Protein Science (1998), 7:1768–1771. Cambridge University Press. Printed in the USA.

Enzyme-mononucleotide interactions: Three different folds share common structural elements for ATP recognition

KONSTANTIN A. DENESSIOUK,[1,2,3] JUKKA V. LEHTONEN,[1,2] AND MARK S. JOHNSON[1,2]

[1]Department of Biochemistry and Pharmacy, Åbo Akademi University, BioCity 3A, P.O. Box 66, FIN-20521 Turku, Finland
[2]Centre for Biotechnology, University of Turku and Åbo Akademi University, BioCity 5B, FIN-20521 Turku, Finland
[3]Moscow Institute of Physics and Technology, 141700, Dolgoprudny, Russia

(RECEIVED April 9, 1998; ACCEPTED May 5, 1998)

Abstract

Three ATP-dependent enzymes with different folds, cAMP-dependent protein kinase, D-Ala:D-Ala ligase and the α-subunit of the $\alpha_2\beta_2$ ribonucleotide reductase, have a similar organization of their ATP-binding sites. The most meaningful similarity was found over 23 structurally equivalent residues in each protein and includes three strands each from their β-sheets, in addition to a connecting loop. The equivalent secondary structure elements in each of these enzymes donate four amino acids forming key hydrogen bonds responsible for the common orientation of the "AMP" moieties of their ATP-ligands. One lysine residue conserved throughout the three families binds the alpha-phosphate in each protein. The common fragments of structure also position some, but not all, of the equivalent residues involved in hydrophobic contacts with the adenine ring. These examples of convergent evolution reinforce the view that different proteins can fold in different ways to produce similar structures locally, and nature can take advantage of these features when structure and function demand it, as shown here for the common mode of ATP-binding by three unrelated proteins.

Keywords: allosteric effector-binding site; ATP-dependent enzymes; convergent structural similarities; local structural comparisons; similar ligand-binding sites

Adenosine 5′-triphosphate (ATP) is necessary for the function of a wide variety of enzymes. Currently, the Brookhaven Protein Data Bank (PDB) (Bernstein et al., 1977) contains atomic coordinates for over 140 X-ray and NMR structures of ATP-dependent proteins solved as complexes mainly with bound ATP-analogues, bound AMP, bound ADP, and in some cases with bound ATP.

Several distinct classes of folds of ATP-dependent enzymes have been identified (Schulz, 1992). A characteristic feature of these folds is the presence of a β-sheet, parallel or antiparallel, in the vicinity of the ATP-binding site. One class of folds, the classic mononucleotide-binding fold (Schulz & Schirmer, 1974; Rossmann et al., 1975; Schulz et al., 1986), has an ATP-binding site placed at the edge of a parallel β-sheet. Two other classes of folds, the protein kinase family fold (Hanks et al., 1988; Hanks & Quinn, 1991; Hubbard et al., 1994; Xu et al., 1995), and the glutathione synthetase fold (Yamaguchi et al., 1993; Waldrop et al., 1994; Wolodko et al., 1994; Fan et al., 1995; Artymiuk et al., 1996; Herzberg et al., 1996; Hibi et al., 1996; Matsuda et al., 1996; Thoden et al., 1997; Esser et al., 1998), also called the "ATP-grasp" fold (Murzin, 1996), have ATP-binding sites positioned on the surface of an antiparallel β-sheet (Kobayashi & Go, 1997b). These last two fold classes are particularly interesting. Although the folds are different, two of their typical members, cAMP-dependent protein kinase (cAPK) from the one family (Bossemeyer et al., 1993) and D-Ala:D-Ala ligase (DD-ligase) from the other family (Fan et al., 1994, 1997), have over 100 structurally-equivalent residues from ten segments that form two identical supersecondary structures between which the cofactor ATP is bound in a similar way (Denessiouk et al., 1998).

Reprint requests to: Mark S. Johnson, Department of Biochemistry and Pharmacy, Åbo Akademi University, Tykistökatu 6, BioCity 3A, P.O. Box 66, FIN-20521 Turku, Finland; e-mail: mark.johnson@btk.utu.fi.

