Minireview

Protein kinases – structure and function

Dirk Bossemeyer*

Department of Pathochemistry, German Cancer Research Centre, INF 280, D-69120 Heidelberg, Germany

Received 10 May 1995

Abstract **The solution of crystal structures from half a dozen protein kinases during the last four years in different laboratories has deepened our understanding of the catalysis and regulation of this enzyme class, and given a vigorous impetus to the whole field. Due to the great degree of sequence conservation among protein kinases the informational yield with every new structure is high, as each is a representative of the enzyme family in general and most often of a subclass in particular. This review will focus on the active site structure of cAMP-dependent protein kinase (cAPK) with special regard to two new crystal structures; one of an active protein kinase CK1*, which may represent an as yet unsolved step in the kinetic pathway, and the other of the insulin receptor kinase domain, the first structure of a tyrosine kinase.**

Key words: Crystal structure; cAMP-dependent protein kinase; Protein kinase CK1; Insulin receptor; Catalytic site; Conserved sequence motif

1. Introduction

Protein kinases are critically involved in almost every regulatable cellular process. Several hundreds of members of the protein kinase family all possess a sequence of eleven conserved subdomains, the catalytic core, in conjunction with their shared task of modulating substrate proteins by phosphorylating serine (S), threonine (T) or tyrosine (Y) residues [1]. Until very recently all protein kinase crystal structures have been from members of the ST-subfamily; for reviews see [2–4]. Now Hubbard et al. have filled this gap, and reported a structure of the protein tyrosine kinase (PTK) domain of the insulin receptor IRK [5]. The various crystal structures confirm the structural conservation of the protein kinase catalytic core, at the same time demonstrating how protein kinases reconcile the high degree of conservation essential for the catalytic function with the need for individual solutions for their diverse biological roles. As a rule, conservation in sequence indicates conserved spatial structure and function, while individual properties relate to residues, loops or insertions not conserved in the family as a whole. However, exceptions are apparant. In IRK one of the most highly conserved sequence motifs, the triad Asp-Phe-Gly in subdomain VII, is displaced and part of an autoinhibitory mechanism. On the other hand, protein kinase CK1, the structure of which has been solved recently [6], lacks the conserved motif APE in subdomain VIII and a conserved arginine residue in subdomain XI without significant conformational restrictions. More surprises may be ahead.

*Corresponding author. e-mail bossemeyer@dkfz-heidelberg.de

Abbreviations: cAPK, cAMP-dependent protein kinase; CK1, protein kinase CK1, formerly casein kinase I; CKi1, variant of CK1; ERK, extracellular signal regulated kinase; IRK, kinase domain fragment of insuline receptor; CDK2, cyclin dependent kinase 2. Residue numbers of homologous residues of cAPK are in brackets.

2. The catalytic situation in PKs

The most comprehensive set of structural data on a single enzyme has been accumulated for porcine and recombinant mouse catalytic subunit of cAPK, which provided the first crystal structure of a protein kinase [7], and still is the only protein kinase which has been crystallised in ternary complexes, presumably showing the catalytic conformation with nucleotide and (pseudo)substrate bound [8–10]. Yet, we have good reason to assume that during catalysis all protein kinases will adopt a conformation very similar to that of the ternary complex of cAPK, making the cAMP-dependent PK a universal structural reference.

One or several crystal structures were also reported from cyclin dependent kinase CDK2 [11], the extracellular signal regulated kinase ERK2 [12,13] and the myosine light chain kinase twitchin [14]. The inactive states of these enzymes revealed many important features of enzyme regulation and specificity [2–4]. The recently reported structures of CKi1 and IRK now are from potentially active enzymes, in addition to providing new insights into the specific and unique features of each enzyme they allow to evaluate the active site in comparison to that of cAPK more closely.

