selection of substrate. *Science* **280**, 578–582.

Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G. & Steitz T. A. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* **313**, 762–766.

Onesti, S., Miller, A. D. & Brick, P. (1995). The crystal structure of the lysyl-tRNA synthetase (LysU). from *Escherichia coli*. *Structure* **3**, 163–176.

Owen, D. J., Noble, M. E.M, Garman, E. F., Papageorgiou, A. C. & Johnson, L. N. (1995). Two structures of the catalytic domain of phosphorylase kinase: an active protein kinase complexed with substrate analogue and product. *Structure* **3**, 467–482.

Pace, H. C., Garrison, P. N., Robinson, A. K., Barnes, L. D., Draganescu, A., Rösler, A., Blackburn, G. M., Siprashvili, Z., Croce, C. M., Huebner, K. & Brenner, C. (1998). Genetic, biochemical and crystallographic characterization of Fhit-substrate complexes as the active signaling form of Fhit. *Proc. Natl. Acad. Sci.* **95**, 5484–5489.

Pai, E. F., Krengel, U., Petsko, G. A., Goody, R. S., Kabsch, W. & Wittinghofer, A. (1990). Refined crystal structure of the triphosphate conformation of H-*ras* p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* **9**, 2351–2359.

Parsell, D. A., Sanchez, Y., Stitzel, J. D. & Lindquist, S. (1991). Hsp104 is a highly conserved protein with two essential nucleotide binding sites. *Nature* **353**, 270–273.

Patel, S. & Latterich, M. (1998). The AAA team: related ATPases with diverse functions. *Trends in Cell Biology* **8**, 65–71.

Pedersen, L. C., Benning, M. M. & Holden, H. M. (1995). Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase. *Biochemistry* **34**, 13305–13311.

Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H. & Kraut, J. (1994). Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. *Science* **266**, 2025–2026.

Pelletier, H., Sawaya, M. R., Wolfle, W., Wilson, S. & Kraut, J. (1996). Crystal structures of human DNA polymerase β complexed with DNA: implications for catalytic mechanism, processivity and fidelity. *Biochemistry* **35**, 12742–12761.

Perona, J., Rould, M. A. & Steitz, T. A. (1993). Structural basis for transfer RNA aminoacylation by *Escherichia coli* glutaminyl-tRNA synthetase. *Biochemistry* **32**, 8758–8771.

Poland, B. W., Silva, M. M., Serra, M. A., Cho, Y., Kim, K. H., Harris, E. M. S. & Honzatko, R. B. (1993). Crystal structure of adenylosuccinate synthetase from *Escherichia coli*. *J. Biol. Chem.* **268**, 25334–25342.

Poland, B. W., Hou, Z., Bruns, C., Fromm, H. J. & Honzatko, R. B. (1996). Refined crystal structure of guanine nucleotide complexes of adenylosuccinate synthase from Escherichia coli. *J. Biol. Chem.* **271**, 15407–15413.

Prodromou, C., Roe, S. M., Piper, P. W. & Pearl, L. H. (1997a). A molecular clamp in the crystal structure of the N-terminal domain of the yeast Hsp90 chaperone. *Nature Struc. Biol.* **4**, 477–482.

Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W. & Pearl, L. H. (1997b). Identification and characterization of the ATP/ADP binding site in the Hsp90 molecular chaperone. *Cell* **90**, 65–75.

Rao, B. D. N., Cohen, M. & Scopes, R. K. (1978). ^{31}P NMR study of bound reactants and products of yeast

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

3-phosphoglycerate kinase at equilibrium and the effect of sulfate ion. *J. Biol. Chem.* **253**, 8056–8060

Raushel, F. M., Thoden, J. B., Reinhart, G. D. & Holden, H. M. (1998). Carbamoyl phosphate synthetase – a crooked path from substrates to products. *Curr. Op. Chem. Biol.* **2**, 624–632.