Abbreviations: ATP, adenosine 5′-triphosphate; DD-ligase, D-Ala:D-Ala ligase (EC 6.3.2.4); cAPK, cAMP-dependent kinase (EC 2.7.1.37); RNR R1, R1 subunit of ribonucleotide reductase (EC 1.17.4.1); PDB, Brookhaven Protein Data Bank; RMSD, root-mean-square deviation.

Results and discussion

We have used structure-based alignment techniques and a newly developed computer program GENFIT (Lehtonen JV, Denessiouk KA, Johnson MS, unpubl. results) to compare the structure of cAPK (PDB code: 1CDK, A chain) against 146 ATP-dependent protein complexes in the Brookhaven Protein Data Bank on the basis of the optimal superposition of the ligands bound to the ATP-binding sites. cAPK itself and nine additional kinase-related complexes with similar folds (1ATP, 1CSN, 1FIN, 1GOL, 1HCK, 1IR3, 1JST, 1PHK, and 1AD5) have been identified among the 146 complexes. Among the proteins having folds distinct from the kinase family fold, cAPK was found to share partial structural similarities with DD-ligase (1IOW) and four other homologous proteins (1IOV, 1AUX, 1GSA,

```
                    &                                  □ *
cAPK          56 RVMLVK 61   S2(55-62)        68 HFAMKI 73   S3(67-75)
DD-ligase    153 GMSKVV 158  S7(155-157)     140 PVIVKP 145  S6(141-145)
RNR R1        56 STKIGD 51   S3(50-54)         5 LLVTKR 10   S1(6-8)

                  #*□*
cAPK         116 LYMVMEYVPGG 126  S5(115-120)    128 MFSHLRR 134  HD(128-135)
DD-ligase    176 EVLIEKWLSGP 186  S8(176-181)
RNR R1        11 DGSTERINLDK 21   S2(14-16)      24 RVLDWAE 30    H1(19-30)
```

Fig. 1. Structure-based alignment of 23 structurally equivalent residues in three segments that were found to be common to all three enzymes, and seven additional equivalent amino acids superimposed in cAPK and RNR R1. The start–stop positions of the secondary structures are noted at the right of the segments according to the X-ray crystallographic data (Bossemeyer et al., 1993; Eriksson et al., 1997; Fan et al., 1997). Positions of equivalent residues forming hydrogen bonds and hydrophobic interactions with the cofactors are marked by "*" and "□," respectively. Positions of equivalent residues Met154 (in DD-ligase) and Val57 (in cAPK), participating in hydrophobic interactions with the cofactors, and Thr55 (in RNR R1), which binds to the 4′, 5′, and α-phosphate group oxygens, are indicated with an "&." Positions of equivalent residues Glu180 (in DD-ligase) and Glu15 (in RNR R1), forming equivalent hydrogen bonds with the N6 purine nitrogens are labeled with a "#."

and 2DLN) from the glutathione synthetase family fold. The extent of similarity between cAPK and DD-ligase has already been reported (Kobayashi & Go, 1997a, 1997b; Denessiouk et al., 1998). Surprisingly, the comparison also revealed an unexpected structural similarity between the cofactor binding site in cAPK and the allosteric effector binding site in the α-subunit of ribonucleotide reductase (RNR R1) (3R1R, A chain). RNR R1 is responsible for converting ribonucleotides to deoxyribonucleotides for use in DNA synthesis during cell proliferation and ATP functions as a general positive allosteric regulator (Eriksson et al., 1997).

The first 100 residues of the N-terminal domain of the RNR R1 α-subunit form a four-helix bundle and a three-stranded mixed β-sheet between which the allosteric active site cleft is located (Eriksson et al., 1997). The most meaningful similarity common to cAPK, DD-ligase and RNR R1 was found over 23 structurally equivalent residues in each of these three proteins (Fig. 1). In RNR R1, these residues correspond to most of the N-terminal three-stranded β-sheet, while in cAPK and DD-ligase, these residues reside within their antiparallel β-sheets. In addition to these strands, the loops located at the ends of β-strands S5 in cAPK, S8 in DD-ligase and S2 in RNR R1 are held in common. Interestingly, seven additional amino acids from α-helices HD in cAPK and H1 in RNR R1 are superimposed too (Fig. 2A). No structure equivalent to HD and H1 is present in DD-ligase, instead, β-strand S8 is followed by strand S9, which veers off in a different direction than the HD and H1 helices. The similarities resulting from the superposition of these enzymes and their bound ATP-ligands, in terms of the root-mean-square deviation (RMSD) over equivalent sets of superposed atoms, are given in Table 1. The RMSD values were obtained with the "Fit Atoms" command of the computer program SYBYL (Tripos Associates, Inc., St. Louis, Missouri).