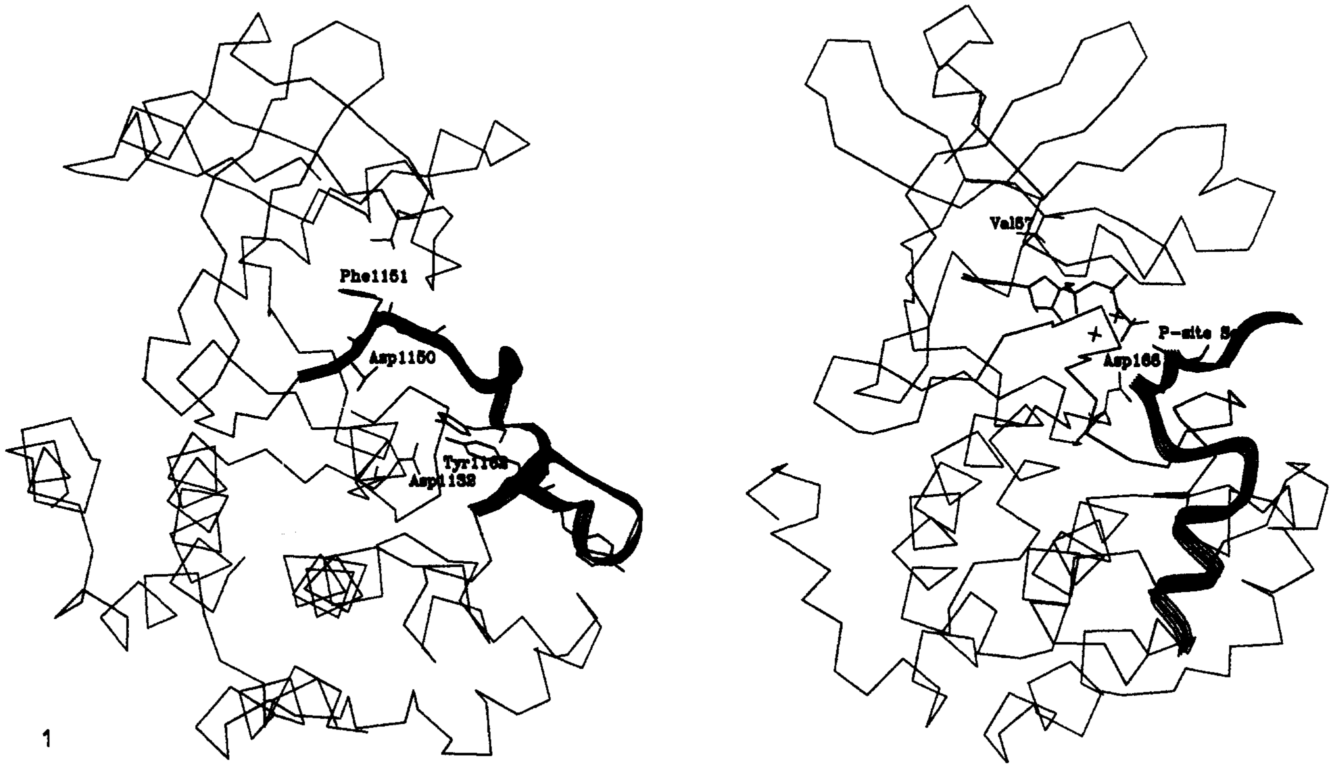
2.1. The catalytic protein kinase core

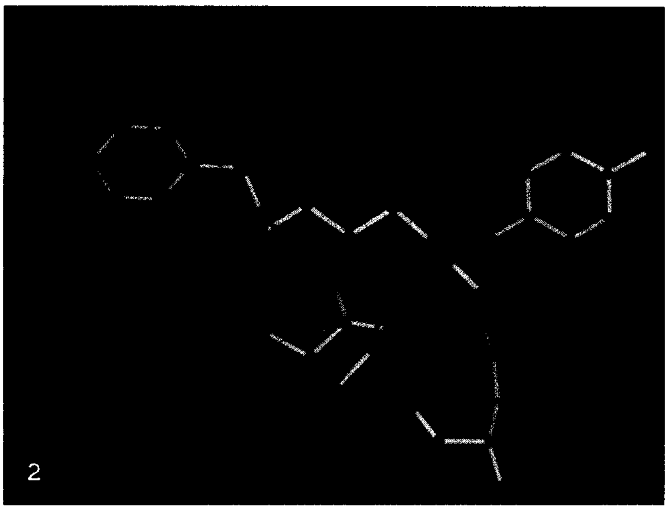
Protein kinases are highly flexible enzymes, large rotational movements of the conserved kinase lobes [15] as well as movements of flexible loops and domains accompany binding of substrates, cofactors, auto inhibitory domains, or interacting proteins, during catalysis or for regulation. The enzyme is bilobal with the active site located in the mouth region of a deep cleft between the lobes (Fig. 1). This cleft completely accommodates the ATP molecule, with the γ-phosphate oriented outwards. The protein substrate binding site is vis-à-vis at the opening of the cleft. Most of the highly conserved residues cluster in this active site, approaching from different parts of both lobes.

2.2. The N-lobe

The smaller amino terminal N-lobe is dominated by a five stranded antiparallel β-sheet (β1–β5) and a long α-helix (C). In cAPK an additional helix B inserts between strand β3 and helix C. In some PKs, including cAPK, this region is used for pro-

0014-5793/95/$9.50 © 1995 Federation of European Biochemical Societies. All rights reserved.
SSDI 0014-5793(95)00580-3





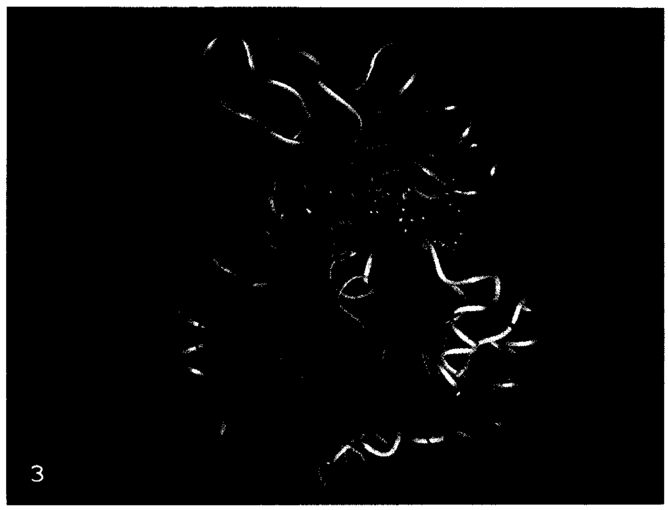