Ray, B. D. & Rao, B. D. N. (1988). ^{31}P NMR studies of enzyme-bound substrate complexes of yeast 3-phosphoglycerate kinase. 1. Effects of sulfate and pH. Mg(II) affinity at the two ATP sites. *Biochemistry* **27**, 5574–5578

Rayment, I., Rypniewski, W. R., Schmidt-Bäse, K., Smith, R., Tomchick, D. R., Benning, M. M., Winkelmann, D. A., Wesenberg, G. & Holden, H. M. (1993). Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science* **261**, 50–65.

Redinbo, M. R., Stewart, L., Kuhn, P., Champoux, J. J. & Hol, W. G. J. (1998). Crystal structure of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science* **279**, 1504–1513.

Ren, J., Esnouf, R., Hoplins, A., Ross, C., Jones, Y., Stammers, D. & Stuart, D. (1995). The structure of HIV-1 reverse transcriptase complexed with 9-chloro-TIBO: lessons for inhibitor design. *Structure* **3**, 915–926.

Riordan, J. R., Rommens, J. M., Kerem, B.-S., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.-L., Drumm, M. L., Iannuzzi, M. C., Collins, F. S. & Tsui, L.-C. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* **245**, 1065–1073.

Rittinger, K., Walker, P. A., Eccleston, J. F., Nurmahomed, K., Owen, D., Laue, E., Gamblin, S. J. & Smerdon, S. J. (1997a). Crystal structure of a small G protein in complex with the GTPase-activating protein RHOGAP. *Nature* 388, 693–697.

Rittinger, K., Walker, P. A., Eccleston, J. F., Smerdon, S. J. & Gamblin, S. J. (1997b). Structure at 1.65 Å of RhoA and its GTPase-activating protein in complex with a transition-state analogue. *Nature* **389**, 758–762.

Rizzi, M., Nessi, C., Mattevi, A., Coda, A., Bolognesi, M. & Galizzi, A. (1996). Crystal structure of NH3-dependent NAD+ synthetase from Bacillus subtilis. *EMBO J.* **15**, 5125–5134.

Roca, J. R. & Wang, J. C. (1992). The capture of a DNA double helix by an ATP-dependent protein clamp: a key step in DNA transport by type II DNA topoisomerases. *Cell* **71**, 833–840.

Roca, J. & Wang, J. C. (1994). DNA transport by a type II DNA topoisomerase: evidence in favor of a two-gate mechanism. *Cell* **77**, 609–616.

Rodgers, D. W., Gamblin, S. J., Harris, B. A., Ray, S., Culp, J. S., Hellmig, B., Woolf, D. J., Debouck, C. & Harrison, S. C. (1995). The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* **92**, 1222–1226.

Roseman, A. M., Chen, S., White, H., Braig, K. & Saibil, H. R. (1996). The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. *Cell* **87**, 241–251.

Rothman, J. E. (1994). Mechanisms of intracellular protein transport. *Nature* **372**, 55–63

Rould, M., Perona, J. J. & Steitz, T. A. (1989). Structure of *Escherichia-coli* glutaminyl transfer RNA synthetase complexed with glutaminyl transfer RNA and ATP at 2.8 Å resolution. *Science* **246**, 1135–1142.

Rossmann, M. G., Liljas, A., Brändén, C.-I. & Banaszak, L. J. (1975). Evolutionary and structural relationships among dehydrogenases. In: *The Enzymes* (ed. P. D. Boyer), vol. **11**, 3rd edn, pp. 61–102. Academic Press, New York.

Ruff, M., Krishnaswamy, S., Boeglin, M., Potersz-man, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C. & Moras, D. (1991). Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA^{Asp}. *Science* **252**, 1682–1689.

Russo, A. A., Jeffrey, P. D., Patten, A. K., Massagué, J. & Pavletich, N. P. (1996). Crystal structure of the p27^{Kip1} cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 coplex. *Nature* **382**, 325–331.

Rypniewski, W. R. & Evans, P. R. (1989). Crystal structure of unliganded phosphofructokinase from *Escherichia coli*. *J. Mol. Biol.* **207**, 805–821.

