Chemical approaches towards unravelling kinase-mediated signalling pathways

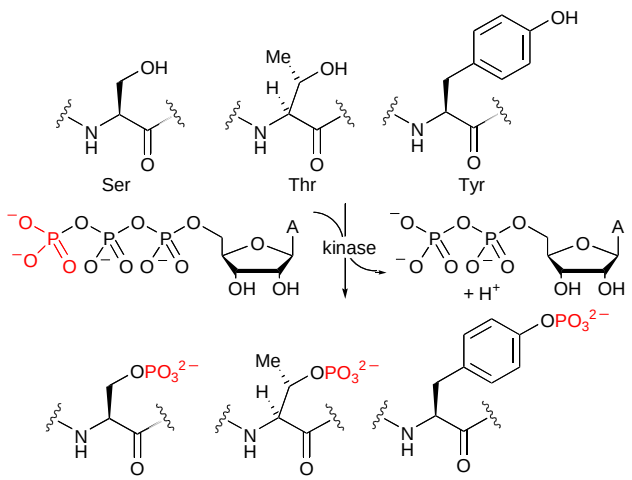
David R. W. Hodgson[a] and Martin Schröder[b]

Received (in XXX, XXX) Xth XXXXXXXXX 200X, Accepted Xth XXXXXXXXX 200X
First published on the web Xth XXXXXXXXX 200X
DOI: 10.1039/b000000x

Protein kinases control the function of about one third of cellular proteins by catalysing the transfer of the γ-phosphate group of ATP onto their substrate proteins. Protein phosphatases counter this action and also control the activation status of many kinases. Cellular responses to environmental changes, or signalling events, temporarily tilt the balance of protein phosphorylation and dephosphorylation to one side or the other. The identification of protein-kinase substrate pairs and substrate-phosphatase pairs is critical to understanding cell function and how cells respond to environmental changes. Identification of these substrate-enzyme pairs is non-trivial, because of the structural and mechanistic conservation of the catalytic cores of protein kinases. Here we review recent progress towards identifying protein kinase-substrate pairs by emphasising the use of chemical genetics and purpose-designed ATP analogues that target one particular protein kinase. In addition, we discuss activity-based chemical profiling approaches, based on ATP analogues, for the detection of active kinases.

A Introduction

Protein kinases catalyse the transfer of the γ-phosphate group of ATP onto serine, threonine or tyrosine residues of substrate proteins (Scheme 1). Phosphorylation can have profound effects on the phosphorylated protein, altering its enzymatic activity, ability to recognize and interact with other proteins, or subcellular localization. Phosphatases remove phosphoryl groups, making phosphorylation a reversible modification. Consequentially, the balance of protein kinase and phosphatase activities controls the phosphorylation state of many, if not all, cellular proteins. Temporary perturbation of this balance when cells respond to stimuli is the basis for protein kinase/phosphatase-mediated signal transduction.



Scheme 1 Kinase-catalysed transfer of the γ-phosphoryl group from ATP to substrate residues.

The identification of kinase-substrate interactions is central to understanding phosphorylation-based. This task is complicated by the huge number of protein kinases and phosphorylated substrates involved. The human 'kinome' consists of 518 protein kinases and an estimated ~20,000 phosphorylation events.[1] The catalytic cores of kinases adopt a common, conserved fold and mechanism to catalyse γ-phosphate transfer. This conservation complicates the identification of kinase-substrate pairs, because it makes it difficult to achieve selectivity for the γ-phosphate transfer reaction. In a cellular context, subcellular co-localization of a kinase and its substrate(s), and long-distance interactions between the kinase and its substrate(s) facilitate control over the selectivity of a particular protein kinase. Some approaches, notably the use of labelled ATP analogues, require cells to be lysed, because of the membrane impermeability of these hydrophilic molecules. Thus, the spatial control over kinase-substrate pairs is lost. For all these reasons faithful identification of all kinase-substrate pairs, especially those where the substrate is present at low abundance, is technically challenging. In this review we will summarise methods to identify these kinase-substrate pairs. Emphasis will be given to recently-developed chemical genetic approaches, and the structural and enzymological basis of these approaches. In addition, we will also discuss activity-based profiling of active protein kinases that employ chemical, ATP-based, probes for active protein kinases.

B Classical approaches towards identifying kinase-substrate pairs

The importance of kinase-substrate relationships has inspired many strategies towards identifying these pairs[1] that can be classified into biochemical and genetic approaches. To appreciate the advantages of chemical genetic approaches over these classical approaches, we will first briefly summarize these classical approaches and their limitations.

Biochemical approaches

Phosphorylation of potential substrates by a kinase is most often used to identify substrates for the kinase of interest. To identify substrates for a kinase, the kinase can be expressed in host cells such as bacteria, yeasts, or insect cells, and purified. The purified kinase can then be incubated with candidate substrate proteins, chosen on the basis of current knowledge or candidate libraries, and an ATP molecule labelled at the γ-phosphate position. The γ-phosphate label often takes the form of a radioactive phosphate or a thiophosphate group, where the choice of label determines the strategy for detection and isolation of the substrate(s). Candidate substrates can be provided as individual, recombinantly expressed and purified proteins. Substrates can also be supplied as libraries displayed on phages, arrayed as peptides or complete proteins on chips or microtiter plates, or even as whole cell lysates or lysates of subcellular fractions. Successful phosphotransfer can then be demonstrated by a variety of methods such as autoradiography after gel electrophoresis or precipitation of the substrate and quantitative analysis of incorporated radioactivity by scintillation counting, followed by identification of the protein by mass spectrometry (MS)-based methods.

Expression of protein kinases is not always successful, necessitating purification and isolation of some kinases from cells or tissues. This, however, comes at the price of potentially introducing contaminating kinases into the assay. This problem can be addressed if a given kinase has been expressed as a recombinant protein because quantitative comparison of γ-phosphate incorporation catalysed by wild type (WT) and a catalytically inactive mutant can be used to exclude phosphorylation events catalysed by contaminating kinases.

Careful optimisation of the experimental conditions for γ-phosphate transfer, such as using high concentrations of recombinant kinase, γ-^{32}P-ATP of high specific activity, short incubation times, Mn^{2+} instead of Mg^{2+} as cofactor, and chromatographic fractionation of the extract before the kinasing reaction, often allows successful identification of kinase substrates. Adoption of this careful optimisation of reaction conditions, together with parallel analysis of several isoforms of a kinase to identify substrates preferentially phosphorylated by only one of the isoforms is the basis of the KESTREL method.[2]

The major concern with *in vitro* identification of kinase-substrate pairs is identification of a large number of false-positive substrates, because the loss of tight control over the kinases that is present in cells *in vivo* often leads to promiscuous substrate phosphorylation.[2-4] Other limitations, such as non-specific adsorption of proteins to chips or nitrocellulose membranes and incorrect folding of recombinant proteins also need to be considered. Furthermore, for convenience, peptide analogues, rather than protein substrates, are often employed in *in vitro* studies, however, they often display K_M values that are several orders of magnitude higher than their protein progenitors.

To identify kinase substrates *in vivo* substrates phosphorylated by a given protein kinase in response to a stimulus can be identified by incubating cells with radioactive phosphate and comparing the spectrum of proteins that incorporated the radioactive phosphate in the presence and in the absence of the stimulus. This approach often suffers from poor selectivity and the use of heroic amounts of radioactive phosphate. The use of phosphorylation-specific antibodies or phospho-specific binding domains[5] allows enrichment of phosphopeptides and thus limits the need to use high doses of radioactivity. Antibodies with high affinity and specificity against phosphotyrosine are available, but preparation of antibodies of comparable quality against phosphoserine or phosphothreonine has been less successful.[3] All of these approaches require careful, case-specific optimisation and further validation, for example through *in vitro* kinase assays, to ensure that substrates for the kinase of interest are genuinely identified. This is because activation of downstream kinases and secondary effects can easily lead to the identification of false positives. Identification of kinase substrates by *in vivo* phosphorylation of substrates in cells treated with a protein kinase inhibitor often leads to false positive results because kinase inhibitors usually target conserved ATP-binding/catalytic sites in a promiscuous manner.[6, 7]

Kinase-substrate pairs can also be identified by exploiting the, often transient, interaction between the kinase and its substrate by any kind of methodology that allows studying of protein-protein interactions, such as pull-down assays, co-immunoprecipitation assays, cross-linking of kinase and substrate, yeast two-hybrid or split-ubiquitin assays. Interaction is not equivalent to phosphorylation, thus, substrates identified through interaction assays must be confirmed in phosphotransfer assays.

Instead of assaying for substrate phosphorylation, dephosphorylation of a phosphorylated substrate in the presence of high ADP concentrations can also be used to identify a kinase substrate.[8] This approach exploits the principle that kinases, as catalysts, do not alter the equilibrium position of the phosphotransfer reaction and that, in consequence, the reaction can be drawn to the side of the reactants by provision of a huge excess of ADP over ATP.

Genetic approaches

Genetic approaches can also be used to identify kinase-substrate pairs. These approaches are most advanced for lower eukaryotes such as the yeast *Saccharomyces cerevisiae*. However, these organisms do not harbour the full spectrum of orthologues of mammalian kinases or their substrates, thus, many potential kinases and substrates of biomedical significance are, necessarily, 'missed' when these model organisms are used.

In *S. cerevisiae*, synthetic genetic analysis (SGA) has been used to identify all pairwise genetic interactions between kinases, phosphatases, and their substrates[9]. SGA is based on robotic generation of mutant yeast cells that have two genes deleted, for example a kinase and its suspected (or known) substrate(s). Analysis then centres on quantification of a desired phenotype, usually growth on standard media. Comparison of phenotypes then allows the identification of aggravating, buffering, or neutral interactions between both mutations.