Fig. 1. Autoinhibition by IRK. The structure of IRK [5] (left) with the autophosphorylation loop (plus the DFG motif) which connects subdomain VII and subdomain VIII is depicted as a blue ribbon. While Tyr1162 is in a position seemingly poised for autophosphorylation, Phe1151 blocks the ATP-binding site. For comparison the situation in cAPK [9] (right) is shown. β-strands are in red, helices in pink, and loops in green. The autophosphorylation loops are in blue.

Fig. 2. Alignment of catalytic residues. Superimposition of residues from cAPK [9] (thin sticks) and IRK [5] (balls and sticks) as well as the P-site substrate amino acids Tyr1162 and Ala21 (modeled as serine) shows almost identical positions of the guanidinium group of Arg1136 and the ammonium group of Lys168. Phe1151 aligns with the adenosine moiety of the AMP-PNP molecule from the ternary complex of cAPK.

Fig. 3. Overlay of the structures of CKi1 [6] and cAPK [9]. Comparison of the binary MgATP-complex of CKi1 (red) with the ternary complex of cAPK with MnAMP-PNP and protein kinase inhibitor peptide PKI (white) shows non-conserved elements and conformational differences in the glycine-rich sequence above the ATP. The nucleotides are in ball and stick representation, the metal ions (two in cAPK) are bigger balls. An interaction of Ser53 in the glycine-rich sequence with the pseudosubstrate amino acid Ala17 of PKI, which possibly contributes to synergistic binding of cosubstrate and substrate in cAPK, is depicted.