Since the relative positions of the secondary structure elements about the ATP-binding sites in cAPK, DD-ligase and RNR R1 are similar, and since the "AMP" moieties of the ligands in the three enzyme complexes closely superimpose too, it is perhaps not too surprising that the key interactions between the proteins and the ATP-ligands are similar as well. In fact, the adenine bases are bound to the enzymes through equivalent hydrogen bonds between the N1 nitrogen of the purine and equivalent backbone amide NH groups of Val123 in cAPK, Leu183 in DD-ligase and Asn18 in RNR R1, and between the N6 nitrogen of the purine and the backbone carbonyls of Glu121 (cAPK), Lys181 (DD-ligase), and Arg16 (RNR R1) (Figs. 2A,C). In addition, the nitrogenous bases are fixed by interactions with Glu180 and Glu15 in DD-ligase and RNR R1, respectively, where the N6 nitrogen of the purine forms hydrogen bonds with oxygen OE2 of glutamate in each case (Fig. 2C). The oxygen O2 of the pentose ring is tightly bound via hydrogen bonds to the side chains of Glu127 (cAPK) and Glu187 (DD-ligase), and the main-chain carbonyl oxygen of Lys21 in RNR R1 (Figs. 2B,C). The α-phosphate group forms equivalent hydrogen bonds with the side chains of Lys72 (cAPK), Lys144 (DD-ligase), and Lys9 (RNR R1). These lysine residues are conserved throughout each of these three families (Hanks et al., 1988; Fan et al., 1994; Stapleton et al., 1996; Eriksson et al., 1997; Galperin & Koonin, 1997; Dideberg & Bertrand, 1998; Esser et al., 1998).

The adenine moieties of the ATP-ligands are also stabilized by a number of equivalent hydrophobic interactions (Figs. 2B,C). Nonpolar interactions occur between the adenine moiety and both Ala70 and Tyr122 in cAPK, Ile142 and Trp182 in DD-ligase, and Val7 and Ile17 in RNR R1, and with Val57 (cAPK) and Met154 (DD-ligase) that are matched in the alignment to the polar residue Thr55 in RNR R1. The side-chain oxygen of Thr55 in RNR R1 serves to anchor the O4′ oxygen of the ribose, the O5′ oxygen between the ribose and the α-phosphate group, and the O2α oxygen of the α-phosphate group of bound AMPPNP. Furthermore, the hydrophobic amino acids Leu49 in cAPK, Phe209 in DD-ligase and Ile58 in RNR R1 are similarly placed about the adenine moieties of the ATP-ligands in the three proteins. Note that Leu49 and Ile58 are adjacent to β-strands S2 and S3, respectively, in the

Table 1. *The superposition of cAPK, DD-ligase, and RNR R1 and the resulting similarity in the disposition of their cofactors*

	Enzyme[a]	Adenine[a]	Ribose[a]	Pd[a]	"AMP"[a]
cAPK–DD-ligase	1.8	1.4	2.1	2.0	1.8
DD-ligase–RNR R1	2.2	0.6	1.2	1.2	0.9
cAPK–RNR R1	1.8	1.6	2.9	2.4	2.2
cAPK–RNR R1[b]	2.0	1.5	2.6	2.2	2.1

[a]RMSD values between superimposed sets of 23 Cα atoms, adenine moieties (10 atoms), ribose moieties (7 atoms), α-phosphate moieties (C5, O5, Pα, and OA, four atoms) and "AMP" moieties of cofactors (21 atoms), respectively.

[b]The superimposed sets consist of 30 Cα atoms.

