Control by phosphorylation
Louise N Johnson* and Marc O'Reilly

The two examples of phospho and dephospho proteins for which structural data were previously available (glycogen phosphorylase and isocitrate dehydrogenase) demonstrated two different mechanisms for control. In glycogen phosphorylase, activation by phosphorylation results in long-range allosteric changes. In isocitrate dehydrogenase, inhibition by phosphorylation is achieved by an electrostatic blocking mechanism with no conformational changes. During the past year, the structures of the phospho and dephospho forms of two more proteins, the cell cycle protein kinase CDK2 and yeast glycogen phosphorylase, have been determined. The new results highlight the importance of the phosphoamino acids both in the organization of local regions of protein structure through phosphate–arginine interactions and in the promotion of long-range conformational responses.

Introduction

The modification of proteins by phosphorylation and dephosphorylation reactions is the major mechanism for the regulation of enzyme activity in the response of the cell to extracellular signals. Binding of hormones, such as adrenaline or insulin, or mitogens, such as growth factors, to their cognate receptors on the outside of the cell membrane stimulates a series of reactions and recognition phenomena on the inside of the cell that results in the modification of proteins by a cascade of phosphorylation/dephosphorylation reactions. Protein phosphorylation is now recognized to affect diverse and important processes such as metabolism, growth, differentiation, motility, membrane transport, learning and memory. Phosphorylation provides a reversible process in which the forward and the backward reactions are catalyzed by different enzymes operating with different specificities so that reactions can be turned on or off in response to different stimuli. Phosphorylation on single or multiple sites can elicit a variety of molecular responses; it can result in enzyme activation or inhibition; it can alter the association/dissociation properties of protein–protein assemblies; and it can alter the surface-recognition properties of a protein.

The response of a protein upon phosphorylation is dictated by the special properties of the phosphate group that distinguish it from the naturally occurring amino acids. The phosphate group, with four oxygen atoms, can participate in extensive hydrogen-bond interactions which can link different parts of the polypeptide chain. The phosphate group (pK_a~6.7) is likely to be dianionic at physiological pH. The property of a double-negative charge is one that is not available to the naturally occurring amino acids, and electrostatic effects are therefore important in control by phosphorylation. Analysis of the protein–phosphate interactions in existing protein structures [1,2] has shown that the most common interaction is between the phosphate oxygens and the main-chain nitrogens at the start of a helix [3], where the most frequently found residue is glycine. In non-helix interactions, phosphate groups most commonly interact with arginine residues. The guanidinium group is suited for interactions with phosphate by virtue of its planar structure and its ability to form multiple hydrogen bonds [4]. Because of its resonance stabilization, the guanidinium group is a poor proton donor (pK_a~12) and cannot function as a general-acid catalyst in the hydrolysis of phosphorylated amino acids. Electrostatic interactions between arginine and phosphate groups provide tight binding sites that appear to play a dominant role in recognition and stabilization of protein conformations.

In this review we summarize our current understanding of the structural basis of control by phosphorylation with reference to the new work on the cell cycle protein kinase CDK2 in complex with cyclin A (CDK2–cyclin A: [5**,6**]), and on yeast glycogen phosphorylase (yGP: [7**]) and compare the mechanisms demonstrated by these enzymes to those already established for rabbit muscle glycogen phosphorylase (rmGP: [8]) and isocitrate dehydrogenase (IDH; reviewed in [9]). We shall not discuss the recent structural work on the bacterial phosphotransferase system which involves phosphoryl transfer between histidine residues. Work in this field has recently been summarized by Liao et al. [10].

Phosphate recognition sites and control mechanisms

The phosphate recognition sites in the four enzyme systems, and their mechanism for control are summarized in Table 1 and in Figure 1 and several features are apparent. In rmGP, yGP and CDK2–cyclin A, activation by phosphorylation involves significant conformational changes in the protein structure and the rearrangement or generation of new protein–protein interactions. It is...
Control by phosphorylation

Johnson and O'Reilly 763

noteworthy that each of the phosphate recognition sites in these three enzymes involve two or three arginine residues and that the constellation of positively charged residues provides a tight phosphate binding site. In rmGP, the two arginines come from different subunits of the physiologically active dimer and the interactions help strengthen the subunit-subunit interface (Fig. 1a). In yGP (Fig. 1b) the phosphate recognition site is partially buried and localization of the phosphoamino acid (Thr-P10) and surrounding residues also locates a hydrophobic core for subunit-subunit interactions. In CDK2-cyclin A, the arginine residues shift on phosphorylation to form an almost ideal coordination site for the phosphothreonine residue (Fig. 1c).