tein-protein interactions. In cAPK His87 at the N-terminus of helix C contacts the C-terminal lobe via the constitutive phosphorylated Thr197. This contact is lost in the open conformation of the enzyme [15,16]. Interaction of His87 with the regulatory RI-subunit is indicated by complementary effects of mutations in His87 of the C-subunit and Ser99 in the RI-subunit [17]. Cyclin binding, a precondition for the activation of CDKs, is affected by mutations of residues in the region between 3 and helix C in CDK [18]. The first two β-strands contain the glycine-rich sequence GXGXXGXV, a turn connecting both strands is formed by the central four amino acids of this motif. The conserved glycine-rich sequence contributes in many ways to protein kinase function [19]. The invariant conserved glycines allow a close approach of the nucleotide to the peptide backbone. From the third β-strand an invariant lysine, Lys72 in cAPK, contacts α- and β-phosphoryl groups of the bound ATP. Protein kinases are most sensitive to mutations of this lysine, which probably contributes to the correct stereochemical orientation of the triphosphate and the γ-phosphate for catalysis [20,21]. The spatial position of Lys72, and in turn that of the bound phosphoryl groups, is secured by a salt bridge to Glu91, a residue in the middle of helix C. It is assumed that the interaction between Lys72 and Glu91 is conserved in the catalytic conformation of PKs. This salt bridge is missing in some structures of inactive kinases, and also in the potentially active IRK: in CDK2, helix C is rotated, and Glu51 points to the solvent. Lys33 (Lys72 in cAPK) interacts with the α-phosphate only, the β-phosphate is oriented differently. Cyclin binding might allow the active site residues and the triphosphate to achieve the active conformation found in cAPK. In the twitchin kinase, Lys5971 (Lys72) interacts with residues of a C-terminal intrasteric autoinhibitory peptide, which traverses the active site and is assumed to dislodge on calmodulin binding [22].

2.3. The C-lobe

Several residues of the larger helical C-terminal lobe interact with the triphosphate group of ATP. The conserved core of this lobe has seven α-helices in cAPK, and four short β-strands, with a central bundle of four antiparallel α-helices. The β-strands β6 to β9 form the bottom of the cleft beneath the adenosine moiety. The catalytic loop between β6 and β7, and the metal binding loop between β8 and β9 are of primary importance for the active site of PKs. The catalytic loop from Arg165 to Asn171 carries residues which are directly involved in catalysis. Here resides the catalytic base, Asp166, which is believed to abstract the proton from the hydroxyl group of the substrate amino acid. Of special significance is Lys168. Its sidechain contacts the γ-phosphate, neutralises the negative charge, and stabilises the presumed intermediate state. This residue is conserved in kinases with ST-specificity, including members of the STY-kinases [23]. In PTKs an arginine in equivalent position or two residues further in the primary sequence (Glu170 in cAPK) has been predicted to take up the functional role of this lysine [9].

The metal-binding loop contains the conserved motif DFG (Asp184PheGly). In cAPK the carboxylate of Asp184 serves as a ligand of the metal ion M2 which co-ordinates β- and γ-phosphoryl oxygens, and a pair of water molecules. Two metal ions are found liganded to ATP; one is essential for activity, the other is inhibitory and binds with low affinity. Because of its nearly optimal octahedral co-ordination scheme, and the relative number of charged ligands, the M2 site is assumed to be the primary metal binding site. Asp184 may serve to establish the stereochemistry of the triphosphate group and to neutralise charge at the γ-phosphate via the metal ion. Mutating this residue results in inactive kinase [24]. The secondary metal ion co-ordinates the α- and γ-phosphoryl groups, a water molecule, and the sidechain of invariant Asn171 at the end of the catalytic loop. The co-ordination scheme of this metal ion is trigonal bipyramidal, if one of the oxygens of Asp184 at a distance of 2.6 Å (compared to about 2.0 Å for the metal co-ordination sphere) is taken as a further weak ligand. Both metal ions potentially contribute to charge dispersal at the γ-phosphate, which is supported by kinetic data which relate the inhibitory effect of the secondary metal ion largely to a reduction in release of ADP, with little effect on the actual phosphoryl transfer reaction [25,26].