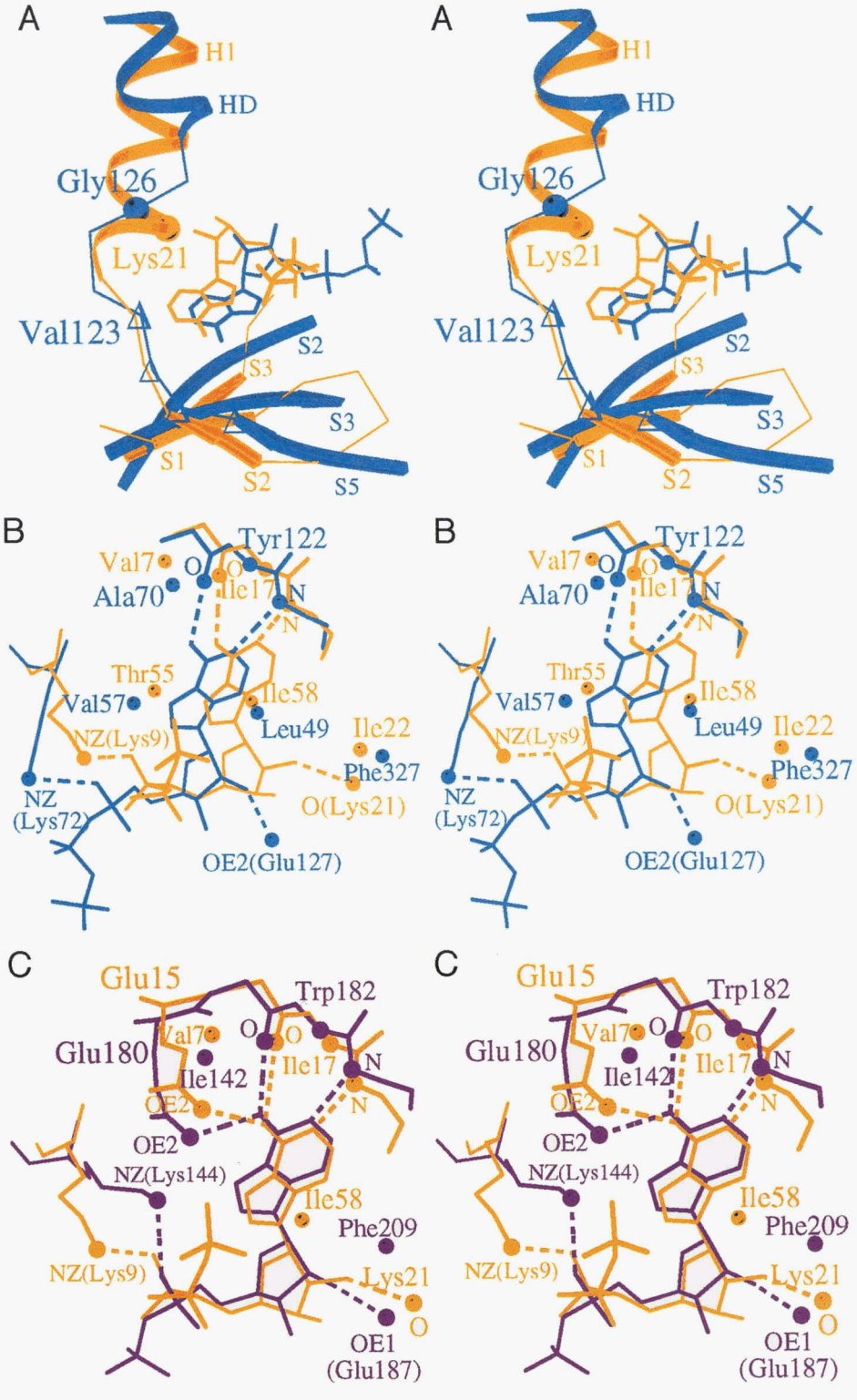

Fig. 2. The pairwise superposition of **(A)** four segments of cAPK (blue) and RNR R1 (light brown) in the alignment of Figure 1, and the equivalent polar and hydrophobic environments around the cofactors bound in **(B)** cAPK (blue) and RNR R1 (light brown) and **(C)** DD-ligase (dark red) and RNR R1 (light brown). In (**A**) four common segments from cAPK and from RNR R1 form a three-stranded β-sheet (lying below the cofactors) and an α-helix (lying above the cofactors) that are connected by a loop. Bound cofactors are shown as stick models. The Cα-atoms of Gly126 and Lys21, indicated by balls, show the limit of the common structural similarity when DD-ligase is also considered. The location of four amino acids from cAPK reported by Kobayashi and Go (1997a, 1997b) to be held in common with DD-ligase are drawn as triangles; the position of Val123 is indicated for reference. In (**B**) and (**C**) the atoms of amino acids forming equivalent hydrogen bonds (dashed lines) with the cofactors and pairs of Cα-atoms of equivalent amino acids forming a hydrophobic environment about the adenine moieties of the cofactors are drawn as circles. The figure was drawn with the program MOLSCRIPT (Kraulis, 1991).

