Protein kinases and their therapeutic exploitation

L. Johnson¹
Laboratory of Molecular Biophysics, University of Oxford, U.K., and Diamond Light Source, Chilton, Oxon., U.K.

Abstract
This review focuses on the recognition properties of protein kinases at the molecular level. Phosphorylation of the substrate protein by a protein kinase can result in enzyme activation or inhibition, conformational changes that change recognition properties, or the creation of a surface with distinct binding properties. Protein kinases have become important targets for the development of inhibitors with potential therapeutic application. Various examples are considered in this review, and I discuss our own work on glycogen phosphorylase and phosphorylase kinase, and the structures of proteins involved with the cell cycle, including cyclins and cyclin-dependent kinases.

Introduction
Most protein kinases are switched on in response to a cellular signal. They catalyse the transfer of the γ-phosphate from ATP to a serine, threonine or tyrosine hydroxy group in the target protein. The change from a non-phospho- to a phosho-protein can result in activation of an enzyme, inhibition of an enzyme, conformational changes that change recognition properties, or the creation of a surface with distinct binding properties. Protein kinases have become important targets for the development of inhibitors with potential therapeutic application. Various examples are considered in this review, and I discuss our own work on glycogen phosphorylase and phosphorylase kinase, and the structures of proteins involved with the cell cycle, including cyclins and cyclin-dependent kinases.

Key words: cyclin, cyclin-dependent kinase (CDK), glycogen phosphorylase, phosphorylase kinase, protein kinase; protein phosphorylation.

Abbreviations used: APC, anaphase-promoting complex; CDK, cyclin-dependent kinase; EGF(R), epidermal growth factor (receptor); PDGFR, platelet-derived growth factor receptor; pRb, retinoblastoma family protein; VEGF(R), vascular endothelial growth factor (receptor).

¹email louise.johnson@biop.ox.ac.uk

Glycogen phosphorylase and phosphorylase kinase
Phosphorylase kinase was the first protein kinase to be described, by Ed Fisher and Ed Krebs in 1955 [1]. Phosphorylase kinase phosphorylates glycogen phosphorylase. Phosphorylation turns phosphorylase from its inactive state to its active state, allowing it to metabolize glycogen to supply energy for muscle contraction.

We solved the structure of the enzyme in each of its two states, the unphosphorylated glycogen phosphorylase b (GPb) and the phosphorylated glycogen phosphorylase a (GPa), and showed that there was a significant conformational change between them [2]. On phosphorylation of a single serine residue (Ser-14) in the N-terminal region, the phosphate group docks at the interface between the monomers of this dimeric molecule, drawing the subunits together at the interface. The negative charge of the phosphate group interacts with positive charges of two arginine residues, one from one subunit and the other from the other subunit. Changes at this interface region are accompanied by changes in other interface regions, which signal down to the catalytic site over 40 Å (≈4 nm) away, thus perturbing the equilibrium from an inactive to an active conformation, in a remarkable demonstration of an allosteric mechanism first envisaged by Monod, Wyman and Changeux [3]. Thus we could understand how addition of a single phosphate group to a protein that contains 847 amino acids could produce such profound effects. Subsequent work with other systems has shown that stabilization of phosphate groups by clusters of arginine residues is a common feature for tight phosphate-binding sites.

Our structure of phosphorylase kinase (Figure 1) shows the basic two-lobe kinase fold, with the N- and C-terminal lobes joined by a hinge region [4]. ATP binds at a site between the two lobes, and the main recognition for the substrate protein is through interaction with the activation segment, a region in the C-terminal lobe. From the structure of phosphorylase kinase in complex with a peptide similar to that recognized in phosphorylase, we can understand the enzyme’s specificity, in terms of the basic and non-polar residues arranged either side of the serine that is to be phosphorylated.

The structure of phosphorylase kinase has enabled us to understand many of its recognition properties. But it has not provided us with all the answers. In particular, we do not yet back to the original state. It is remarkable that the addition of a four-atom phosphate group, to a protein containing tens of thousands of atoms, can have such a profound effect.
understand why the kinase is much more efficient against the intact glycogen phosphorylase, with both a higher $k_{cat}$ and a much lower $K_m$ than with a small peptide substrate. This suggests that there are additional sites in the intact protein that the kinase recognizes and which improve its catalytic efficiency, but knowledge of the structures has not yet yielded information on these sites. Many other kinases, including the cyclin-dependent kinases (CDKs), utilize substrate-binding regions that are remote from the catalytic site to improve catalytic efficiency.