3. IRK

IRK is the fragment of the cytoplasmic kinase domain of the insulin receptor β-chain, an α2β2 heterotetramer, which is assumed to undergo a change in quarternary structure on insulin binding to the α-chain, leading to activation of the kinase and trans-autophosphorylation of a number of tyrosine residues in the β-chain [5,27]. The overall structure shows great similarity to that of the known structures of ST-kinases, confirming the universal validity of the catalytic core. IRK, although unphosphorylated, is potentially active, and becomes autophosphorylated upon addition of MgATP. The enzyme is in an open conformation, which is stabilised by steric interactions between the glycine rich sequence and the DFG motif (Fig. 1). The open conformation leads to relative displacement of catalytic residues from the small lobe compared to those of the large lobe, and also does not allow formation of the salt bridge between Lys1030 (Lys72) and Glu1047 (Glu91). Clearly rigid body rotations and in addition conformational changes are necessary for IRK to take up the catalytic conformation upon activation.

One of the most interesting aspects of this tyrosine kinase structure is the mechanism of autoinhibition it revealed. Many kinases use the loop between DFG (subdomain VII) and APE (subdomain VIII) for regulation, it differs, however, largely in size and conformation in each enzyme [2]. In IRK this loop contains the three autophosphorylation site tyrosines. We have the intriguing situation of finding one of them, Tyr1162, in the active site, occupying a position seemingly poised for autophosphorylation. However, cis-autophosphorylation cannot occur because the ATP binding site is efficiently blocked by Phe1151 of the DFG motif (Fig. 1). Apparently, 'self' binding of the activation loop transmits serious distortions to the DFG motif. Hubbard et al. suggest trans-autophosphorylation dislodges the autoinhibitory loop, in agreement with biochemical findings [28,29].

Although not in the catalytic conformation, the structure of IRK allows examination of interactions of the residues of the catalytic loop. The hydroxyl group of Tyr1162 is hydrogen bonded to the carboxylate of Asp1132 (Asp166), supporting the role of this residue as the catalytic base in the kinase reaction [8,9,30], in contradiction to the results of others [31]. In addition, the hydroxyl group has contact to the guanidinium group of Arg1136 (168 + 2), as a direct evidence for the proposed [9] functional replacement of the catalytic loop lysine Lys168 (cAPK) by arginine. Superimposition of the kinase structures shows similar locations of the guanidinium group of this arginine and the ammonium group of Lys168 in cAPK (Fig. 2). Apparently a positive charge is conserved throughout the protein kinase family in three-dimensional space, rather then in the primary sequence.

The bound autophosphorylation substrate also provides an explanation how tyrosine selectivity is determined. The rather long sidechain of a tyrosine residue compared to serine or threonine is compensated for by a different position of its main chain, governed by PTK specific residues of the P + 1 loop.

The IRK mechanism of autoinhibition by a real substrate, instead of a pseudosubstrate, is novel, and may apply for other PTKs as well [5]. However, inhibition by a bona fide substrate is also known from cAPK. The RII-subunit inhibits the C-subunit in the absence of ATP, but becomes readily phosphorylated in the presence of ATP, with little change in the inhibitory potential [32], adding an 'inhibitory product' to the varied collection of protein kinase regulatory mechanisms.