A common limitation of loss-of-function-based genetic approaches, such as SGA, is that gene pairs/families with redundant or overlapping functions escape investigation, plus it is limited to mutations that maintain viability of the cells. In addition, the mutant cells may adapt and mask the desired phenotype. Redundant gene pairs/families can be studied using overexpression approaches,[10] but the level of overexpression of a protein or a functional enzyme itself may introduce artefacts into the experimental system. In addition, genetic approaches do not provide information on whether an interaction is direct or indirect, in other words, mediated between the products of two interacting genes or by an unknown number of other proteins. The read-out of genetic interaction screens depends on the phenotype that is being investigated, meaning that many legitimate, but more subtle interactions that do not affect growth may not be uncovered in a screen for growth defects at normal growth temperatures.

Despite their limitations, classical biochemical and genetic approaches have made, and will continue to make, valuable contributions towards identifying kinase-substrate pairs. To obtain a complete picture of the kinome, the intrinsic shortcomings of classical approaches towards the identification of kinase-substrate pairs have to be addressed. Only then will reliable identification of all kinase-substrate pairs be possible, including the identification of those pairs where one or both partners are present at low abundance. Chemical genetics of protein kinases promises to fill this gap. Protein kinase chemical genetics combines the genetic engineering of a protein kinase and chemical synthesis of ATP analogues tailored as substrates or inhibitors towards only the engineered, or analogue-sensitised (*as*)-kinase, as it is known. This approach relies on our understanding of the mechanism of kinase-catalysed γ-phosphoryl group transfer from ATP onto substrates, binding of ATP to the active site, and synthetic strategies towards the preparation of ATP analogues. In the following sections we will briefly review the structure and catalytic mechanism of protein kinases, then discuss synthetic strategies towards ATP analogues and limitations of *as*-kinases, before concluding with examples of other chemical methods for understanding protein-kinase-mediated signalling pathways.

C Structure and catalytic mechanism of the protein kinase domain

The structures of the kinase domains of all eukaryotic protein kinases known to date are similar. This high degree of sequence conservation suggests that all eukaryotic protein kinases share a common structure and use similar mechanisms to catalyse γ-phosphate transfer.[11]

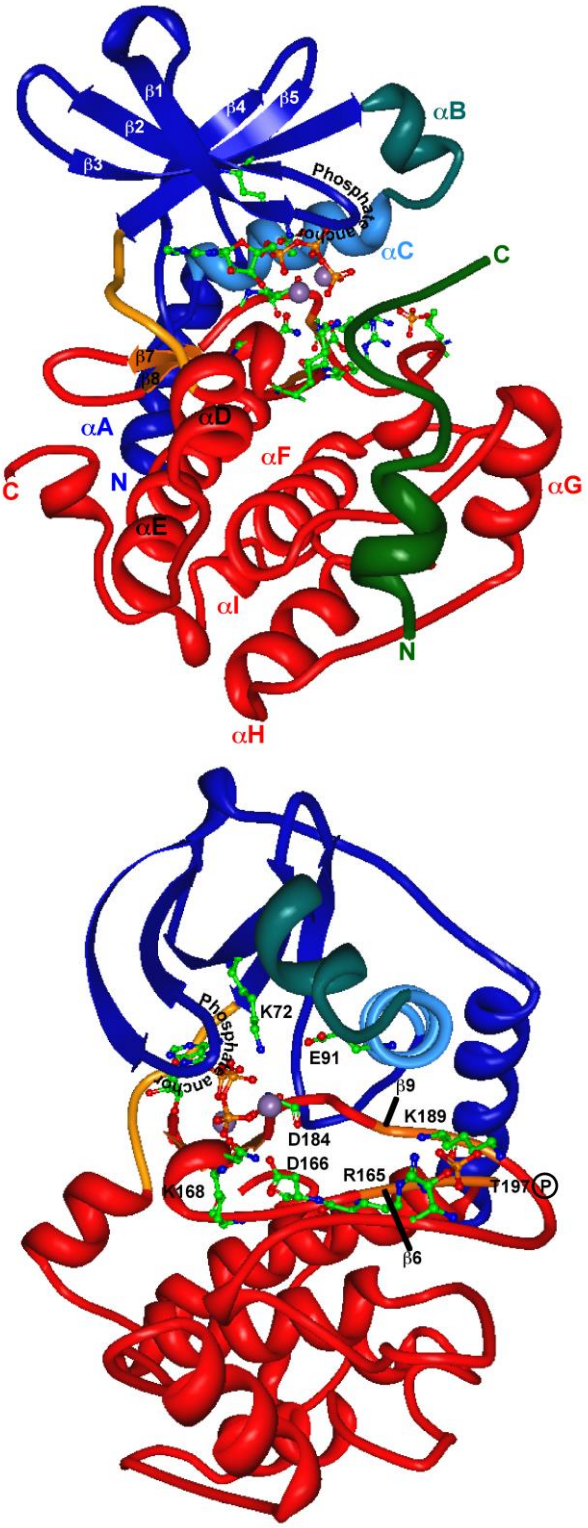


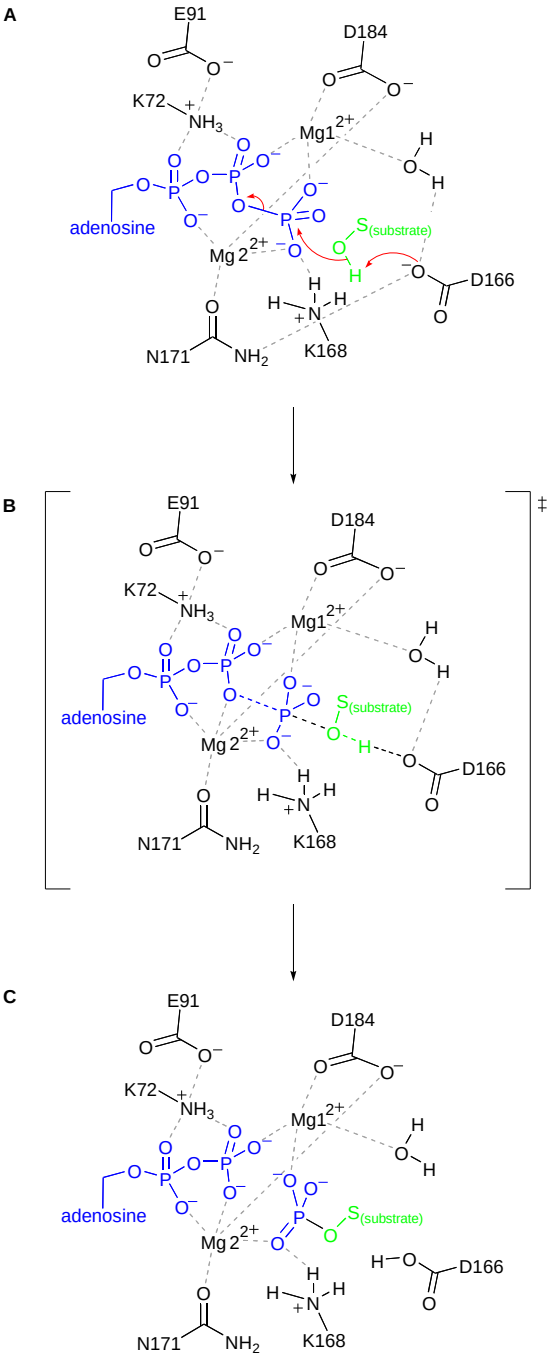
Fig. 1 Ribbon diagrams of the X ray structure of PKA (PDB no. 1ATP). The N lobe is blue, the C lobe red, and the linker between the N and C lobes orange. The protein kinase inhibitor peptide PKI (5-24) is shown in green. The non-conserved helix αB is in dark green, helix αC in light blue and β sheets in the C lobe in dark orange. In the upper picture, helices and β sheets are labelled. In the lower picture residues involved in catalysis of γ-phosphate transfer and phosphorylated T197 positioned in the basic RD pocket formed by R165 and K189 are labelled.

Structure

A typical protein kinase domain consists of two lobes, a small *N*-terminal lobe (the N lobe), primarily involved in ATP binding, and a large *C*-terminal lobe (C lobe) that is largely responsible for substrate binding and catalysis (Fig. 1), with an additional role in ATP binding. Comparison of the X-ray structures of protein kinase A (PKA), the cyclin-dependent kinase CDK2, and the mitogen-activated protein kinase ERK2, reveals a largely conserved fold for both lobes.[4] The N lobe consists of a twisted five-stranded antiparallel β-sheet and a lateral flanking αC helix. The phosphate anchor (GxGxxG, x = any amino acid) is located in a loop between sheets β1 and β2. Sheet β3 harbours an invariant lysine, K72, and the αC helix provides the invariant glutamate, E91. The 'phosphate anchor' positions the α- and β-phosphates of ATP. The C lobe consists of four β strands and several α helices. Its important role in catalysis is illustrated by the presence of the 'catalytic loop' between β6 and β7 of PKA (Fig. 1) and the Mg^{2+} loop between β8 and β9 in which an aspartate coordinating a Mg^{2+} ion, D184, is situated. The catalytic loop harbours K168, which contacts the γ-phosphate of ATP, and N171, which binds a second Mg^{2+} ion.