Sack, S., Muller, J., Marx, A., Thormahlen, M., Mandelkow, E. M., Brady, S. T. & Mandelkow, E. (1997). X-ray structure of motor and neck domains from rat brain kinesin. *Biochemistry* **36**, 16155–16165.

Sakon, J., Liao, H. H., Kanikula, A. M., Benning, M. M., Rayment, I. & Holden, H. M. (1993). Molecular structure of kanamycin nucleotidyltransferase determined to 3.0 Å resolution. *Biochemistry* **32**, 11977–11984.

Saraste, M., Sibbald, P. R. & Wittinghofer, A. (1990). The P-loop – A common motif in ATP and GTP binding proteins. *Trends Biochem. Sci.* **15**, 430–434.

Savage, H., Montoya, G., Svensson, C., Schwenn, J. D. & Sinning, I. (1997). Crystal structure of phosphoadenyl sulphate (PAPS) reductase: a new family of adenine nucleotide α hydrolases. *Structure* **5**, 895–906.

Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H. & Kraut, J. (1994). Crystal structure of rat DNA polymerase β: evidence for a common polymerase mechanism. *Science* **264**, 1930–1935.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Scheffzek, K., Kliche, W., Wiesmüller, L. & Reinstein, J. (1996a). Crystal Structure of the complex of UMP/CMP kinase from Dictyostelium discoideum and the bisubstrate inhibitor P1-(5′-Adenosyl) P5-(5′-Uridyl) pentaphosphate (UP5A) and Mg^{2+} at 2.2 Å: implications for water-mediated specificity. *Biochemistry* **35**, 9716.

Scheffzek, K., Lautwein, A., Kabsch, W., Ahmadian, M. R. & Wittinghofer, A. (1996b). Crystal structure of the GTPase-activating domain of human p120GAP and implications for the interaction with Ras. *Nature* **384**, 591–596.

Scheffzek, K., Ahmadian, M. R., Kabsch, W., Wiesmüller, L., Lautwein, A., Schmitz, F. & Wittinghofer, A. (1997). Ras-RasGAP communication: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science* **277**, 333–338.

Schierbeck, B. & Larsson-Raznikiewicz, M. (1979). Product inhibition studies of yeast phosphoglycerate kinase evaluating properties of multiple substrate binding sites. *Biochim. Biophys. Acta*. **568**, 195–204.

Schindelin, N., Kisker, C., Schlessman, J. L., Howard, J. B. & Rees, D. C. (1997). Structure of ADP·AlF(4)(−)-stabilized nitrogenase complex and its implications for signal transduction. *Nature* **387**, 370–376.

Schlichting, I. & Reinstein, J. (1997). Structure of active conformations of UMP kinase from Dictyostelium discoideum suggest phosphoryl transfer is associative. *Biochemistry* **36**, 9290–9296.

Schirmer, T. & Evans, P. R. (1990). Structural basis of the allosteric behaviour of phosphofructokinase. *Nature* **343**, 140–145.

Schmidt, M., Rutkat, K., Rachel, R., Pfeifer, G., Jaenicke, R., Viitanen, P., Lorimer, G. & Buchner, J. (1994). Symmetric complexes of GroE chaperonins as part of the functional cycle. *Science* **265**, 656–659.

Schweins, T. & Wittinghofer, A. (1994). Structures, interactions and relationships. *Current Biology* **4**, 547–550.

Schweins, T., Geyer, M., Scheffzek, K., Warshel, A., Kalbitzer, H. R. & Wittinghofer A. (1995). Substrate-assisted catalysis as a mechanism for GTP hydrolysis of p21ras and other GTP-binding proteins. *Nature Structural Biology* **2**, 36–44.

Schweins, T., Geyer, M., Kalbitzer, H. R., Wittinghofer, A. & Warshel, A. (1996). Linear free energy relationships in the intrinsic and GAP activating protein-stimulated guanosine 5′-triphosphate hydrolysis of p21ras. *Biochemistry* **35**, 14225–14231.