In rmGP, yGP and CDK2-cyclin A each of the arginine residues is involved in contacts to other side chains so that the phosphate-arginine interactions are connected to other parts of the polypeptide chain. This extends the sphere of influence of the phosphate group from its immediate vicinity. For example, in rmGP, Arg69 is hydrogen bonded to Gln72, which in turn is hydrogen bonded to Asp42' (Fig. 1a) and in CDK2-cyclin A, Arg50 and Arg150 form hydrogen bonds to main-chain carbonyl oxygen atoms on cyclin A (Fig. 1c).

The three enzymes rmGP, yGP and CDK2-cyclin A exhibit different mechanisms for activation by phosphorylation but there is a common theme in that the phosphate group of each enzyme directs the organization of the local peptide region and thereby promotes more long-range changes (Table 1).

Isocitrate dehydrogenase is unique (Table 1 and Fig. 1d). Phosphorylation of Ser113 inhibits the enzyme by an electrostatic blocking mechanism. The phosphate group prevents binding of the charged citrate substrate but it does not use any of the charged amino acids in the vicinity to promote binding of the phosphate. The phosphate group may interact with the positive charge at the N-terminal end of the helix which starts with Ser115 but the contacts from the phosphate oxygens to the main-chain nitrogens are >3.7Å. The effect of the Ser-phosphate can be demonstrated by replacing Ser113 with glutamate or aspartate which reduces activity by a factor of 107 [11]. Recent results have shown that the inactivation by an acidic residue can be suppressed by mutation of a nearby residue, Asn115, to leucine [12]. Structural results indicate that the leucine mutation causes a conformational change that moves the negative charge away from the catalytic site. Although the constellation of atoms observed in the native conformation can lead to an effective control mechanism, it appears that this system may easily be perturbed by mutagenesis.

 Cyclin-dependent protein kinase CDK2

The eukaryotic cell cycle is coordinated by several Ser/Thr protein kinases, each consisting of a catalytic cyclin-dependent protein kinase (CDK) subunit and a regulatory cyclin subunit. Levels of cyclin vary throughout the cell cycle, and this provides a temporal control on the kinase activity. The CDK-cyclin complex may exhibit some activity but full kinase activity is dependent upon the phosphorylation of the protein kinase by the CDK activating kinase, CAK. CDK2, the CDK for which

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Phosphoamino acid</th>
<th>Phosphate contact residues</th>
<th>Mechanism</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rabbit muscle glycogen phosphorylase</td>
<td>Ser14</td>
<td>Arg43' Arg69</td>
<td>Activation by allosteric transitions. On phosphorylation, the N-terminal residues shift with a change in conformation from mobile to ordered, which is accompanied by changes in the tertiary and quaternary structure that lead to activation at the catalytic site, 45 Å away from the Ser14-phosphate site. Ser14 shifts 50 Å on phosphorylation.</td>
<td>[8]</td>
</tr>
<tr>
<td>Yeast glycogen phosphorylase</td>
<td>Thr10</td>
<td>Arg309 Arg310</td>
<td>Activation is achieved by competition of the phosphoamino acid with the allosteric inhibitor, glc-6-P, for the same phosphate recognition site. The large conformational change of the N-terminal peptide relieves steric blocking by the N-terminal residues of the catalytic site, and there are changes at the subunit interfaces and some movements between the two domains.</td>
<td>[7**]</td>
</tr>
<tr>
<td>CDK2-cyclin A</td>
<td>Thr160</td>
<td>Arg50 Arg126 Arg150</td>
<td>Phosphorylation results in reordering of the activation segment, which shifts by as much as 7 Å. This affects the putative substrate-binding site and promotes stronger CDK2-cyclin A interactions.</td>
<td>[5**,6**]</td>
</tr>
<tr>
<td>Isocitrate dehydrogenase</td>
<td>Ser113</td>
<td>None</td>
<td>Inhibition is achieved by an electrostatic blocking mechanism with no conformational changes, in which the serine-phosphate partially occupies the site recognized by the negatively charged substrate, isocitrate, but does not make strong stabilizing interactions with positively charged residues in the vicinity.</td>
<td>[9]</td>
</tr>
</tbody>
</table>
structural data are available, is important at the G1/S boundary of the cell cycle, where it is complexed with cyclin E, and at the start of S phase, where it is complexed with cyclin A.