Another important feature of the kinase domain is the activation loop. The activation loop is located between the last β strand and helix αF. Certain protein kinases possess a conserved arginine immediately preceding the catalytic aspartate (D166), a basic residue in strand β9, and basic residues in helix αC. In these 'RD kinases' phosphorylation of serine, threonine, or tyrosine residues located in the activation loop induces positioning of the phosphorylated activation loop in to basic pockets.[12] Activation loop movement can also be induced through interaction of the protein kinase with regulatory subunits, for example interaction of cyclin-dependent kinases with cyclins.[12] The activation loop gates access to the ATP or substrate binding pockets. Consequently, chemical probing for accessibility of the ATP binding pocket can be used to identify the set of active kinases in a certain situation, for example in cancer cells[13] (cf. section E).

Several residues interact with ATP, however, those contributing to recognition of the adenine portion of ATP and the catalytic core for phosphoryl transfer are of most significance to chemical genetic studies and will be the primary focus here. Hydrogen bonds between the N7 nitrogen and the hydroxyl group of T183 and between the N6 nitrogen of the adenine ring and the backbone carbonyl of E121 bind the adenine ring in an otherwise hydrophobic pocket. L49, V123, and a large hydrophobic residue, M121 in PKA, contribute to this hydrophobic pocket. Filling-in the space created in kinase mutants in which the bulky, hydrophobic M121 residue has been replaced by a smaller amino acid using chemically synthesised adenine nucleotide analogues carrying a bulkier, hydrophobic adenine residue is the basis of protein kinase chemical genetics developed by Shokat and co-workers (Fig. 2).



Scheme 2 Mechanism of kinase-mediated phosphoryl transfer. **A** Protein kinase binding ATP and substrate in readiness for phosphoryl transfer. **B** Model of the phosphoryl transfer transition structure with a highly dissociative phosphoryl transfer and general base-assisted nucleophilic attack of an aliphatic (serine or threonine) or aromatic hydroxyl (tyrosine) group on the γ-phosphoryl group of ATP. **C** Products bound within the kinase active site. Models are based on[11, 14]

Catalytic roles of conserved residues

The catalytic mechanism for γ-phosphate transfer is highly dissociative (Scheme 2)[11], meaning that bond breaking between the leaving group *O*-atom and the γ-phosphorus atom is more advanced than the formation of the bond between the *O*-atom of the incoming nucleophile and the γ-phosphorus atom in the transition structure. The first Mg^{2+} ion ($Mg1^{2+}$,

Scheme 2A) is involved in optimal positioning of the γ-phosphate for attack by the incoming *O*-nucleophile or neutralization of the charges on the γ-phosphate to facilitate nucleophilic attack, or both. The second Mg^{2+} ion ($Mg2^{2+}$, Scheme 2A) is bound to N171, D184, a water molecule and the α- and γ-phosphates of ATP. It has been proposed to act as a Lewis acid stabilising the negative charge formed on the oxygen atom by bridging the β- and γ-phosphates in the transition structure (Scheme 2B).[14] D166 forms a hydrogen bond with a hydroxyl group of an inhibitor peptide in crystal structures of PKA, suggesting that D166 orients the hydroxyl group on the substrate for nucleophilic attack,[15] and subsequently acts as a general base that abstracts the proton from the hydroxyl group of the *O*-nucleophile late in the reaction[14]. The development of a large amount of negative charge on the *O*-nucleophile at an early point along the reaction co-ordinate may actually hinder attack on the γ-phosphate because of charge repulsion between the (di)anionic γ-phosphate and the incoming nucleophile.[16] The cationic K168 makes an electrostatic contact with the γ-phosphate in PKA throughout the reaction, but is replaced by an arginine or even alanine in other protein kinases[17]. D184 is conserved in all kinases where it coordinates both Mg^{2+} ions. A mutation of the analogous residue in the essential *S. cerevisiae* PKA isoform Tpk1p is incompatible with life, suggesting that D184 is critical for enzymatic activity[18]. The cationic K72 contacts the α- and β-phosphates and its mutation to arginine or alanine does not affect $K_{M,ATP}$, but decreases k_{cat} up to 840-fold, highlighting the importance of this residue in catalysis. K72 may also be required for formation of the active conformation by forming a salt bridge with E91 and sequestering E91 away from basic residues of the RD pocket.

D Chemical genetics of protein kinases

Chemical genetics of protein kinases has two aims:[19, 20] (*i*) to identify the physiological functions of protein kinases, and (*ii*) to identify all substrates for a given protein kinase. Common to both aims is the use of an engineered protein kinase, in which a large hydrophobic or polar residue in the adenine-binding pocket in the kinase domain (Fig. 2) has been mutated to a smaller amino acid, such as alanine or glycine, and adenine-enlarged ATP analogues. Cell permeable ATP analogues lacking the triphosphate group are used to identify physiological roles of the protein kinase in a temporally-controlled manner. Base-enlarged ATP analogues carrying transferable γ-phosphate groups are used to identify substrates for the protein kinase.

Mutation of a large hydrophobic or polar residue contacting the N^6 position of ATP generates access to an aryl pocket immediately behind the ATP binding pocket[21]. For this reason, this residue, which is conserved in most protein kinases as a large or polar amino acid[22], has been termed the 'gatekeeper' residue. Gatekeeper mutations decrease interactions between the adenine ring and the adenine-binding pocket of the protein kinase, and as a consequence decrease the affinity of the kinase for ATP ~10 fold[23]. However, the resulting K_M values remain well below the cellular ATP concentration of 1-3 mM. For example, the K_M value for ATP of engineered Rous sarcoma virus tyrosine kinase (*v*-Src) rises from 12 to 150 μM.[23] Therefore, mutation of the gatekeeper residue is not expected to affect kinase activity *in vivo*.

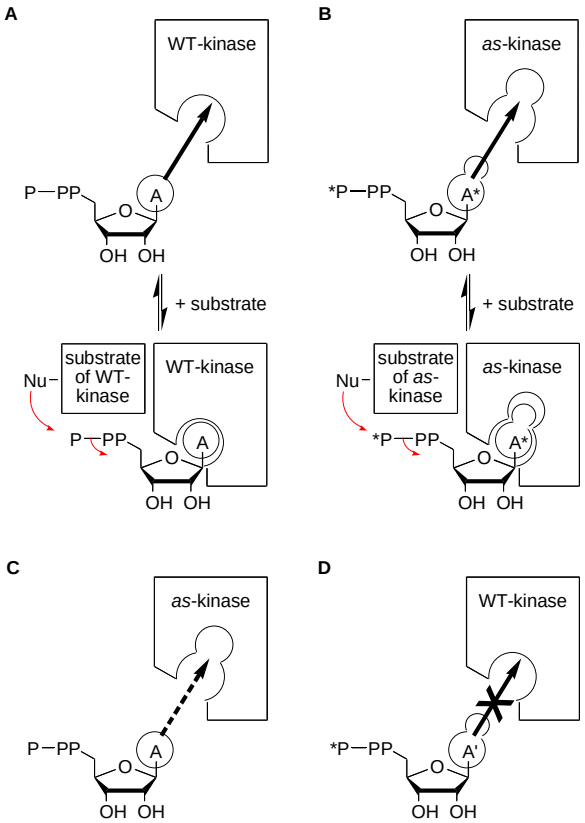


Fig. 2 Competitive binding of ATP and adenine-enlarged ATP to WT and *as*-kinases is the principle underlying chemical genetics of protein kinases developed by Shokat and co-workers. **A** Binding of ATP to WT kinases is unaffected allowing unlabelled phosphate to be transferred to substrates of WT kinases. **B** Adenine-enlarged A*TP binds effectively to *as*-kinase, for example M121A PKA, allowing labelled phosphate to be transferred to substrates of *as*-kinase alone. **C** Affinity of ATP for *as*-kinase is reduced compared to WT kinase, reducing competition between ATP and the adenine-enlarged A*TP for *as*-kinase active site. **D** Binding of adenine-enlarged A*TP to WT kinases is precluded on steric grounds.

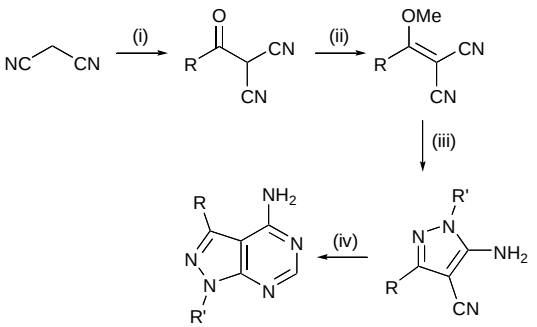
The enlarged adenine binding site can be filled by providing an ATP analogue with a bulkier, hydrophobic adenine moiety. These ATP analogues can engage in additional hydrophobic and van-der-Waals interactions with the enlarged adenine binding pocket and, thus, display an increased affinity for the *as*-kinase compared to ATP. At the same time, the bulkier heterocyclic ring systems preclude interaction of these ATP analogues with the ATP binding pockets of most other endogenous protein kinases or adenine nucleotide utilising enzymes. For example, a series of enlarged adenine analogues (analogous to those outlined in Scheme 3) was found to have IC_{50} values for WT *v*-Src or Fyn kinase of >300 μM[24], strongly suggesting that they had no effect on other endogenous kinases. Mutation of the gatekeeper residue decreased IC_{50} 30-50 fold meaning that the *as*-kinase was selectively inhibited.[24] In summary, the biochemical data characterising the interactions of WT and *as*-kinases with ATP and adenine-enlarged ATP analogues

(Fig. 2) form the basis of the idea that *as*-kinases are fully functional in cells; and that these kinases can either be selectively inhibited by providing base-enlarged ATP analogues that lack transferable phosphoryl groups, or can selectively transfer the γ-phosphate group of ATP analogues carrying triphosphate groups onto their substrates.