Schulz, G. (1992). Binding of nucleotides by proteins. *Current Op. Struct. Biol.* **2**, 61–67.

Schulz, G. E., Elzinga, M., Marx, F. & Schirmer, R. H. (1974). Three-dimensional structure of adenylate kinase. *Nature* **250**, 120–123.

Schulz, G. E., Müller, C. W. & Diederichs, K. (1990). Induced-fit movements in adenylate kinases. *J. Mol. Biol.* **213**, 627–630.

Schultz, P, Olland, S., Oudet, P. & Hancock, R. (1996). Structure and conformational changes of DNA topoisomerase II visualized by electron microscopy. *Proc. Natl. Acad. Sci. USA* **93,**

Schulze-Gahmen, U., De Bondt, H. L. & Kim, S.-H. (1996). High resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design. *Journal of Medicinal Chemistry* **39**, 4540–4506.

Scopes, R. K. (1978). Binding of substrates and other anions to yeast phosphoglycerate kinase. *Eur. J. Biochem.* **91**, 119–129.

Sharma, A., Hauai, R. & Mondragon, A. (1994). Crystal structure of the amino-terminal fragment of vaccinia virus DNA topoisomerase I at 1.6 Å resolution. *Structure* **2**, 767–777.

Shirakihara, Y. & Evans, P. R. (1988). Crystal structure of the complex of phosphofructokinase from *Escherichia coli* with its reaction products. *J. Mol. Biol.* **204**, 973–994.

Shuman, S. (1996). Closing the gap on DNA ligase. *Structure* 4, 653–656.

Sicheri, F., Moarefi, I. & Kuriyan, J. (1997). Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385**, 602–609.

Sicherie, F. & Kuriyan, J. (1997). Structures of Src-family tyrosine kinases. *Curr. Op. Struc. Biol.* **7**, 777–785.

Silva, M. M., Poland, B. W., Hoffmann, C. R., Fromm, H. J. & Honzatko, R. B. (1995). Refined crystal structures of unligated adenylosuccinate synthetase from *Escherichia coli*. *J. Mol. Biol.* **254**, 431–446.

Smith, C. A. & Rayment, I. (1996). X-ray structure of the magnesium(II)·ADP·Vanadate complex of the *Dictyostelium discoideum* myosin motor domain to 1.9 Å Resolution. *Biochemistry* **35**, 5404–5417.

Smith, K. E. & Fisher, M. T. (1995). Interactions between the GroE chaperonins and rhodanese. Multiple intermediates and release and rebinding. *J. Biol. Chem.* **270**, 21517–21523.

Sondek, J., Lambright, D. G., Noel, J. P., Hamm, H. E. & Sigler, P. B. (1994). GTPase mechanism of G proteins from the 1.7-Å crystal structure of transducin aGDPAIF$_4^-$. *Nature* **372**, 276–279.

Sousa, R. (1996). Structural and mechanistic relationships between nucleic acid polymerases. *TIBS* **21**, 186–190.

Sousa, R., Chung, Y. J., Rose, J. P. & Wang, B. C. (1993). Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Å resolution. *Nature* **364**, 593–599.

Sozzi, G., Veronese, M. L., Negrini, M., Baffa, R., Cotticelli, M. G., Inoue, H., Tornielli, S.,

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Pilotti, S., Dr Gregorio, L., Pastorini, U., Pierotti, M. A., Ohta, M., Huebner, H. & Croce, C. M. (1996). The FHIT gene 3p14.2 is abnormal in lung cancer. *Cell* **85**, 17–26.

Sozzi, G., Sard, L., De Gregorio, L., Marchetti, A., Musso, K., Buttitta, F., Tornielli, S., Pellegrini, S., Veronese, M. L., Manenti, G., Incarbone, M., Chella, A., Angeletti, C. A., Pastorino, U., Huebner, K., Bevilaqua, G., Pilotti, S., Croce, C. M. & Pierotti, M. A. (1997a). Association between cigarette somking and FHIT gene alterations in lung cancer. *Cancer Res.* **57**, 2121–2123.