**Protein kinases and drug design**

Protein kinases are intimately involved in cell signalling, and defects in these processes result in many important diseases, including cancer, inflammation and diabetes. Several protein kinase genes are mutated in cancer cells, leading to uncontrolled activation of cell proliferation [5]. Genes that are defective in cancer include proteins involved with the control of cell cycle protein kinases, such as cyclins D and E, inhibitors of protein kinases (e.g. p16) or their substrates (e.g. pRb (retinoblastoma family protein)), or genes involved in regulation at checkpoint controls (e.g. p53). Hence protein kinases have become important targets for the development of inhibitors with potential therapeutic application.

The steps in designing a potential drug involve selection of an appropriate target, generation of a library of candidate molecules as potential inhibitors and rapid screening with an effective assay. Knowledge of the three-dimensional structure of the target protein is most useful at the stage of selecting a lead compound from the hits and then modifying it into a potent inhibitor. Structure can contribute most effectively to the design of enzyme inhibitors, and all major pharmaceutical companies now have a structural biology programme. Structure can show where a compound can be modified to improve potency and where it should not be modified. It can explain mechanism of action in terms of detailed stereochemical interactions and conformation. But for a compound to be a successful drug, it must have ‘drug-like’ properties, i.e. low molecular mass, appropriate hydrophobic properties, not too many hydrogen-bond donors or acceptors, and acceptable solubility; it must exhibit potency, efficacy in cell-based assays, selectivity against related targets and acceptable toxicity; and it must also exhibit suitable pharmacokinetic properties, such as cell permeability, stability in vivo, suitable clearance properties and bioavailability. It must have accessible groups for synthetic modifications, and there must be a clear patent strategy. Making an inhibitor a drug is not an easy process.

Brown and Flocco [6] recently published a table showing drugs that are currently in the clinic where the structure of the drug target is known and has either informed drug development or explained the mechanism of action. There are over 60 such compounds, and 25 of these are among the top 200 compounds by sales. Structure was the main inspiration behind about seven drugs. These include: captopril for regulating blood pressure, based on knowledge of the active site of carboxypeptidase; Trusopt for glaucoma, based on the structure of carbonic anhydrase; HIV protease inhibitors developed on the basis of the viral enzyme’s structure; and the anti-influenza compounds Relenza and Tamiflu, based on the knowledge of the structure of influenza virus neuraminidase.

There is a growing list of drugs in the clinic that target protein kinases (reviewed in [7]). The first to be licensed, in 1999, was Fasudil, a Rho kinase inhibitor developed for the treatment of cerebral vasospasm. Very recently, determination of its structure has shown how this compound binds to the ATP-binding site of the active confirmation of the kinase, and shows sites where its potency can be improved. The best example of a clinically important kinase inhibitor is undoubtedly the Abl tyrosine kinase inhibitor, Glivec, for the treatment of Philadelphia chromosome positive chronic myeloid leukaemia. The compound was developed at Ciba Geigy based on clever chemistry starting from the pyrimidine–aminophenyl scaffold and rapid physiological assays. The determination of the Abl structure by John Kuriyan and colleagues showed that Glivec bound to the ATP-binding site of the kinase, but extended into a pocket behind the C-helix of the N-terminal lobe. The binding mode could only be achieved in the inactive confirmation of the kinase. Because chronic myeloid leukaemia is a disease that arises from a single genetic defect that creates the aberrant Abl kinase, the kinase inhibitor Glivec is enormously effective and the response can be 100% if the disease is caught at an
early stage. However, as time has gone on, there are instances of resistance to Glivec. Structure has been informative in showing the significance of the mutations. In particular, mutation of a threonine residue that guards the pocket behind the C-helix to a leucine prevents the binding of Glivec and allows the aberrant cells to escape from the drug. Glivec is also effective against another tyrosine kinase, c-Kit, and against the PDGFR (platelet-derived growth factor receptor) kinase. Its inhibitory properties against c-Kit allowed swift approval of the drug for the treatment of gastrointestinal stromal tumours.

In 2004, the two EGFR [EGF (epidermal growth factor) receptor] tyrosine kinase inhibitors, Iressa and Tarceva, were approved as anticancer agents. The structure of Tarceva bound to its target showed that it binds to an active conformation of the kinase. Certain patients have been found to respond more favourably to treatment with Iressa than others. These patients have mutations that activate the kinase domain of EGFR. This may explain why drugs that bind to the active conformation are more effective in these patients. Two final examples are Sorafenib and Sunitinib, both licensed against renal cell carcinoma. Sorafenib was originally developed against Raf kinase for melanoma treatment, but both these drugs inhibit a number of kinases, including VEGFR [VEGF (vascular endothelial growth factor) receptor] and PDGFR, and this widening of specificity has allowed them to be effective against angiogenesis.