Synthesis of *as*-kinase inhibitors

Adenine-enlarged ATP analogues lacking the triphosphate group of ATP have been prepared, and they function as selective inhibitors of *as*-kinases (Scheme 3, R´=^{t}Bu; R= substituted phenyl, naphthyl or -CH_2-aryl). Owing to the absence of triphosphate groups these compounds are membrane-permeable and have been used to inhibit an *as*-kinase in intact yeast or mammalian cells, and have even been fed to mice to inhibit the *as*-kinase *in situ*.[25]

Many kinase inhibitors aim to mimic the adenine portion of ATP through the use of fused aza-heterocycles that show membrane permeability. There are many well-established routes for the preparation of aza-heterocycles and these can also be equally well applied to the preparation of *as*-kinase-targeted inhibitors. Most synthetic routes towards *as*-kinase inhibitors use nitrogen nucleophiles in combination with carbonyl-, cyano- or electron-deficient aromatic electrophiles, and a few illustrative synthetic examples are cited below. Hanefeld's method for the preparation of [3,4-*d*]pyrimidines, which have been used widely as *as*-kinase inhibitors, exemplifies the ease with which adenine analogues can be constructed (Scheme 3).[26]



Scheme 3 Hanefeld's method for the preparation of substituted adenine-analogues. (i) NaH, RCOCl; (ii) Me_2SO_4, reflux; (iii) $H_2NNHR´$, NEt_3, reflux; (iv) formamide, reflux.

Sandford's use of highly electron-deficient pyridine precursors illustrates the scope of sequential nucleophilic aromatic substitutions for the preparation of aza-heterocycles (Figure 3A),[27] using chemistry that is sufficiently robust to be used in simple flow systems.[28] Recently, a mixture of organolithium, carbonyl and nucleophilic aromatic substitution methods has allowed automated, multistep synthesis of a small library of potential kinase inhibitors, based on imidazo[1,2-*b*]pyridazines, using a microflow reactor (Figure 3B).[29] Taken together, these examples give a clear indication of the scope of simple, reliable chemical methods for the preparation of potential *as*-kinase inhibitors that can be used to selectively modulate the activities of chosen kinases. The limiting factor in the application of these potential inhibitors is the ability to screen libraries of compounds to determine their inhibition characteristics and selectivity towards a given *as*-kinase system.

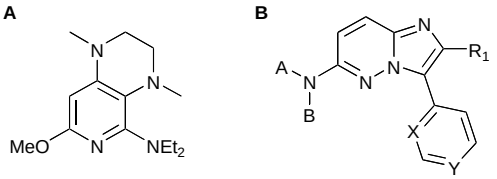
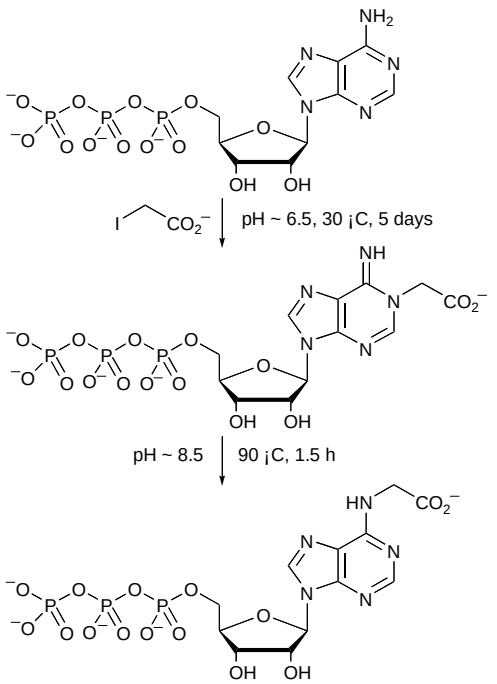


Fig. 3 A An example of an aza-heterocycle system assembled by sequential nucleophilic aromatic substitution steps. **B** Aza-heterocyclic scaffold for a range of potential kinase inhibitors assembled using flow chemistry.

Synthesis and use of ATP analogues with substrate properties in *as*-kinase systems

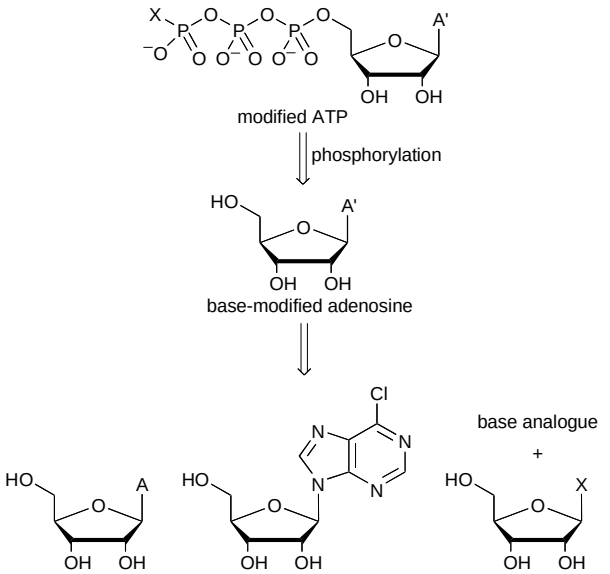
Base-enlarged analogues with a transferable γ-phosphoryl-group are used to identify unknown substrates of a given *as*-kinase by exploiting the enlarged active site's selectivity toward systems containing enlarged bases. In the simplest cases, it is possible to generate enlarged adenine nucleoside phosphates directly from phosphorylated adenosine systems through alkylation of the adenine and subsequent Dimroth rearrangement to give N^6-alkylated-5´-phosphorylated adenosines (Scheme 4).[30]



Scheme 4 Direct alkylation of the N^6-position of ATP.

However, the preparation of nucleotide triphosphates and their analogues is often a much more laborious task, even though clear, prescriptive procedures are available.[31]

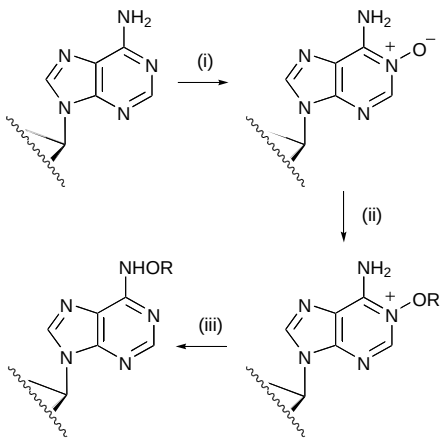
Usually, the assembly of ATP analogues can be broken down into two major disconnective steps; (*i*) preparation of the modified nucleoside, and (*ii*) elaboration to the triphosphate or its analogue (Scheme 5). These two conceptual steps will be considered in turn by way of a small selection of representative examples.



Scheme 5 Retrosynthetic analysis of modified ATP.

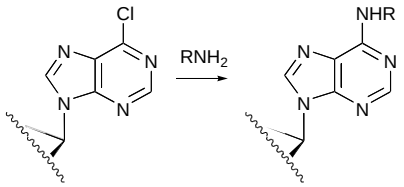
Preparation of base-modified adenosine analogues

Routes to adenine analogues can be loosely categorised into three classes. Conceptually, the simplest approach employs the use of alkylating agents for the modification of the exocyclic N^6-position of adenosine[32] or even ATP[30] (Scheme 4) via preliminary alkylation of N^1 of the adenine followed by Dimroth rearrangement. In addition, Dimroth chemistry can also be employed after preliminary oxidation of the N^1-position of adenosine. The resulting *N*-oxide can be *O*-alkylated, and the intermediate can be rearranged to form a N^6-hydroxylamine system (Scheme 6).



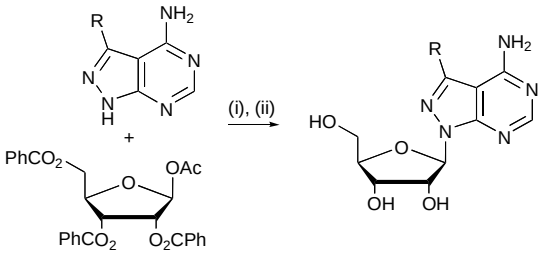
Scheme 6 Dimroth chemistry for the preparation of N^6-hydroxylamines. (i) oxidation to form N^1-oxide; (ii) *O*-alkylation of *N*-oxide; (iii) rearrangement to form N^6-*O*-alkyl hydroxylamine.

Secondly, nucleophilic aromatic substitution of a leaving group from the C^6-atom on an adenosine-based precursor, such as 6-chloropurine, by good nucleophiles is also a well-established approach towards modification at the 6 position of the purine ring (Scheme 7).[32]



Scheme 7 Displacement of 6-chloro-group by an amine nucleophile.

Finally, an adenine analogue can be assembled as a stand-alone entity before being added to the ribose ring via a glycosylation procedure. For example, Shokat and co-workers have employed Hanefeld's pyrazolopyrimidine synthesis to assemble a series of adenine analogues which have then been glycosylated using a protected D-ribose system to generate base-enlarged adenosine analogues. (Scheme 8).[33]



Scheme 8 Construction of base-enlarged adenosine analogues from pyrazolopyrimidines. R= -^{i}Pr, -Ph, -Bn, -CH_2-(1)Np. Reagents: (i) $SnCl_4$; (ii) NH_3 in MeOH.

Once assembled, the adenosine system must be phosphorylated.

Phosphorylation

Whilst many procedures are available for the preparation of triphosphate analogues,[31, 34] we will focus on two widely adopted approaches. In the first, a nucleoside diphosphate is assembled, then converted to the triphosphate analogue using enzymatic methods. In the second, a triphosphate is chemically ligated to a modifying group to generate the triphosphate analogue.