Sozzi, G., Tornielli, S., Tagliabue, E., Sard, L., Pezzella, F., Pastorini, U., Minoletti, F., Pilotti, S., Radcliffe, C., Veronese, M. L., Goldstraw, P., Huebner, K., Croce, C. M. & Pierotti, M. A. (1997b). Absence of Fhit protein in primary lung tumors and cell lines with FHIT gene abnormalities. *Cancer Res.* **57**, 5207–5212.

Stehle, T. & Schulz, G. E. (1992). Refined structure of the complex between guanylate kinase and its substrate GMP at 2.0 Å resolution. *J. Mol. Biol.* **224**, 1127–1141.

Steitz, T. A. (1998). A mechanism for all polymerases. *Nature* **391**, 231–232.

Steitz, T. A., Shoham, M. & Bennett, W. S. (1981). Structural dynamics of yeast hexokinase during catalysis. *Phil. Trans. R. Soc. London* **293**, 43–52.

Steitz, T. A. & Steitz, J. A. (1993). A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. USA* **90**, 6498–6502.

Steves, R. C., Gouaux, J. E. & Lipscomb, W. N. (1990). 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-ångstrom resolution. *Biochemistry* **29**, 7691–7701.

Stewart, L., Redinbo, M. R., Qiu, X., Hol, W. G. J. & Champoux, J. J. (1998). A model for the mechanism of human topoisomerase I. *Science* **279**, 1534–1541.

Story, R. M., Weber, I. T. & Steitz, T. A. (1992). Structure of the recA protein-ADP complex. *Nature* **355**, 374–376.

Story, R. M. & Steitz, T. A. (1992). The structure of the *E. coli* RecA protein monomer and polymer. *Nature* **355**, 318–325.

Sträter, N., Lipscomb, W. N., Klabunde, T. & Krebs, B. (1996). Two-metal ion catalysis in enzymatic acyl- and phosphoryl-transfer reactions. *Angewandt Chemie/International Edition in English* **35**, 2024–2055.

Stuart, D. I., Lavine, M., Muirhead, H. & Stammers, D. K. (1979). Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Ångstroms. *J. Mol. Biol.* **134**, 109–142.

Subramanya, H. S., Bird, L. E., Brannigan, J. A. & Wigley, D. B. (1996a). Crystal structures of DExx box helicase. *Nature* **384**, 379–383.

Subramanya, H. S., Doherty, A. J., Ashford, S. R. & Wigley, D. B. (1996b). Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7. *Cell* **85**, 607–615.

Suzuki, C. K., Rep, M., van Dijl, J. M., Suda, K., Grivell, L. A. & Schatz, G. (1997). ATP-dependent proteases that also chaperone protein biogenesis. *TIBS* **22**, 118–123.

Takusagawa, F., Kamitori, S. & Markham, G. D. (1996). Structure and function of S-adenosylmethionine synthetase with ADP, BrADP, and PPi at 2.8 Å resolution. *Biochemistry* **35**, 2586–2596.

Tampé, R. (1998). Das 'ABC' der Immunologie – Funktion des Peptid-Transportkomplexes TAP bei der zellulären Immunerkennung. *Biospektrum* **1**, 23–28.

Tanaka, T., Saha, S. K., Tomomori, C., Ishima, R., Lin, D., Tong, K. I., Park, H., Dutta, R., Quin, L., Swindell, S. M. B., Yamazaki, T., Ono, A. M., Kainosho, M., Inouye, M. & Ikura, M. (1998). NMR structure of the histidine kinase domain of the *E. coli* osmosensor EnvZ. *Nature* 396, 88–92.

Tang, W. J. & Gilman, A. G. (1995). Construction of a soluble adenylyl cyclase activated by G(S)Alpha and forskolin. *Science* **268**, 1769–1772.

Tari, L. W., Matte, A., Pugazhenthi, U., Goldie, H. & Delbaere, L. T. (1996). Snapshot of an enzyme reation intermediate in the structure of the ATP-MG^{2+}-oxalate ternary complex of *Escherichia coli* PEP carboxykinase. *Nat. Struct. Biol.* **3**, 355–363.