The crystal structure of the nonphosphorylated inactive form of CDK2 was solved in 1993 [13]. This study showed, as anticipated from the wealth of comparative-sequence data on the protein kinases, that the fold of CDK2 was
similar to that of the archetypal protein kinase, cyclic AMP dependent protein kinase (cAPK), whose structure had been solved in 1991 [14–16]. The kinase core consists of a bilobal scaffold that has an N-terminal lobe composed almost entirely of β sheet and a C-terminal lobe in which α helices dominate (Fig. 2). The two lobes are joined by a single strand of polypeptide chain. ATP is bound in a deep cleft between the lobes; the cleft contains the catalytic residues that are conserved in nearly all eukaryotic protein kinases. The activation segment spans the two highly conserved motifs present in almost all kinases (DFG to APE, using amino acid single-letter code), and is an important control element in protein kinases that frequently carry a phosphorylatable residue (e.g. Thr160 in CDK2) [17,18***]. In the inactive CDK2 structure the enzyme exhibited a closed conformation in which residues of the activation segment (145-172) were folded so that they blocked the substrate binding site.

In 1995 the structure of the partially active CDK2-cyclin A complex was solved and demonstrated some of the key structural changes that lead to activation of the enzyme [5**]. In the CDK2-cyclin A complex the cyclin A molecule is bound to one side of the catalytic cleft and interacts with both lobes of CDK2 to form an extensive protein–protein interface (Fig. 2). The extensive contacts are consistent with biological data which suggest that the complex, once formed, is long lived [19]. The cyclin A molecule consists of two identical helical domains, and an independent structure determination of the unbound cyclin A has shown that the structure does not change between the free and complexed forms [20**]. Residues 45–51 of CDK2, which have the sequence PSTAIRE (using amino acid single-letter code), form an identifying motif of the cyclin-dependent kinases, and are known from mutational studies to be important for recognition of cyclin A. The structure shows that this sequence adopts a helical conformation which is central to the interface with cyclin A (Fig. 2). In the complex, the PSTAIRE helix shifts, with a rotation about its axis, and moves several angstroms into the catalytic-site cleft. The CDK2–cyclin A complex has a more open conformation in which the activation segment has been removed from the catalytic site. The conformation resembles that of the activation segment in the active ternary complex of cAPK, except for the three residues around the site of phosphorylation on Thr160. Thus, the structural changes in CDK2 on recognizing cyclin A result in the removal of the activation segment from the catalytic site (which would otherwise prevent access of the protein substrate) and the repositioning of the PSTAIRE helix, which leads to the correct alignment of ATP-recognizing groups. Thr160, the site of phosphorylation, is directed towards the solvent and is accessible to CAK, in contrast to its position in the free CDK2 (where it is buried).

Phosphorylation of Thr160 in the CDK2–cyclin A complex results in a 300-fold increase in activity. The latest
structural studies on the phosphorylated CDK2-cyclin A complex have shown the key changes that lead to full activation of the enzyme [6**]. The phosphorylated threonine, Thr160-P, swings in to occupy a buried site surrounded by the PSTAIRE helix, cyclin A, the strand containing the catalytic Asp127 and parts of the activation segment. As predicted [17,18**], the phosphate group contacts three arginines: Arg50, from the PSTAIRE helix; Arg126, adjacent to the catalytic aspartate; and Arg150, from the activation segment (Fig. 1c). The conformational changes involve the C-terminal lobe, the CDK2-cyclin A interface and the activation segment in the vicinity of Thr160. The activation segment shifts further away from the catalytic site and there are increased CDK2-cyclin A interactions (Fig. 2). These shifts further expose the likely substrate-recognition site for the consensus sequence Ser/Thr-Pro-X-Lys (where X stands for any amino acid). The activation segment now closely resembles the conformation of the segment in the active cAPK ternary complex. In further work, Russo et al. [21**] have determined the structure of the phosphorylated CDK2-cyclin A complex bound to the cellular inhibitor protein p27Kip1, which shows a fascinating mechanism for inhibition, but which is outside the scope of this review.

In the phosphorylated CDK2-cyclin A complex, the phosphate group plays a crucial role as an organizing centre within the structure. The extensive contacts with the arginine residues, and the other contacts these residues make extend the sphere of influence of phosphate beyond its immediate vicinity. The conformational changes are driven by shifts in a polypeptide region which is partially flexible in the nonphosphorylated form but which is ordered in the phosphorylated form.