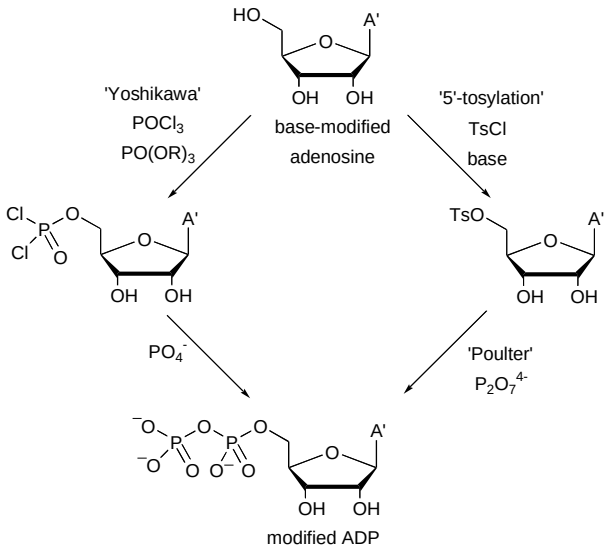
Diphosphates can be accessed by several routes, however, two commonly used routes will be discussed in detail here: (*i*) Yoshikawa phosphorylation of the modified adenosine followed by introduction of the second phosphoryl group through the use of tri- or tetraalkylammonium salts of inorganic phosphate, and (*ii*) preparation of a 5′-tosylate of the adenosine analogue and subsequent displacement of this leaving group using a tri- or tetraalkylammonium salt of the inorganic pyrophosphate ion (Scheme 9).

When employing the Yoshikawa route, the adenosine substrate must be dissolved in dry trialkyl phosphate solvent. Phosphorus oxychloride is then introduced in excess, often with a tertiary amine base that neutralises the acid that is evolved.[31] When following the Poulter route,[35] protection of the 2′/3′-*cis*-diol system of the adenosine substrate is usually required in order to ensure exclusive 5′-tosylation, which adds protection and deprotection steps to this route. At this stage, both routes require sources of mono- or di-phosphate nucleophiles, which must be soluble in organic solvents. In order to overcome the intrinsic insolubility of metal salts of phosphates in organic solvents, ion exchange procedures are

used to prepare tri- or tetraalkylammonium salts that show improved solubility in organic solvents. Preparation of the alkylammonium salts is a laborious task, however, some systems are commercially available. The mono- or diphosphate-sources must be scrupulously dry before addition to the nucleoside-monophosphodichloridate intermediate or the 5´-tosylate, and this drying is achieved through time-consuming repeated evaporations with dry pyridine.

Crude reaction mixtures from both procedures usually require extensive chromatographic purification. Crude mixtures from the Poulter route are usually less complex, however, attempts towards the preparation of diphosphates via the Yoshikawa route often result in complex mixtures of mono-, di- and triphosphates that present a significant chromatographic challenge.[36]

Other approaches towards the assembly of di- and triphosphates use phosphoramidates, such as morpholidates,[37] or carbonyl diimidazole (CDI)[38] as agents for the formation of phospho-anhydrides. There is usually a requirement for scrupulously dry conditions that must be maintained for long reaction times. Furthermore, tri- or tetraalkylammonium salts of phosphates and careful, time-consuming chromatography steps are also needed, and, even then, these efforts are often poorly rewarded with moderate yields.[39]



Scheme 9 Preparation of nucleoside diphosphates.

Enzymatic labelling of the triphosphate group

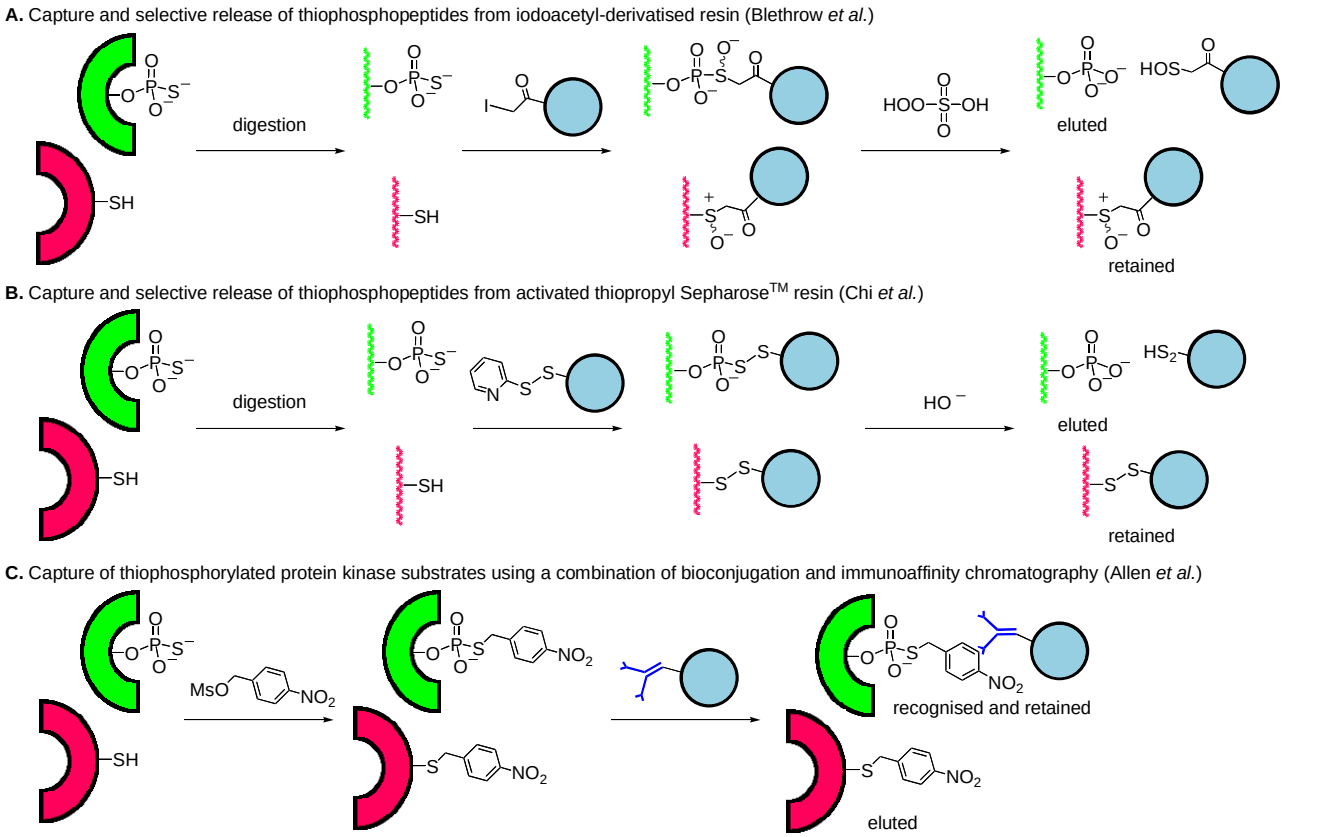
Base-enlarged ATP analogues labelled with a γ-^{32}P-phosphoryl group or a γ-thiophosphoryl group are conveniently prepared by nucleoside diphosphate kinase (NDPK)-mediated transfer of a labelled phosphoryl group from γ-^{32}P-ATP or γ-S-ATP. NDPK operates via a ping-pong mechanism where the labelled phosphoryl group is first transferred from the donor ATP substrate to N1 of an active site histidine giving a phosphoryl-enzyme intermediate. After dissociation of the ADP by-product from the phosphorylated active site, the base-enlarged ADP substrate binds and the labelled phosphoryl group is then transferred to give the labelled, base-enlarged ATP analogue.[33] Immobilisation of NDPK on a solid support can be used to ensure that excess labelled ATP donor and ADP by-product can be fully removed before introduction of the base-enlarged ATP analogue. It is essential to avoid the presence of ATP donor because this could be used by non-*as*-kinases and lead to false positives.

ATP analogues labelled with a γ-^{32}P-phosphoryl group and γ-thiophosphoryl group have been used by Shokat and co-workers. In the case of *v*-Src the catalytic efficiency (k_{cat}/K_M) of an *as*-kinase with γ-^{32}P-N^6-cyclopentyl-ATP as substrate was indistinguishable from the catalytic efficiency of the WT kinase with ATP.[23, 40] Subsequently, γ-^{32}P-based triphosphate analogues derived from Hanefeld's pyrazolpyrimidines (Scheme 8) have also been used.[33] γ-^{32}P-based systems have several disadvantages. Firstly, detection and isolation of labelled proteins requires the use of 2D gel methods. Isoelectric focussing conditions in the first dimension have to be carefully chosen to minimize bias against highly acidic or basic proteins. Buffers to solubilize transmembrane proteins have to be carefully optimised to maximise solubility of these proteins in gel-based analysis systems. Secondly, ^{32}P-phosphorylated substrates are susceptible to hydrolysis by phosphatases, which decreases sensitivity and may introduce bias.

The use of chemically labelled γ-phosphates promises to circumvent some of these technical limitations. The most obvious choice to label the γ-phosphate is replacement of one terminal oxygen atom by a sulfur atom in γ-S-ATP. The nucleophilic sulfur anion of the thiophosphate, when transferred to a substrate protein, offers a 'handle' for affinity purification, bioconjugation or, indeed, a combination of these methods. Furthermore, thiophosphates are more resistant to hydrolysis by phosphatases, thus, once formed, thiophosphopeptides are more likely to be detected than phosphopeptides. In addition, γ-S-ATP is not readily accepted as a substrate by ATPases and adenine-enlarged γ-S-ATP analogues can be generated with relative ease using NDPK as described above.