4. CKi1

The crystal structure of a truncated variant, CKi1, of protein kinase CK1, a kinase ubiquitous to all eukaryotes, in its MgATP complex has been solved [6]. The protein is unphosphorylated, though active and the peptide binding site is fully accessible. Two sulfate ions from the crystallisation medium show potential phosphate binding sites, one of them being homologous to the Thr197 phosphate binding site in cAPK. The CKi1 substrate selectivity is directed towards phosphate groups N-terminal of the P-site. The other sulfate ion, bound in the cleft, seems to indicate the binding site for this substrate recognition phosphate. The overall architecture resembles the known protein kinases, with largest differences observed in surface loops. Again it is a loop between subdomain VII and VIII (L-9D), which contains three potential phosphorylation sites. In contrast to CDK2, where the unphosphorylated T-loop blocks the active site, this loop in CKi1 is in an open conformation, allowing access to the substrate binding site (Fig. 3). Xu et al. discuss the possibility of an in vivo mechanism to stabilise L-9D in a CDK2 like conformation for inactivation, perhaps in connection with a cell cycle dependent localisation of CK1. The absence of the APE motif (which is Ser186IleAsn in CKi1) seems to be compensated for by an H-bond from Ser186 and Asn188 to an arginine and an aspartate of helix E. A novel ion pair between Glu202 of helix E and Arg261 of helix H may confer additional stability.

The potentially active CKi1 may represent the conformation of the first step in the kinetic pathway, the enzyme–MgATP complex, which has not been solved for any other active kinase yet. The structure is in a closed conformation, with the glycine-rich sequence tilted away from the cleft, relative to its position in ternary cAPK (Fig. 3). This supports previous postulations of induced-fit movements of the glycine-rich sequence to explain the synergism of substrate binding in cAPK [9,19]. The conformation of the nucleotide is very similar to that found in the ternary complex of cAPK, β- and γ-phosphates appear somewhat more flexible. The adenosine moiety is placed in a similar hydrophobic pocket like in cAPK, with identical H-bonds from N1 and N6 to main chain groups of the enzyme. A third H-bond between N7 and the sidechain of Thr183 (cAPK) is not established in CKi1, perhaps indicating a site for specific interaction with CK1 inhibitors [33]. The triphosphate interactions resemble those in ternary cAPK, with identical contacts to Lys41 (72), Lys133 (168), Asp154 (184), and Ser22 (53) and Gly24 (55) from the glycine-rich sequence. Also formed is the salt bridge between Lys41 and Glu55 (Glu91), and the catalytic loop is retained. This together indicates that very little conformational changes are necessary, to adopt the catalytic conformation found in cAPK. However, in contrast to cAPK, electron density for only one Mg^{2+} ion is observable in CKi1. It locates in the position of the secondary M1 metal site of cAPK, interacts to oxygens of both α- and γ-phosphates and to the sidechain of Asn136 (Asn171). The observation of a single metal ion in the presumed low affinity site is somewhat surprising: Xu et al. [6] discuss the possibility that the primary metal binding site may not be conserved in protein kinases, and perhaps can vary between M1 and M2. Another possibility may be that the metal position is just less defined in the binary complex, and M2 becomes occupied after peptide substrate binding. Rather weak metal binding may be indicated by the relatively large co-ordination sphere distances in CKi1. Still unclear is the role of the metal ligands asparagine or aspartate in the absence of 'their' metal ion at low magnesium. The invariance of both residues suggests essential functions. For the asparagine, however, a 'metal free' role can be imagined: with M1 unoccupied, Asn171 in cAPK is still able to contact the phosphoryl oxygens as via the metal ion in order to stabilise the triphosphate conformation.

The solution of protein kinase crystal structures in the future may complete not only our present picture of the enzymatic pathway, but very likely will give us also new insights into more possible variations on a common conserved catalytic core.

Acknowledgements: Atomic coordinates of CKi1 were kindly provided by Xiaodong Cheng, the coordinates of IRK were taken from Brookhaven Protein Data Bank entry pdb1irk.ent. I should like to express my thanks to Volker Kinzel and Robert Huber for continuous support.