structural alignment (Fig. 1), overlap in space in the superposed structures, and are positioned close to the side chain of the equivalent Phe209 of DD-ligase. The side chains of Phe327 in cAPK and Ile22 in RNR R1 also coincide and contribute to the common hydrophobic environments about the bound ligand in the two proteins (Fig. 2B); no counterpart is present in the structure of DD-ligase. In RNR R1, Val7, Ile22, and Ile58 have been identified as playing key roles in stabilizing the adenosine base of the bound allosteric effector (Eriksson et al., 1997).

The similarity common to all three proteins is less extensive than that between cAPK and DD-ligase (Denessiouk et al., 1998). This is mostly due to the fact that both the β- and γ-phosphates of the ATP-allosteric regulator in RNR R1 adopt a different conformation from that seen for the corresponding phosphates in either cAPK or DD-ligase (Figs. 2B,C). In the alignment of cAPK and DD-ligase (Denessiouk et al., 1998), residues from five common segments are responsible for interactions of the proteins with the β- and γ-phosphates of the ATP-cofactors. In RNR R1, these segments are not present in the structure. Instead, the common segments that are observed in these three proteins are responsible for the common interactions of the "AMP" moiety of ATP with the enzymes. In particular, this includes the equivalent key loop containing a similar four-residue segment about the adenine moieties, originally found in cAPK and DD-ligase by Kobayashi and Go (1997a), and three strands that participate in most of the interactions with the "AMP" moiety.

In our previous article (Denessiouk et al., 1998), we suggested that if the equivalent structural features found in cAPK and DD-ligase were conserved for a particular structural or functional reason, then we might observe these same common features repeatedly in other ATP-binding proteins. In reality, among the current repertoire of ATP-binding proteins in the Brookhaven Protein Data Bank, only one other family of proteins, represented by RNR R1, was found to share these common similarities involved in extensive interactions with the cofactor. Interestingly, one of the key conserved residues in each of these three families, a lysine residue, is responsible for making important interactions with the α-phosphate group of the common cofactor to these three families of enzymes. These examples of convergent evolution reinforce the view that different proteins can fold in different ways to produce similar structures locally, and nature can take advantage of these features when structure and function demand it, as shown here for the common mode of ATP-binding by three unrelated proteins.

Acknowledgments

This research was supported from grants from the Academy of Finland, the Technology Development Centre Finland 'TEKES', the Hasselblad Foundation, and the Graduate School in Informational and Structural Biology.

References

Artymiuk PJ, Poirrette AR, Rice DW, Willett P. 1996. Biotin carboxylase comes into the fold. *Nature Struct Biol 3*:128–132.

Bernstein FC, Koetzle TF, Williams GJB, Meyer EJ, Jr, Brice MD, Rodgers JK, Kennard O, Shimanouchi I, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Bossemeyer D, Engh RA, 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.

Denessiouk KA, Lehtonen JV, Korpela T, Johnson MS. 1998. Two "unrelated" families of ATP-dependent enzymes share extensive structural similarities about their cofactor binding sites. *Protein Sci 7*:1136–1146.

Dideberg O, Bertrand J. 1998. Tubulin tyrosine ligase: A shared fold with the glutathione synthetase ADP-forming family. *Trends Biochem Sci 23*:57–58.

Eriksson M, Uhlin U, Ramaswamy S, Ekberg M, Regnström K, Sjöberg B-M, Eklund H. 1997. Binding of allosteric effectors to ribonucleotide reductase protein R1: Reduction of active-site cysteines promotes substrate binding. *Structure 5*:1077–1092.

Esser L, Wang C-R, Hosaka M, Smagula CS, Südhof TC, Deisenhofer J. 1998. Synapsin I is structurally similar to ATP-utilizing enzymes. *EMBO J 17*:977–984.