Structure has also been influential in explaining the interactions of monoclonal antibodies to tyrosine kinases as anticancer drugs. Herceptin, licensed for Her2 receptor-positive breast cancer, is the best known of these. Herceptin binds to the extracellular domain of the EGFR at domain IV, the domain closest to the membrane where it could interfere with signalling, but it is thought that its major action is as an antibody to target the cells for the immune response. In contrast, the antibody Erbitux was shown to compete with the EGF-binding region of the receptor, but keeping it in its closed inactive confirmation and hence preventing a response. Finally, the anti-VEGF antibody, Avastin, binds to the growth factor itself at a site that is close to the receptor-binding site, hence preventing interaction.

**CDKs (cyclin-dependent kinases)**

Our work at Oxford has led to structures of proteins involved in the cell cycle, both cyclins and CDKs. In response to mitotic signals, cells progress from the quiescent phase G0 to enter the first growth phase (G1), where the centrosomes are duplicated. After passing the checkpoint, cells enter S phase, where the chromosomes are duplicated. There follows a second growth or gap phase (G2) leading to a further checkpoint, after which cells enter mitosis (M), cell division in which the chromosomes are distributed equally to two daughter cells. Each transition is catalysed by CDKs: the G2/G1 transition by the D-type cyclins in association with CDK4 or CDK6; the G1/S transition by CDK2/cyclin E; S phase and G2 by CDK2/cyclin A; and the G2/M transition by CDK1/cyclin B. There are strict regulatory controls on the cyclins both for their synthesis by transcription and their degradation by ubiquitinylation signals that target them to the proteasome.Thus cyclin levels cycle during the cell cycle. The CDKs are also controlled by phosphorylation on a threonine residue in the activation segment. The ‘master kinase’ that phosphorylates CDK4, CDK6, CDK2 and CDK1 is CDK7 in association with cyclin H and MAT1.

We have considered CDK2 as a potential target for anticancer drugs. (This work was carried out by Tom Davies, Jane Endicott, Martin Noble and others in Oxford, and in collaboration with Herbie Newell and Roger Griffin and their colleagues at the University of Newcastle.) CDK2 is not generally mutated in cancer, and is not essential for the cell cycle. Mouse knockout experiments have shown that cells in which CDK2 has been ablated can complete the cell cycle, presumably because its functions are taken over by CDK1, and possibly by CDK4 [8]. Most CDK2 inhibitors also inhibit CDK1, and cells in which both these kinases have been ablated arrest in G2. Furthermore, in tumours where the tumour suppressor protein pRb has been mutated, CDK2 inhibition prevents exit from S phase and leads to apoptosis, and in those aggressive tumours where cyclin E is overexpressed, inhibition of its interaction partner, CDK2, can help mitigate its effects. There are reasons why this kinase might be a target for anticancer drugs, and several companies have compounds in clinical trials showing proof of principle.

We have solved the structure of phospho-CDK2 bound to cyclin A and have targeted the ATP-binding site. ATP makes two essential hydrogen bonds to the hinge region of the kinase. Our collaborators in Newcastle discovered that a compound, Nu2058 (Figure 2a), which they had developed as a thymidylate synthase inhibitor, actually inhibited CDK2. It was not very potent, but it was selective for CDK2 and CDK1 over CDK4. Structural studies showed that the most productive way of modifying this compound would be to build down into a pocket towards Lys-89. The resulting compound, Nu6102 (Figure 2b), has an IC50 of 5 nM against CDK2 and 10 nM against CDK1 [9]. It has a sulfamoylanilino substituent bound into this pocket, with the anilino group making two essential hydrogen bonds to the hinge region of the protein in the hinge. The orientation of this anilino ring against the purine ring is exact, which is observed in small-molecule crystal structures of compounds of this type, indicating that there is no entropy penalty when it binds. This compound is effective in inhibiting cell growth with an IC50 of 5 µM. It is less effective in cells than as an inhibitor in vitro, probably because in cells it must compete with millimolar concentrations of ATP. Our collaborators are introducing further modifications to improve bioavailability and activity in cell culture.

A further strategy to arrest the cell cycle could be inhibition of the ‘master’ CDK, CDK7. Graziano Lolli, Ed Lowe and Nick Brown in my group solved the structure of this enzyme [10]. Comparing its ATP-binding site with that of CDK2, we saw that the part of this site that accommodates the adenine and ribose is almost identical, but that the equivalent residue to Lys-89 is a valine. This, together with other sequence
Figure 2 | Structures and binding mechanisms of the CDK2 inhibitors Nu2058 (a) and Nu6102 (b)
The Figure shows how modifying Nu2058 by building into a pocket towards residue Lys-89 has produced a more potent inhibitor.

changes, makes the pocket into which the sulfamoylanilino group of Nu6102 would bind more open and non-polar. It is not surprising that this inhibitor is about three orders of magnitude less potent against CDK7 than against CDK2. Inhibitor development is progressing.

The role of cyclins in the cell cycle

Why does the cell need both cyclins E and A? Both cyclins associate with and activate CDK2. In 2003, mouse knockout experiments showed that cyclin E is not essential for cycling cells, but is essential for cells to enter the cell cycle from the quiescent phase G0. Cyclin E-deficient mice die in utero.