Simple reversible thiol affinity chromatography of systems that have been exposed to base-enlarged-γ-S-ATP analogues, where captured thiophosphorylated peptides are released via disulfide reduction, is limited by the fact that cysteine-containing proteins are also likely to be captured. Judicious control of pH can be used to attenuate the reactivity of thiols in comparison to thiophosphates,[41] however, selective release of thiophosphorylated peptides from a capture resin represents a more reliable strategy. This approach was adopted in parallel by Blethrow *et al.* and *Chi et al.*, who used different thiol-affinity resins and release strategies in order to accomplish this goal (Scheme 10). Blethrow *et al.* chose to capture both thiophosphorylated peptides and cysteine-containing peptides after digestion of a pool of potential substrates.[42] An iodoacetyl resin was employed, which led to the formation of carbon-sulfur bonds. In the case of the thiophosphate system, the resulting *S*-bridging phosphodiester was susceptible to hydrolysis after oxidation of the sulfur atom. On the other hand, the thioethers resulting from capture of cysteine-containing peptides were oxidised to stable sulfoxides that remained attached to the affinity resin (Scheme

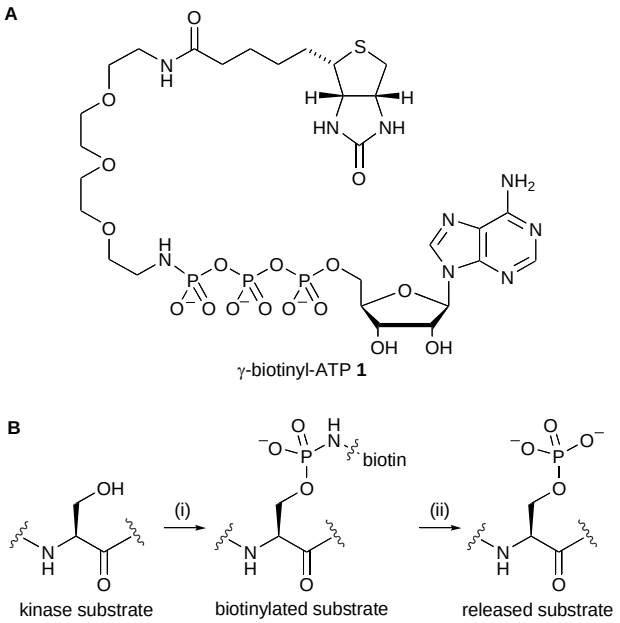
10A). Chi *et al.* captured both thiophosphorylated peptides and cysteine-containing peptides after digestion, however, a thiopyridyl disulfide resin was employed in order to form mixed disulfides between the resin and thiophosphopeptides or cysteine-containing peptides.[43] The mixed disulfides between simple thiols (on the resin) and thiophosphoryl groups are labile under basic conditions, however, simple mixed disulfides between resin and cysteine groups are not, thus, thiophosphopeptides were selectively eluted from the affinity resin (Scheme 10B). In a third approach, Allen *et al.* chose to alkylate the nucleophilic thiophosphate ion and then used antibodies raised against an *O*-thiophosphorylated, *S*-alkylated thiophospho-threonine-based hapten to immunoprecipitate thiophosphorylated, alkylated kinase substrates (Scheme 10C).[44]



Scheme 10 Methods for selective purification of thiophosphorylated proteins and peptides. Green semi-circles represent thiophosphoproteins, green lines digested phosphoprotein fragments, red semi-circles non-thiophosphorylated proteins contained cysteine residues, red lines digested protein fragments derived from these proteins and cyan circles represent affinity matrices. Blue lines represent antibodies raised against S-(*p*-nitrobenyl)thiophosphopeptides.

Chemical labelling of the triphosphate group

Green and Pflum extended the use of γ-modified ATP analogues by using a γ-biotinyl-ATP **1** (Scheme 11A).[45] This approach relied upon the open nature of the active sites of many kinases around the position of the γ-phosphoryl group of ATP (Fig. 1). When γ-biotinyl-ATP **1** was incubated with protein kinase A or casein kinase 2 and their respective substrates, the biotinylated γ-phosphoryl group was transferred to the substrates in both cases (Scheme 11B). Once attached, the biotinyl group was used as an affinity tag for purification of labelled substrates. The tag was then cleaved from the substrates using acid to allow for identification of the substrate by MS. The use of a phosphoramidate-linked label also gave conclusive evidence that significantly elaborated γ-phosphoryl groups can be transferred by kinases. This result is initially surprising given the differenccs in mechanisms of phosphoryl group transfer between phosphate mono- and diesters. The fact that phosphoramidate-based γ-groups can be transferred can perhaps be explained by virtue of the fact that the γ-amino group is likely to be protonated at the pHs that are used in labelling assays. This, in turn, probably leads to an increase in the electrophilicty of the adjacent phosphoryl group and a more dissociative mechanism for phosphoryl tranfer being retained,[46] with the net result being only minor reductions in observed k_{cat} values (see dansyl systems below).



Scheme 11 A γ-biotinyl-ATP **1**. **B** Green and Pflum's use of γ-biotinyl-ATP **1** in kinase catalysed biotinyl-phosphate transfer for kinase substrate isolation. (i) γ-biotinyl-ATP **1** and kinase, (ii) affinity chromatography then CF_3CO_2H.

In further extensions of this approach, Pflum and co-workers have developed ATP analogues with fluorophore- and photo-probe-modified γ-phosphoryl groups in the form of dansyl- and azido-phosphoramidates, respectively.[47, 48] The dansyl-ATP system was obtained commercially, whereas, the azide was prepared via carbodiimide-based coupling of ATP to an amino-based linker connected to an aryl azide. With the dansyl system, efficient dansylation of both peptide and protein substrates was demonstrated where only relatively small reductions in k_{cat} values (~6-25-fold) were observed. With the azide system, a given kinase can transfer the azide-modified phosphoryl group to a substrate, then, when photolysed, the azide gives a reactive intermediate that can cross-link to the kinase. In this way, a covalent link can be established between a given kinase and its substrate(s) to form a high molecular weight complex between the kinase and substrate which migrates as a single entity during gel electrophoresis. The substrate and kinase components can then be identified using classical blotting techniques.[47]

Gouverneur and co-workers[49] have also developed phosphoramidate-based azide- and alkyne-γ-modified ATP analogues that can be used for 'click'-style cycloadditions and Staudinger ligations. These functional groups provide handles for further modification, such as addition of an affinity purification tag, when the γ-phosphoryl has been transferred to a peptide substrate. Synthetic methods were based on CDI activation of ATP followed by reaction of the activated ATP with a functionalised amine. In addition, phosphoester-based, rather than phosphoramidate-based, systems were also explored, and these were prepared by sequential phosphorylation of a functionalised alcohol followed by ligation of the resulting phosphate to ADP using CDI. Despite the expected difference in reactivity of the phosphoester-based substrates, phosphorylation of $p27^{kip1}$ by CDK2 could be shown using an antibody against phosphorylated $p27^{kip1}$.[49] Considering the tight steric requirements of the phosphospecific antibody during selective recognition of phosphorylated epitopes, these results are surprising and may require confirmation using more sensitive techniques, such as MS.

Limitations of *as*-kinases

as-Kinases are powerful tools for identifying kinase substrates and the functions of protein kinases. The methodology is, however, limited by (*i*) mutation of the gatekeeper inactivating some protein kinases; (*ii*) a few kinases possessing alanine or glycine as the gatekeeper amino acid,[50] (*iii*) provision of the *as*-kinase at endogenous expression levels, and as the sole source for the kinase; (*iv*) competition between ATP and adenine-enlarged ATP analogues in cell lysates as phosphate donors, (*v*) the membrane impermeability of charged ATP analogues (*vi*) and the rapid hydrolysis and reincorporation of their labelled phosphates into nucleotides.

Considerations (*i*), (*ii*) and (*iii*) affect the use of ATP analogues that lack transferable phosphoryl groups when trying to identify physiological functions of the *as*-kinase. Limitation (*i*) is significant because loss of control over the time point of inactivation of the *as*-kinase by administration of the inhibitor largely eliminates the advantage that the use of *as*-kinases promises compared to classical approaches based on kinase-inactivating amino acid mutations. Limitation (*ii*) indicates that interpretation of data arising from inhibition of *as*-kinases may not be foolproof.