References

[1] Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) Science 241, 42–51.
[2] Morgan, D.O. and De-Bondt, H.L. (1994) Curr. Opin. Cell. Biol. 6, 239–246.
[3] Taylor, S.S. and Radzio-Andzelm, E. (1994) Structure 2, 345–355.
[4] Goldsmith, E.J. and Cobb, M.H. (1994) Curr. Opin. Struct. Biol. 4, 833–840.
[5] Hubbard, S.R., Wei, L., Ellis, L. and Hendrickson, W.A. (1994) Nature 372, 746–754.
[6] Xu, R.-M., Carmel, G., Sweet, R.M., Kuret, J. and Cheng, X. (1995) EMBO J. 14, 1015–1023.
[7] Knighton, D.R., Zheng, J.H., Ten-Eyck, L.F., Ashford, V.A., Xuong, N.H., Taylor, S.S. and Sowadski, J.M. (1991) Science 253, 407–414.
[8] Zheng, J., Knighton, D.R., Ten-Eyck, L.F., Karlsson, R., Xuong, N., Taylor, S.S. and Sowadski, J.M. (1993) Biochemistry 32, 2154–2161.
[9] Bossemeyer, D., Engh, R.A., Kinzel, V., Ponstingl, H. and Huber, R. (1993) EMBO J. 12, 849–859.
[10] Madhusudan, Trafny, A., Xuong, N.H., Adams, J.A., Ten-Eyck, L.F., Taylor, S.S. and Sowadski, J.M. (1994) Protein Sci. 3, 176–187.
[11] De-Bondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgan, D.O. and Kim, S.H. (1993) Nature 363, 595–602.
[12] Zhang, F.M., Strand, A., Robbins, D., Cobb, M.H. and Goldsmith, E.J. (1994) Nature 367, 704–711.
[13] Zhang, J., Zhang, F., Ebert, D., Cobb, M.H. and Goldsmith, E.J. (1995) Structure 3, 299–307.
[14] Hu, S.H., Parker, M.W., Lei, J.Y., Wilce, M.C.J., Benian, G.M. and Kemp, B.E. (1994) Nature 369, 581–584.
[15] Cox, S., Radzio-Andzelm, E. and Taylor, S.S. (1994) Curr.. Opin. Struct. Biol. 4, 893–901.

[16] Zheng, J.H., Knighton, D.R., Xuong, N.H., Taylor, S.S., Sowadski, J.M. and Ten-Eyck, L.F. (1993) Protein Sci. 2, 1559–1573.
[17] Cox, S. and Taylor, S.S. (1994) J. Biol. Chem. 269, 22614–22622.
[18] Ducommun, B., Brambilla, P., Felix, M.-A., Franza Jr, B.R., Karsenti, E. and Draetta, G. (1991) EMBO J. 10, 3311–3319.
[19] Bossemeyer, D. (1994) Trends Biochem. Sci. 19, 201–205.
[20] Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammarström, L., Kinnon, C., Levinsky, R., Bobrov, M, Smith, C.I.E. and Bentley, D.R. (1993) Nature 361, 226–233.
[21] Bossemeyer, D. (1993) Nature 363, 590.
[22] Kemp, B.E. and Pearson, R.B. (1991) Biochim. Biophys. Acta 1094, 67–76.
[23] Hanks, S.K. and Quinn, A.M. (1991) Methods Enzymol. 200, 38–81.
[24] Gibbs, C.S. and Zoller, M.J. (1991) J. Biol. Chem. 266, 8923–8931.
[25] Adams, J A. and Taylor, S.S. (1993) Protein Sci. 2, 2177–2186.
[26] Kong, C.-T. and Cook, P.F. (1988) Biochemistry 27, 4795–4799.
[27] Mcdonald, N.Q., Murrayrust, J. and Blundell, T.L. (1995) Structure 3, 1–6.
[28] Frattali, A.L., Treadway, J.L. and Pessin, J.E. (1992) J. Biol. Chem. 267, 19521–19528.
[29] Levy-Toledano, R., Caro, L.H.P., Accili, D. and Taylor, S.I. (1994) EMBO J. 13, 835–842.
[30] Yoon, M.-Y. and Cook, P.F. (1987) Biochemistry 26, 4118–4125.
[31] Adams, J.A. and Taylor, S.S. (1993) J. Biol. Chem. 268, 7747–7752.
[32] Bossemeyer, D., Kinzel, V. and Reed, J., submitted
[33] Chijwa, T., Hagiwara, M. and Hidaka, H. (1989) J. Biol. Chem. 264, 4924–4927.