Interest in cyclin E is stimulated by the demonstration that cyclin E is frequently up-regulated in cancer, particularly in some of the more aggressive cancers. Furthermore, and remarkably, cyclin E-deficient cells are resistant to oncogenic transformation by Ras, showing the importance of cyclin E in transformation.

Both cyclin E and cyclin A activate CDK2 by binding to it and promoting the key conformational changes in both the C-helix and the activation segment of CDK2 (Figure 3). Active CDK2 specifically phosphorylates the serine residue in a Ser/Thr-Pro-Xaa-Arg/Lys motif, where 'Xaa' is any amino acid, but there is a preference for arginine or lysine in this position. Structural studies led by Nick Brown showed that the proline docks into a pocket created by the activation segment and a lysine in the P + 3 position of the substrate interacts with the phosphorylated threonine [11]. The residue in the P + 2 position (arginine in our peptide complex) points into solution, thus explaining a preference for a charged polar group in this position.

In order to find a favourable construct that would allow CDK2–cyclin E crystallization, we used a construct based on a stable fragment obtained by limited proteolysis of full-length cyclin E that was similar to one that had been observed by Keyomarsi et al. [12] in breast cancer cells. We co-crystallized and solved the structure of CDK2 bound to this fragment of cyclin E to 2.25 Å resolution [13] (Figure 3). There are no significant differences between phospho-CDK2 when bound to either cyclin E or cyclin A, an observation that is consistent with kinetic experiments, which had demonstrated no significant difference in kinetic parameters between phospho-CDK2–cyclin E and phospho-CDK2–cyclin A. Cyclin E makes slightly more contacts with...
CDK2, and this is consistent with observations that, in cells, almost all cyclin E is associated with CDK2, whereas cyclin A distributes between CDK2 and CDK1.

The main difference between cyclins A and E is in the way in which they are targeted for degradation; cyclin A is recognized by the APC (anaphase-promoting complex) through its ‘destruction box’, whereas cyclin E is recognized by the SCF (Skp1/cullin/F)-box complex through phosphorylation. In both cases the result is ubiquitylation, which targets the complex for breakdown in the proteasome. In several cancer cell lines, mutations have been observed in residues in the F box protein that recognize the phospho-Thr-380 residue of cyclin E [14], thus allowing cyclin E to become immortal and to promote chromosomal instabilities by untimely activation of CDK2.

In 2002, experiments with a cell free system showed that CDK2/cyclin E had to be added to CDK2 before cyclin A, rather than the other way around, if DNA synthesis was to occur [15], and that the crucial step probably involved the assembly of the pre-replication complex before the start of S (DNA-replication) phase. The formation of the pre-replication complex involves the recruitment of the ORC (origin of replication complex) on to DNA, which then recruits proteins Cdc6 and Cdt1, followed by recruitment of the MCM (mini-chromosome maintenance) proteins, MCM1–7, to form a functional complex. Ironically, Cdc6 is a substrate for the APC in complex with its Cdh1 subunit. APC(Cdh1) becomes active at anaphase, and remains active as cells re-enter G1. On exit to G0, Cdc6 is degraded and has to be resynthesized. As it is a substrate for APC(Cdh1), Cdc6 would be degraded as soon as it was formed, unless there were mechanisms to prevent it. CDK2/cyclin E phosphorylates Cdc6 at two sites in the N-terminal region that are close to the ‘destruction box’ signals of Cdc6 [16]. This prevents recognition of Cdc6 and its degradation and allows assembly of the pre-replication complex. CDK2–cyclin A cannot catalyse this reaction in the cell, because cyclin A is a substrate for APC(Cdh1) and levels of cyclin A only increase on entry into S phase when the Cdh1 subunit of the APC is degraded. Thus the differences between cyclin E and cyclin A lie not in their structures (overall they are 29% identical in sequence) or in their substrate recognition properties, but in their different mechanisms for ubiquitylation that allow these cyclins to be present at different times during the cell cycle. However, it is possible that cyclin E has other roles that are not dependent on CDK2. Cyclin E has been shown to bind to the centrosomes in a non-CDK-dependent manner, and that this process is also important for DNA synthesis [17]. We identified this binding region on cyclin E, and showed that it is partly exposed on a loop that has a major difference in sequence and conformation to the region in cyclin A. It is not yet known which of the many proteins of the centrosome is recognized. There is still much to be discovered about the molecular mechanisms by which the cell cycle is regulated.

Most of this work over the years has been supported by the Medical Research Council together with other organizations. I am most grateful for their continuing support. I also wish to acknowledge my splendid colleagues whose names are in the publications and without whom none of this work would have been possible.

References


Received 31 August 2006