Fan C, Moews PC, Shi Y, Walsh CT, Knox JR. 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.

Fan C, Moews PC, Walsh CT, Knox JR. 1994. Vancomycin resistance: Structure of D-alanine:D-alanine ligase at 2.3 Å resolution. *Science 266*:439–443.

Fan C, Park IS, Walsh CT, Knox JR. 1997. D-alanine:D-alanine ligase: Phosphonate and phosphinate intermediates with wild type and the Y216F mutant. *Biochemistry 36*:2531–2538.

Galperin MY, Koonin EV. 1997. A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. *Protein Sci 6*:2639–2643.

Hanks SK, Quinn AM. 1991. Protein kinase catalytic domain sequence database: Identification of conserved features of primary structure and classification of family members. *Methods Enzymol 200*:38–81.

Hanks SK, Quinn AM, Hunter T. 1988. The protein kinase family: Conserved features and deduced phylogeny of the catalytic domains. *Science 241*:42–52.

Herzberg O, Chen CC, Kapadia G, McGuire M, Carroll LJ, Noh SJ, Dunaway-Mariano D. 1996. Swiveling-domina mechanism for enzymatic phosphotransfer between remote reaction sites. *Proc Natl Acad Sci USA 93*:2652–2657.

Hibi T, Nishioka T, Kato H, Tanizawa K, Fukui T, Katsube Y, Oda J. 1996. Structure of the multifunctional loops in the nonclassical ATP-binding fold of glutathione synthetase. *Nature Struct Biol 3*:16–18.

Hubbard SR, Wei L, Ellis L, Hendrickson WA. 1994. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature 372*:746–754.

Kobayashi N, Go N. 1997a. ATP binding proteins with different folds share a common ATP-binding structural motif. *Nature Struct Biol 4*:6–7.

Kobayashi N, Go N. 1997b. A method for similar protein local structures at ligand-binding sites and its application to adenine recognition. *Eur Biophys J 26*:135–144.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:945–949.

Matsuda K, Mizuguchi K, Nishioka T, Kato H, Go N, Oda J. 1996. Crystal structure of glutathione synthetase at optimal pH: Domain architecture and structural similarity with other proteins. *Protein Eng 9*:1083–1092.

Murzin AG. 1996. Structural classification of proteins: new superfamilies. *Curr Opin Struct Biol 6*:386–394.

Rossmann MG, Liljas A, Branden C-I, Banaszak LJ. 1975. Evolutionary and structural relationships among dehydrogenases. In: Boyer PD, ed. *The enzyme, Vol 11*. New York: Academic Press. pp 61–102.

Schulz GE. 1992. Binding of nucleotides by proteins. *Curr Opin Struct Biol 2*:61–67.

Schulz GE, Schiltz E, Tomasselli AG, Frank R, Brune M, Wittinghofer A, Schirmer RH. 1986. Structural relationships in the adenylate kinase family. *Eur J Biochem 161*:127–132.

Schulz GE, Schirmer RH. 1974. Topological comparison of adenylate kinase with other proteins. *Nature 250*:142–144.

Stapleton MA, Javid-Majd F, Harmon MF, Hanks BA, Grahmann JL, Mullins LS, Raushel FM. 1996. Role conserved residues within the carboxy phosphate domain of carbamoyl phosphate synthetase. *Biochemistry 35*:14352–14361.

Thoden JB, Holden HM, Wesenberg G, Raushel FM, Rayment I. 1997. Structure of carbamoyl phosphate synthetase: A journey of 96 Å from substrate to product. *Biochemistry 36*:6305–6316.

Waldrop GL, Rayment I, Holden HM. 1994. Three-dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase. *Biochemistry 33*:10249–10256.

Wolodko WT, Fraser ME, James MNG, Bridger WA. 1994. The crystal structure of succinyl-CoA synthetase from *Escherichia coli* at 2.5-Å resolution. *J Biol Chem 269*:10883–10890.

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

Yamaguchi H, Kato H, Hata Y, Nishioka T, Kimura A, Oda J, Katsube Y. 1993. Three-dimensional structure of the glutathione synthetase from *Escherichia coli* at 2.0 Å resolution. *J Mol Biol 229*:1083–1100.