Taylor, S. S., Knighton, D. R., Zheng, J., Sowadski, J. M., Gibbs, C. S. & Zoller, M. J. (1993). A template for the protein kinase family. *TIBS* **18**, 84–89.

Taylor, S. S. & Radzio-Anselm, E. (1994). Three protein kinase structures define a common motif. *Structure* **2**, 345–355.

Teplyakov, A., Sebastiao, P., Obmolova, G., Perrakis, A., Brush, G. S., Bessman, M. J. & Wilson, K. S. (1996). Crystal structure of bacteriophage T4 deoxynucleotide kinase with its substrates dGMP and ATP. *EMBO J.* **15**, 3487–3497.

Tesmer, J. J. G., Klem, T. J., Deras, M. L., Davisson, V. J. & Smith, J. L. (1996). The crystal structure of GMP synthetase reveals a novel catalytic triad and is a structural paradigm for two enzyme families. *Nature Struct. Biol.* **3**, 74–86.

Tesmer, J. J. G., Sunahara, R. K., Gilman, A. G. & Sprang, S. R. (1997). Crystal structure of the catalytic domains of adenylyl cyclase in a complex with $G_{s\alpha}$·GTPγS. *Science* **278**, 1907–1916.

Tesmer, J. J. G., Sunahara, R. K., Johnson, R. A., Gosselin, G., Gilman, A. G. & Sprang, S. R.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

(1999). Two-metal-ion catalysis in adenylyl cyclase. *Science* **285**, 756–760.

The C. elegans Sequencing Consortium. (1998). Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* **282**, 2012–2018.

Theyssen, H., Schuster, H.-P., Packschies, L., Bukau, B. & Reinstein, J. (1996). The second step of ATP binding to DnaK induces peptide release. *J. Mol. Biol.* **263**, 657–670.

Thoden, J. B., Holden, H. M., Wesenberg, G., Raushelm, F. M. & Rayment, I. (1997). Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product. *Biochemistry* **36**, 6305–6316.

Tsai, F. T. F., Singh, O. M. P., Skarzynski, T., Wonacott, A. J., Weston, S., Tucker, A., Pauptit, R. A., Breeze, A. L., Poyser, J. P., O'Brien, R. O., Ladbury, J. E. & Wigley, D. B. (1997). The high-resolution crystal structure of a 24-kDa gyrase B fragment from *E. coli* complexed with one of the most potent coumarin inhibitors, Chlorobiocin. *Proteins* **28**, 41–52.

Tsai, M.-D. & Yan, H. (1991). Mechanism of adenylate kinase: site-directed mutagenesis versus X-ray and NMR. *Biochemistry* **30**, 6806–6818.

Traut, T. W. (1994). The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. *Eur. J. Biochem.* **222**, 9–19.

Valenzuela, J. G., Charlab, R., Galperin, M. Y. & Ribeiro, J. M. C. (1998). Purification, cloning, and expression of an apyrase from the bed bug *Cimex lectularius*. *J. Biol. Chem.* **273**, 30583–30590.

Virgilio, L., Schuster, M., Gollin, S. M., Veronese, M. L., Ohta, M., Huebner, K. & Croce, C. M. (1996). FHIT gene alteration in head and neck squamous cell carcinomas. *Proc. Natl. Acad. Sci USA* **93**, 9770–9775.

Volz, K. & Matsumura, P. (1991). Crystal structure of *Escherichia coli* CheY refined at 1.7-Ångstrom resolution. *J. Biol. Chem.* **266**, 15511–15519.

Vonrhein, C., Schlauderer, G. J. & Schulz, G. W. (1995). Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. *Structure* **3**, 483–490.

Waldrop, G. L., Rayment, I. & Holden, H. M. (1994). Three-dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase. *Biochemistry* **33**, 10249–10256.

Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951.

Wang, J. C. (1996). DNA topoisomerases. *Annu. Rev. Biochem.* **65**, 635–692.