Activation by phosphorylation results in a 300-fold increase in the activity of the CDK2-cyclin A complex. In the partially active CDK2-cyclin A complex, residue Glu162, occupies a position which is similar to that of the phosphorylated Thr160, and forms hydrogen bonds with Arg126 (Fig. 2). Evidently the single negative charge is not sufficient to drive the full activation process. In phosphorylase kinase, the kinase’s catalytic domain is constitutively active and requires no phosphorylation for activity. Structural results have shown that the residue corresponding to Thr197 in cAPK or Thr160 in CDK2 is a glutamate in phosphorylase kinase [22*]. The glutamate contacts an arginine that is adjacent to the catalytic base. In contrast to cAPK and CDK2, in which a phosphate group is required to neutralize two or even three arginines, in phosphorylase kinase there is only one arginine, and the other residues in the vicinity are neutral. In twitchin kinase [23] there is a further variation. The residue corresponding to the phosphorylatable residue is a valine, and the residue adjacent to the catalytic base is a leucine [23]. Satisfactory organization of the activation segment of twitchin kinase can apparently be achieved by nonpolar packing.

The extent to which protein kinases that are regulated by phosphorylation of a particular residue in the activation segment can be activated following mutagenesis of the residue to an acidic one varies. In mitogen-activated protein kinase, which is phosphorylated on two residues in the activation segment, Thr183 and Tyr185, the mutants Thr183Glu or Tyr185Glu do not show appreciable activity, although the mutant Thr183Glu can achieve 7% of native activity when it is phosphorylated on Tyr185 [24*]. Structural studies on the Tyr185Glu mutants have shown that the negatively charged carboxylate induces disorder in the activation segment, indicating the sensitivity of this region to negative charge, but it appears that only the double phosphorylation of the wild-type protein can induce the right conformation of the activation segment. In the inactive structure of the tyrosine kinase domain of the insulin receptor, the activation segment both blocks the catalytic site and exhibits disorder, and again, a significant structural change is anticipated upon phosphorylation that would allow the phosphate to contact arginine residues and induce the correct orientation of the activation segment [25].

**Yeast glycogen phosphorylase**

The structures of the phospho and dephospho forms of rabbit muscle glycogen phosphorylase (rmGP) are well established [8,26–28]. Within the last year the structures of the phospho and dephospho forms of *Saccharomyces cerevisiae* glycogen phosphorylase (yGP) have been solved [7*,29,30], and they reveal quite a different mechanism. In rmGP, the N-terminal tail changes its conformation upon phosphorylation of Ser14, from an extended conformation in which the peptide makes intrasubunit contacts and exhibits mobility, to an ordered conformation with partly 310 helix in which the Ser14-phosphate makes intersubunit contacts to two arginine residues, Arg69 from its own subunit and Arg43' from the other subunit of the dimer (Figs 1a, 3a). The phosphorylation events are accompanied by changes at the subunit-subunit interface that promote tertiary and quaternary structural changes leading to activation at the catalytic site, which is 45 Å away from the site of phosphorylation. In contrast, in yGP there is direct competition between the phosphoamino acid and an allosteric inhibitor, glucose-6-phosphate (Glc-6-P). On phosphorylation, the N-terminal tail, which is considerably longer in yGP than in rmGP, changes from a conformation in which it makes intersubunit contacts and blocks the catalytic site of the other subunit, to a conformation containing a helix and a hairpin-like loop, in which the phosphorylated threonine makes intrasubunit contacts after displacement of Glc-6-P (Fig. 3b).

In response to nervous or hormonal signals, RmGP is activated by phosphorylase kinase to mobilize the glycogen stores which produce energy to sustain muscle contraction. In yeast it appears that the phosphorylase is mobilized during the approach to stationary phase by transcriptional control in order to enable cells to
utilize glycogen under conditions of substrate depletion [31]. Phosphorylation of yGP results in over a 1000-fold increase in activity. Sequence alignment of yGP and rmGP showed that the former has a unique 39-residue N-terminal extension (numbered as -1 to -39 toward the N terminus relative to residue 1 of rmGP). From acid 80 onwards, the sequences exhibit 49% identity. In yGP, the site of phosphorylation is at Thr-10. There is no sequence identity in the peptides surrounding the sites of phosphorylation in rmGP and yGP, and there is no cross-reaction between the phosphorylases and their kinases, suggesting that these sites of control have originated through different evolutionary mechanisms.