Identification of so-called 'second-site suppressors' has gone some way towards addressing problems with *as*-kinases that do not tolerate mutation of the gatekeeper residue. Mutation of the gatekeeper residue in G protein-coupled receptor kinase 2 (GRK2), cell division cycle 5 (cdc5) kinase, or MEK kinase 1 destroyed kinase activity.[40, 50] Nevertheless, the activity of these mutant kinases was restored by 'second site suppressor' mutations. These second-site suppressors can be identified using random PCR mutagenesis screens ideally coupled to a phenotypical read-out in *E. coli*,[50] yeast, or another genetically traceable organism. Alternatively, candidates for second-site suppressors may be rationally identified by comparing residues lining the ATP-binding pocket (Figure 1).[40]

Identification of *as*-kinase substrates is affected by all six potential limitations. Provision of the *as*-kinase in excess over its normal expression level may lead to the identification of low affinity, false positive substrates. Depending on the organism, expression of the *as*-kinase at endogenous expression levels to minimize identification of false positive, low affinity substrates and to retain spatial, temporal, or developmental control over activity of the *as*-kinase may also be challenging. For inhibition studies, the *as*-kinase also needs to be the sole source of the chosen kinase activity. Homologous recombination to replace the WT allele by the *as*-allele can be used in organisms displaying high frequencies of homologous recombination, for example certain yeast species.[51] The relatively simple gene and promoter structures in these organisms greatly facilitate near endogenous

expression of *as*-alleles on single copy plasmids in cells harbouring a chromosomal deletion of the gene encoding the kinase of interest[51]. This latter approach is not available for mammalian cells, therefore *as*-kinases have been mostly used in yeast.[22, 52-57] In mammalian cells, *as*-kinases have been used to study kinases whose overexpression confers phenotypes.[40, 58] For example, inhibition of the transforming *as*-kinase v-erbB-*as*1 with 1NA-PP1 (Scheme 3, R=1-naphthyl, R´= ^{t}Bu) inhibited the transforming properties of this kinase.[58]

Homologous recombination in murine embryonic stem cells can be used to generate mice carrying the *as*-allele in the chromosomal position of the WT allele. Cell lines, usually embryonic fibroblast lines, can be established from these mice. For differentiated cell types this approach often may not yield sufficient material for biochemical searches for *as*-kinase substrates. In addition, generation of these mice is expensive, time-consuming, and laborious. Zinc finger nucleases (ZFNs), which have been designed to bind to and introduce a DNA strand break near a specific DNA sequence, yield frequencies of homologous recombination of 1-20% in yeast, plant, and mammalian cells.[59, 60] Co-transfection of a vector expressing a ZFN and a donor vector providing the cDNA for the *as*-kinase can be used to integrate the *as*-allele into upstream exons of the kinase gene. Alternatively, the gatekeeper mutation can be introduced directly into the chromosomal locus.

ATP analogues containing phosphate groups cannot cross the cell membrane because of their charged nature, therefore, substrates are often identified in cell lysates. Permeabilisation of cell membranes with low concentrations of the detergent digitonin is a well-established technique and allows ATP analogues to enter the cell.[61] Permeabilisation conditions can be chosen in such a way as to deplete intracellular ATP concentrations, which is important to ensure that relatively small, possibly micromolar concentrations of the ATP analogue, can successfully compete for the *as*-kinase against endogenous millimolar concentrations of ATP. Permeabilisation also offers advantages over cell lysis, because cells often exert spatial control over kinases by separating kinase and substrate in separate compartments. Of course, upon lysis this spatial control is lost.

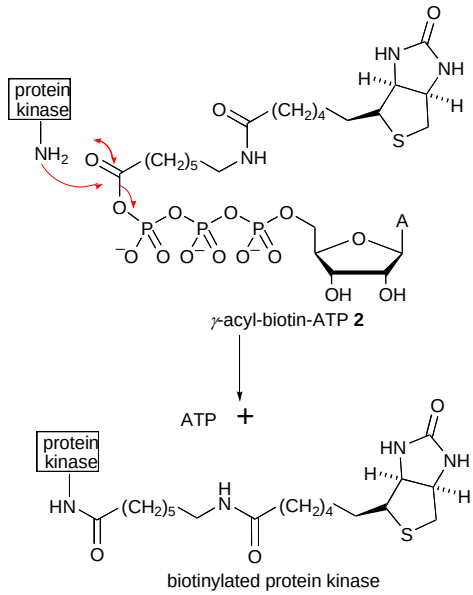
However, significant problems remain with using ATP analogues in cell lysates or permeabilized cells. Cellular ATP, if not efficiently removed, may efficiently compete with precious base-enlarged ATP analogues as phosphate donor, because of its high concentration in cells. Conceptually, ATP can be excluded from the nucleotide binding pocket of an *as*-kinase by introducing bulkier amino acids contacting other portions of the adenine system. These secondary mutations are then compensated for by excavating the adenine system. Progress towards these design aims has been made with the development of N^4-benzyl-5-aminoimidazole-4-carboxamide ribotide triphosphate, which is accepted by an *as*-kinase, but not by the corresponding WT kinase.[62, 63]

Hydrolytic release of the γ-phosphate from ATP analogues such as N^6-phenylethyl-γ-^{32}P-ATP, catalysed by apyrases or ATPases, occurs in <1 min.[61] The liberated phosphate can be reincorporated into the cellular ATP pool in <0.1 s,[64-66], for example by NDPKs, resulting in label dispersion, and a substantial non-specific background. γ-S-ATP is resistant to hydrolysis by ATPases and some apyrases,[67] but is a substrate of NDPKs. Not all kinases readily accept γ-S-ATP as substrate.[44, 48] Thus, further research is required to optimise the performance of ATP analogues to screen for *as*-kinase substrates.

E Activity-based profiling of active protein kinases

In addition to identifying the physiological roles and substrates of all protein kinases, knowledge about which kinases are active in a given experimental setting, for example cancer cells or after exposure of cells to a stimulus, is also valuable. This latter knowledge can be obtained by using activity-based profiling of kinases[68] by exploiting regulation of the activity of many protein kinases by controlling access of ATP to the ATP binding pocket through movement of the activation loop. To identify active kinases cell lysates or permeabilised cells are incubated with an ATP analogue, such as such as γ-acyl-biotin-ATP **2** (Scheme 12).[69] Synthesis of γ-acyl-biotin-ATP **2**, was accomplished via activation of *N*-(+)-biotinyl-6-aminohexanoic acid *via* the chloroformate method followed by ligation of the activated biotin system to the organic solvent-soluble triethylammonium salt of ATP under anhydrous conditions in dipolar aprotic solvent to form a mixed phosphoric-carboxylic acid anhydride.

After labelling, affinity techniques were used to isolate biotin-labelled proteins or peptides for identification and quantification through MS-based techniques. This approach was also used to determine the selectivity of protein kinase inhibitors towards a given protein kinase through competition studies[13]. γ-Acyl-biotin-ATP labels one of two active sites lysines corresponding to K72 and K168 in PKA. K72 is conserved in ~98% of the protein kinases, whereas K168 is less well conserved.[70] Both ATP- and ADP-derived biotinylating agents were used to label >80% of the predicted kinases in cancer cells, indicating that the majority of protein kinases are active in this experimental setting. Rules predicting whether the K72 or K168 analogous residues are preferentially labelled by γ-acyl-biotin-ATP or γ-acyl-biotin-ADP have not yet been developed.



Scheme 12 Kozarich's kinase active site-directed biotinylation strategy for the isolation of kinases. The active site lysine that is labelled corresponds to either K72 or K168 in PKA.

Conclusions

Examples of genetic and classical approaches towards identification of kinase-substrate pairs were used to illustrate the power and limitations of these approaches. In order to address some of these limitations, chemical genetic approaches have been adopted, and these are currently gaining prominence in the area of kinomics. New developments, particularly with respect to addressing the problems associated with the preparation of nucleotides, should overcome some of the obstacles that limit the application of these approaches at present. The development of new, less disruptive approaches towards delivery of charged nucleotide systems through cell walls will remove a significant limitation to the chemical genetic approach where currently cell permeability limits applications to lysates or other *in vitro* systems where spatial control over the action of kinases is lost. However, extension to mammalian systems where biomedical targets can be assessed more robustly will truly allow this field of chemical genetics to achieve its full potential.

Notes and references

[a] *Department of Chemistry, Durham University, Science Laboratories, Durham DH1 3LE, United Kingdom. Fax: 0191 384 4737; Tel: 0191 334 2123; E-mail: d.r.w.hodgson@durham.ac.uk*

[b] *School of Biological and Biomedical Sciences, Durham University, Science Laboratories, Durham DH1 3LE, United Kingdom. Fax: 0191 334 9104; Tel: 0191 334 1316; E-mail: martin.schroeder@durham.ac.uk*

This work was supported by grants from the Biotechnology and Biological Sciences Research Council to D.R.W.H. and M.S. [BB/F003757/1] and M.S. [BB/D01588X/1].