Wang, J., Sattar, A. K. M. A., Wang, C. C., Karam, J. D., Konigsberg, W. H. & Steitz, T. A. (1997). Crystal structure of a pol α replication DNA polymerase from bacteriophage RB69. *Cell* 89, 1089–1098.

Wassarman, D. A. & Steitz, J. A. (1991). Alive with DEAD proteins. *Nature* **349**, 463–464.

Watenpaugh, K. D., Sieker, L. C., Jansen, L. H., Legall, J. & Dubourdieu, M. (1972). Structure of the Oxidized form of a Flavodoxin at 2.5 Å Resolution: resolution of the phase ambiguity by anomalous scattering. *Proc. Natl. Acad. Sci. USA*, **69**, 3185–3188.

Watson, H. C., Bryant, T. N., Walker, N. P., Shaw, P. J. & Wendell, P. L. (1977). The active site of yeast phosphoglycerate kinase. *Biochem. Soc. Transactions* **5**, 652–654.

Wawrzynow, A., Banecki, B. & Zyklicz, M. (1996). The Clp ATPases define a novel class of molecular chaperones. *Molec. Microbiol.* **21**, 895–899.

Weber, C. H., Park, Y. S., Sankar, S., Kent, C. & Ludwig, M. (1999). A prototypical cytidylyltransferase: CTP: glycerol-3-phosphate cytidlyltransferase from *Bacillus subtilis*. *Structure* **7**, 1113–1124.

Weber, D. J., Bhatnagar, S. K., Bullions, L. C., Bessman, M. J. & Mildvan, A. S. (1992). NMR and isotope exchange studies of the site of bound cleavage in the MutT reaction. *J. Biol. Chem.* **267**, 16939–16942.

West, S. C. (1996). DNA Helicases: new breeds of translocating motors and molecular pumps. *Cell* **86**, 177–180.

Wierenga, R. K. & Hol, W. G. J. (1983). Predicted nucleotide-binding properties of p21 protein and its cancer-associated variant. *Nature* **302**, 842–844.

Wigley, D. B., Davies, G. J., Dodson, E. J., Maxwell, A. & Dodson, G. (1991). Crystal structure of an N-terminal fragment of the DNA gyrase B protein. *Nature* **351**, 624–629.

Wilbanks, S. M. & McKay, D. B. (1995). How potassium affects the activity of the molecular chaperone Hsc70. *J. Biol. Chem.* **270**, 2251–2257.

Wild, K., Bohner, T., Aubry, A., Folkers, G. & Schulz, G. E. (1995). The three-dimensional structure of thymidine kinase from herpes simplex virus type 1. *FEBS Lett.* **368,** 289–292.

Williams, J. C., Weijland, A., Gonfloni, S., Thompson, A., Courtneidge, S. A., Superti-Furga, G. & Wierenga, R. K. (1997).The 2.35 Å Crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions. *J. Mol. Biol.* **274**, 757–775.

Wilson, K. P., Fitzgibbon, M. J., Caron, P. R., Griffith, J. P., Chen, W., McCaffrey, P. G., Chambers, S. P. & Su, M. S. (1996). Crystal structure of p38 mitogen-activated protein kinase. *J. Biol. Chem.* **271**, 27696–27700.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press

Wolodko, W. T., Fraser, M. E., James, M. N. G. & Bridger, W. (1994). The crystal structure of succinyl-CoA synthetase from *Escherichia coli* at 2.5-Å resolution. *J. Biol. Chem.* **269**, 10833–10890.

Xie, X. L., Gu, Y., Fox, T., Coll, J. T., Fleming, M. A., Markland, W., Caron, P. R., Wilson, K. P. & Su, M. S.-S. (1998). Crystal structure of NJK3: a kinase implicated in neuronal apoptosis. *Structure* **6**, 983–991.

Xu, R. M., Carmel, G., Sweet, R. M., Kuret, J. & Cheng, X. (1995). Crystal structure of casein kinase-1, a phosphate-directed protein kinase. *EMBO J.* **14**, 1015–1023.