In the nonphosphorylated form of yGP, which was crystallized in the presence of 50 mM Glc-6-P, the cores of the two domains are similar in structure to those in rmGP but there are differences in their relative orientations and the relative disposition of the two subunits of the dimer. The catalytic site of yGP is similar to the catalytic site of activated rmGP. The loop region from residue 280 to 286, which in inactive rmGP blocks access to
the catalytic site, is lifted away in yGP and the residues that form the phosphate substrate recognition site are in their correct position. The nonphosphorylated form of yGP appears to be inactive because the long N-terminal tail first wraps around its own subunit, mimicking closely the conformation of the N-terminal residues in active phosphorylated rmGP, but then continues to wrap around the other subunit so that the N-terminal residues actually contact residues at the entrance to the catalytic-site tunnel, thereby blocking it (Fig. 3b).

In the nonphosphorylated form of yGP, the inhibitor Glc-6-P \( (K_i=0.3\text{mM}) \) is bound at the subunit-subunit interface in the region equivalent to the nucleotide allosteric-effector site of rmGP [30]. The rmGP allosteric site binds the activator AMP with accompanying structural changes that tighten the interface in a similar way to that promoted by serine phosphorylation [27]. It also binds the allosteric inhibitor Glc-6-P [32]. In rmGP, both AMP and Glc-6-P utilize the same arginine residues, Arg309 and Arg310, to bind the phosphate moiety, but the other residues if the AMP- and Glc-6-P-binding sites are different. In yGP, Glc-6-P also utilizes these two arginines but the position of the glucose moiety is different from Glc-6-P bound to rmGP because of changes in sequence.

The position of Thr–10 in nonphosphorylated yGP is surrounded by nonpolar residues, and it was speculated that phosphorylation would result in a displacement that would also remove the N-terminal tail from the catalytic site [29]. The solution of the structure of the phosphorylated form of yGP has shown that these ideas were correct but there is an unexpected twist to the story [34]. It was found that a mutant in which the first 22 residues were deleted (NA22) was still inactive, indicating that activation could not be achieved simply by removing the blocking of the N-terminal tail of the catalytic site [33]. The structure of the phosphorylated form of this mutant showed that there are changes at the subunit interface, corresponding to a tightening around the regulatory site and an opening around the interface closer to the catalytic site, and changes in relative domain orientations upon phosphorylation. In the phosphorylated yGP, Thr–10 displaces the Glc-6-P and utilizes the two arginines, Arg309 and Arg310, as a phosphate binding site (Figs 1b, 3b). In producing this interaction the phosphate group triggers distant changes that result in key nonpolar packing of different parts of the polypeptide chain.

The major route for transmission of events is through the subunit contacts and these contacts can be partially mimicked by engineering a synthetic metal contact site [36]. The new work with yGP has shown that yGP has a further elaboration in that the displacement of the allosteric inhibitor Glc-6-P is crucial to the activation event. In CDK2, phosphorylation of the partially active CDK2–cyclin A structure results in the reorganization of the activation segment to promote a conformation accessible to substrate. The results from these three enzymes highlight the importance of phosphate–arginine interactions at high-affinity phosphate-recognition sites, and the ability of the phosphate group to act as an organizing centre.

We anticipate that further comparisons of phospho and dephospho enzymes will be available in the not too distant future. In particular, structural studies on several other kinases are well advanced. For these it is anticipated that the mechanisms are likely to mimic those observed for CDK2–cyclin A, but both mitogen-activated protein kinase and the insulin receptor tyrosine kinase have the additional elaboration of multiphosphate-site control. There will be a great interest in the understanding of these mechanisms. It may be anticipated that the four control mechanisms described by the presently known systems are unlikely to represent the complete repertoire of mechanisms of control by phosphorylation.

**Conclusions**

In rmGP, the primary driving force for the conformational change on phosphorylation appears to be electrostatic, and activation can be partially achieved by noncovalently bound phosphate or sulphate ions [34]. The interactions of the N-terminal peptide with the protein, however, are also important in transmitting the allosteric activation. Phosphorylase b', which lacks the first 19 residues of the N terminus, cannot be activated by sulphate [35].

**References and recommended reading**

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

The paper describes the structure of the CDK2-cyclin A complex with p27Kip1, a key regulator of cell proliferation.


The structure of a constitutively active protein kinase is presented.


The paper describes the mutants Thr183Glu and Tyr185Glu as the MAP kinase ERK2 is controlled by a flexible surface loop. Structure 1996, 3:299-307.

This review summarizes the recent structural data for eight protein kinases, respectively, regardless of the phase of the cell cycle.