1. S. A. Johnson and T. Hunter, *Nat. Methods.*, 2005, **2**, 17-25.
2. P. Cohen and A. Knebel, *Biochem. J.*, 2006, **393**, 1-6.
3. D. C. Berwick and J. M. Tavare, *Trends Biochem Sci*, 2004, **29**, 227-232.
4. A. Knebel, N. Morrice and P. Cohen, *Embo J*, 2001, **20**, 4360-4369.
5. M. B. Yaffe and A. E. Elia, *Curr. Opin. Cell Biol.*, 2001, **13**, 131-138.
6. J. Bain, H. McLauchlan, M. Elliott and P. Cohen, *Biochem. J.*, 2003, **371**, 199-204.
7. S. P. Davies, H. Reddy, M. Caivano and P. Cohen, *Biochem. J.*, 2000, **351**, 95-105.
8. H. K. Kole, M. Abdel-Ghany and E. Racker, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 5849-5853.
9. D. Fiedler, H. Braberg, M. Mehta, G. Chechik, G. Cagney, P. Mukherjee, A. C. Silva, M. Shales, S. R. Collins, S. van Wageningen, P. Kemmeren, F. C. P. Holstege, J. S. Weissman, M. C. Keogh, D. Koller, K. M. Shokat and N. J. Krogan, *Cell*, 2009, **136**, 952-963.
10. R. Sopko, D. Huang, N. Preston, G. Chua, B. Papp, K. Kafadar, M. Snyder, S. G. Oliver, M. Cyert, T. R. Hughes, C. Boone and B. Andrews, *Mol. Cell.*, 2006, **21**, 319-330.
11. J. A. Adams, *Chem. Rev.*, 2001, **101**, 2271-2290.
12. B. Nolen, S. Taylor and G. Ghosh, *Mol. Cell.*, 2004, **15**, 661-675.
13. M. P. Patricelli, A. K. Szardenings, M. Liyanage, T. K. Nomanbhoy, M. Wu, H. Weissig, A. Aban, D. Chun, S. Tanner and J. W. Kozarich, *Biochemistry*, 2007, **46**, 350-358.
14. M. Valiev, J. Yang, J. A. Adams, S. S. Taylor and J. H. Weare, *J. Phys. Chem. B*, 2007, **111**, 13455-13464.
15. J. Zhou and J. A. Adams, *Biochemistry*, 1997, **36**, 2977-2984.
16. K. Kim and P. A. Cole, *J. Am. Chem. Soc.*, 1998, **120**, 6851-6858.
17. S. K. Hanks and T. Hunter, *FASEB J.*, 1995, **9**, 576-596.
18. C. S. Gibbs and M. J. Zoller, *J. Biol. Chem.*, 1991, **266**, 8923-8931.
19. L. M. Elphick, S. E. Lee, V. Gouverneur and D. J. Mann, *ACS Chem. Biol.*, 2007, **2**, 299-314.
20. M. A. Shogren-Knaak, P. J. Alaimo and K. M. Shokat, *Annual Review of Cell and Developmental Biology*, 2001, **17**, 405-433.
21. Y. Liu, K. Shah, F. Yang, L. Witucki and K. M. Shokat, *Bioorg Med Chem*, 1998, **6**, 1219-1226.
22. A. C. Bishop, J. A. Ubersax, D. T. Petsch, D. P. Matheos, N. S. Gray, J. Blethrow, E. Shimizu, J. Z. Tsien, P. G. Schultz, M. D. Rose, J. L. Wood, D. O. Morgan and K. M. Shokat, *Nature*, 2000, **407**, 395-401.
23. K. Shah, Y. Liu, C. Deirmengian and K. M. Shokat, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 3565-3570.
24. A. C. Bishop, K. Shah, Y. Liu, L. Witucki, C. Kung and K. M. Shokat, *Curr. Biol.*, 1998, **8**, 257-266.
25. H. Wang, E. Shimizu, Y. P. Tang, M. Cho, M. Kyin, W. Zuo, D. A. Robinson, P. J. Alaimo, C. Zhang, H. Morimoto, M. Zhuo, R. Feng, K. M. Shokat and J. Z. Tsien, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 4287-4292.
26. U. Hanefeld, C. W. Rees, A. J. P. White and D. J. Williams, *J. Chem. Soc., Perkin Trans. 1*, 1996, 1545-1552.
27. G. Sandford, R. Slater, D. S. Yufit, J. A. K. Howard and A. Vong, *J. Org. Chem.*, 2005, **70**, 7208-7216.

28. R. D. Chambers, G. Sandford and J. Trmčić, *J Fluorine Chem*, 2007, **128**, 1439-1443.

29. F. Venturoni, N. Nikbin, S. V. Ley and I. R. Baxendale, *Org. Biomol. Chem.*, 2010, **8**, 1798-1806.

30. M. Lindberg and K. Mosbach, *Eur. J. Biochem.*, 1975, **53**, 481-486.

31. D. M. Williams and V. H. Harris, in *Organophosphorus reagents : a practical approach in chemistry*, ed. P. J. Murphy, OUP, Oxford, 2004.

32. M. J. Robins and E. M. Trip, *Biochemistry*, 1973, **12**, 2179-2187.

33. B. C. Kraybill, L. L. Elkin, J. D. Blethrow, D. O. Morgan and K. M. Shokat, *J. Am. Chem. Soc.*, 2002, **124**, 12118-12128.

34. K. Burgess and D. Cook, *Chem. Rev.*, 2000, **100**, 2047-2059.

35. V. J. Davisson, D. R. Davis, V. M. Dixit and C. D. Poulter, *J. Org. Chem.*, 1987, **52**, 1794-1801.

36. N. C. Mishra and A. D. Broom, *J Chem Soc Chem Comm*, 1991, 1276-1277.

37. S. Roseman, J. J. Distler, J. G. Moffatt and H. G. Khorana, *J. Am. Chem. Soc.*, 1961, **83**, 659-663.

38. J. W. Kozarich, A. C. Chinault and S. M. Hecht, *Biochemistry*, 1973, **12**, 4458-4463.

39. G. K. Wagner, T. Pesnot and R. A. Field, *Nat. Prod. Rep.*, 2009, **26**, 1172-1194.

40. D. M. Kenski, C. Zhang, M. von Zastrow and K. M. Shokat, *J. Biol. Chem.*, 2005, **280**, 35051-35061.

41. S. W. Kwon, S. C. Kim, J. Jaunbergs, J. R. Falck and Y. M. Zhao, *Mol. Cell. Proteom.*, 2003, **2**, 242-247.

42. J. D. Blethrow, J. S. Glavy, D. O. Morgan and K. M. Shokat, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 1442-1447.

43. Y. Chi, M. Welcker, A. A. Hizli, J. J. Posakony, R. Aebersold and B. E. Clurman, *Genome Biol.*, 2008, **9**, R149.

44. J. J. Allen, S. E. Lazerwith and K. M. Shokat, *J. Am. Chem. Soc.*, 2005, **127**, 5288-5289.

45. K. D. Green and M. K. H. Pflum, *J. Am. Chem. Soc.*, 2007, **129**, 10-11.

46. T. Lonnberg, M. Ora and H. Lonnberg, *Mini-Rev. Org. Chem.*, 2010, **7**, 33-43.

47. S. Suwal and M. K. H. Pflum, *Angew. Chem. Int. Ed.*, 2010, **49**, 1627-1630.

48. K. D. Green and M. K. H. Pflum, *ChemBioChem*, 2009, **10**, 234-237.

49. S. E. Lee, L. M. Elphick, A. A. Anderson, L. Bonnac, E. S. Child, D. J. Mann and V. Gouverneur, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 3804-3807.

50. C. Zhang, D. M. Kenski, J. L. Paulson, A. Bonshtien, G. Sessa, J. V. Cross, D. J. Templeton and K. M. Shokat, *Nat. Methods*, 2005, **2**, 435-441.

51. M. F. Tuite, *Crit. Rev. Biotechnol.*, 1992, **12**, 157-188.

52. F. R. Papa, C. Zhang, K. Shokat and P. Walter, *Science*, 2003, **302**, 1533-1537.

53. K. R. Benjamin, C. Zhang, K. M. Shokat and I. Herskowitz, *Genes Dev.*, 2003, **17**, 1524-1539.

54. H. Abeliovich, C. Zhang, W. A. Dunn, Jr., K. M. Shokat and D. J. Klionsky, *Mol. Biol. Cell*, 2003, **14**, 477-490.

55. A. S. Carroll, A. C. Bishop, J. L. DeRisi, K. M. Shokat and E. K. O'Shea, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 12578-12583.

56. M. H. Jones, B. J. Huneycutt, C. G. Pearson, C. Zhang, G. Morgan, K. Shokat, K. Bloom and M. Winey, *Curr. Biol.*, 2005, **15**, 160-165.

57. E. L. Weiss, A. C. Bishop, K. M. Shokat and D. G. Drubin, *Nat. Cell Biol.*, 2000, **2**, 677-685.

58. Q. W. Fan, C. Zhang, K. M. Shokat and W. A. Weiss, *Curr. Biol.*, 2002, **12**, 1386-1394.

59. F. D. Urnov, J. C. Miller, Y.-L. Lee, C. M. Beausejour, J. M. Rock, S. Augustus, A. C. Jamieson, M. H. Porteus, P. D. Gregory and M. C. Holmes, *Nature*, 2005, **435**, 646-651.

60. Y. Doyon, J. M. McCammon, J. C. Miller, F. Faraji, C. Ngo, G. E. Katibah, R. Amora, T. D. Hocking, L. Zhang, E. J. Rebar, P. D. Gregory, F. D. Urnov and S. L. Amacher, *Nat. Biotechnol.*, 2008, **26**, 702-708.

61. A. Chaudhary, J. S. Brugge and J. A. Cooper, *Biochem. Biophys. Res. Commun.*, 2002, **294**, 293-300.

62. S. M. Ulrich, O. Buzko, K. Shah and K. M. Shokat, *Tetrahedron*, 2000, **56**, 9495-9502.

63. S. M. Ulrich, N. A. Sallee and K. M. Shokat, *Biorg. Med. Chem. Lett.*, 2002, **12**, 3223-3227.

64. D. J. Smith, B. O. Stokes and P. D. Boyer, *J. Biol. Chem.*, 1976, **251**, 4165-4171.

65. W. S. Thayer and P. C. Hinkle, *J. Biol. Chem.*, 1975, **250**, 5330-5335.

66. W. S. Thayer and P. C. Hinkle, *J. Biol. Chem.*, 1975, **250**, 5336-5342.

67. M. A. Komoszyński, *Acta. Biochim. Pol.*, 1991, **38**, 157-163.

68. B. F. Cravatt, A. T. Wright and J. W. Kozarich, *Ann. Rev. Biochem.*, 2008, **77**, 383-414.

69. M. P. Patricelli, A. K. Szardenings, M. Liyanage, T. K. Nomanbhoy, M. Wu, H. Weissig, A. Aban, D. Chun, S. Tanner and J. W. Kozarich, *Biochemistry*, 2007, **46**, 350-358.

70. N. Kannan, S. S. Taylor, Y. Zhai, J. C. Venter and G. Manning, *PLoS Biol.*, 2007, **5**, e17.