Xu, W., Harrison, S. C. & Eck, M. J. (1997). Three-dimensional structure of the tyrosine kinase Src. *Nature* 385, 595–602.

Xu, Z., Horwich, A. L. & Sigler, P. B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)-chaperonin complex. *Nature* **388**, 741–750.

Yamaguchi, H. & Hendrickson, W. A. (1996). Struc tural basis for activation of the human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature* **384**, 484–489.

Yamaguchi, H., Kato, H., Hata, Y., Nishioka, T., Kimura, A., Oda, J. & Katsube, Y. (1993). Three-dimensional structure of the glutatione synthetase from *Escherichia coli* B at 2.0 Å Resolution. *J. Mol. Biol.* **229**, 1083–1100.

Yamashita, M. M., Almassy, R. J., Janson, C. A., Cascio, D. & Eisenberg, D. (1989). Refined atomic model of glutamine synthetase at 3.5 åangstrom resolution. *J. Biol. Chem.* **264**, 17681–17690.

Yan, H., Shi, Z. & Tsai, M.-D. (1990). Mechanism of adenylate kinase. Structural and functional demonstration of arginine-138 as a key catalytic residue that cannot be replaced by lysine. *Biochemistry* **29**, 6385–6392.

Yan, S.-Z., Hahn, D., Huang, Z.-H. & Tang, W.-J. (1996). Two cytoplasmic domains of mammalian adenylyl cyclase form a Gsα- and forskolin-activated enzyme *in vitro*. *J. Biol. Chem.* **271**, 10941–10945.

Yang, W. & Steitz, T. A. (1995). Crystal structure of the site-specific recombinase gamma-delta resolvase complex with a 34bp cleavage site. *Cell* **82**, 193–207.

Yao, N. H., Hesson, T., Cable, M, Hong, Z., Kwong, A. D., Le, H. V. & Weber, P. C. (1997). Structure of the hepatitis C virus RNA helicase domain. Nat. Struct. Biol. 4, 463–467.

Yoshida, M. & Amano, T. (1995). A common topology of proteins catalyzing ATP-triggered reactions. *FEBS Lett.* **359**, 1–5.

Yu, R. C., Hanson, P. I., Jahn, R. & Brünger, A. T. (1998). Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP. *Nat. Struc. Biol.* **5**, 803–811.

Zarembinski, T. I., Hung, L.-W., Mueller-Dieckman, H.-J., Kim, K.-K., Yokota, H., Kim, R. & Kim, S.-H. (1998). Structure-based assignment of the biochemical function of a hypothetical protein: A test case of structural genomics. *Proc. Natl. Acad. Sci. USA* **95**, 1518–15193.

Zhang, F., Strand, A., Robbins, D., Cobb, M. H. & Goldsmith, E. J. (1994). Atomic structure of the MAP kinase ERK2 at 2.3 Å resolution. *Nature* **367**, 704–711.

Zhang, G., Liu, Y., Ruoho, A. E. & Hurley, J. H. (1997). Structure of the adenylyl cyclase catalytic core. *Nature* **386**, 247–253.

Zhang, P., Toyoshima, C., Yonekura, K., Green, N. K. D. & Stokes, D. L. (1998). Structure of the calcium pump from sarcoplasmic reticulum at 8 Å resolution. *Nature* **392**, 835–839.

Zheng, J., Knighton, D. R., Xuong, N.-H., Taylor, S. S., Sowadski, J. M. & Ten Eyck, L. F. (1993). Crystal structure of the myristylated catalytic subunit of cAMP-dependent protein kinase reveal open and closed conformation. *Protein Science* **2**, 1559–1573.

Zhou, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W. R. & Chapman, M. S. (1998). Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. *Proc. Natl. Acad. Sci. USA* **95**, 8449–8454.

Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E. & Hendrickson, W. A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* **272**, 1606–1614.

https://doi.org/10.1017/S0033583599003480 Published online by Cambridge University Press