Cell. Mol. Life Sci. 63 (2006) 2608–2625
1420-682X/06/222608-18
DOI 10.1007/s00018-006-6202-8


Cellular and Molecular Life Sciences

# Review

# Structural biology of protein tyrosine kinases

**S. W. Cowan-Jacob**

Novartis Institutes for Biomedical Research, Lichtstrasse 35, WSJ-88.9.08a, Novartis Pharma AG, 4056 Basel (Switzerland), Fax: +41 61 324 2686, e-mail: sandra.jacob@novartis.com

Received 2 May 2006; received after revision 21 June 2006; accepted 9 August 2006
Online First 19 October 2006

**Abstract.** Our current understanding of the structure, mechanism of action and modes of regulation of the protein tyrosine kinase family owes a great deal to structural biology. Structures are now available for more than 20 different tyrosine kinase domains, many of these in multiple conformational states. They form the basis for the design of experiments to further investigate the role of different structural elements in the normal function and regulation of the protein and in the pathogenesis of many human diseases. Once thought to be too similar to be specifically inhibited by a small molecule, structural differences between kinases allow the design of compounds which inhibit only an acceptable few. This review gives a general overview of protein tyrosine kinase structural biology, including a discussion of the strengths and limitations of the investigative methods involved.

**Keywords.** Kinase structure, structural biology, protein crystallography, protein nuclear magnetic resonance spectroscopy, conformational flexibility, regulatory domain, drug discovery.

## Introduction

Protein kinases control many different cellular functions by phosphorylating proteins involved in gene expression, metabolic pathways, cell growth and differentiation, membrane transport and apoptosis. Kinase activity can be modulated by phosphatases, upstream kinases, cofactors and ligands and by interactions with other activating or regulatory proteins. With such a critical role in signal transduction, the tight regulation of kinase activity is crucial. A number of diseases and disease symptoms result from the disregulation of kinases or from activating and inactivating mutations in kinases [1].

There are 518 kinase sequences encoded in the human genome (1.7%), of which 430 are expected to be catalytically active [2, 3]. Based on sequence homology, the largest subgroup of the human protein kinases is the tyrosine kinase family, with 90 members [2]. This family can be divided into the receptor tyrosine kinases (RTKs) and the non-receptor tyrosine kinases (nRTKs) (Table 1). The 58 RTKs, which can be further divided into 16 different classes, consist of an extracellular ligand-binding domain linked via a transmembrane section to an intracellular tyrosine-kinase-containing region. Their activation usually involves the binding of a ligand to the extracellular region, which causes dimerization and allows autophosphorylation of the intracellular tyrosine kinase domains. Non-RTKs (32 proteins in nine subgroups) also respond to extracellular stimuli by means of modular units like Src homology 2 (SH2), Src homology 3 (SH3) and Pleckstrin homology (PH) domains, or modification by lipids (e.g. myristate), for appropriate subcellular localization.

Over the past 10–15 years, structural biology has contributed a great deal to our understanding of kinase function. In the early 1990s, X-ray crystallography showed that the catalytic domains of Ser/Thr and Tyr kinases share a common architecture [4, 5]. The kinase domain, which is typically about 300 residues in length, consists of an N-terminal lobe containing a five-stranded $\beta$ sheet and a single $\alpha$ helix ($\alpha$C), plus a larger, mostly helical C-terminal lobe (Fig. 1a). A hinge region, a stretch several residues in length connecting the N- and C-terminal lobes, allows a certain amount of flexibility in their relative ori-

**Table 1.** Publicly available structural data for protein tyrosine kinases (as of February, 2006).

| Family | Sub-family | Protein | PDB entry for kinase domain | PDB entry for non-kinase domain(s) |
|---|---|---|---|---|
| nRTK | Abl | ABL | 1FPU, 1IEP, 1M52, 1OPJ, 1OPK, 1OPL, 2F4J, 2FO0 | SH2: 1AB2, 1OPK, 1OPL, 2ABL, 2FO0<br>SH3: 1ABO, 1ABQ, 1BBZ, 1OPK, 1OPL, 2ABL, 2FO0, 1AWO, 1JU5 |
| | Csk | CSK | 1BYG, 1K9A | SH3: 1K9A, 1CSK, 1JEG<br>SH2: 1K9A |
| | FAK | FAK | 1MP8 | FAT: 1K04, 1K05, 1OW6, 1OW7, 1OW8, 1KTM, 1PV3, 1QVX<br>FERM: 2AEH, 2AL6 |
| | JAK | JAK2 | 2B7A | |
| | | JAK3 | 1YVJ | |
| | Src | FYN | | SH3: 1EFN, 1AVZ, 1FYN, 1G83, 1M27, 1SHF, 1AZG, 1NYF, 1NYG, 1ZBJ, A0N<br>SH2: 1G83, 1AOT, 1AOU |
| | | HCK | 1AD5, 2HCK, 1QCF, 2C0I, 2C0T | SH3: 1BU1, 4HCK, 5HCK, 1AD5, 2HCK, 1QCF<br>SH2: 3HCK, 1AD5. 2HCK, 1QCF |
| | | LCK | 3LCK, 1QPC, 1QPD, 1QPE, 1QPJ | SH3: 1LCK, 1X27, 1H92, 1KIK<br>SH2: 1CWD, 1CWE, 1FBZ, 1IJR, 1LCJ, 1LCK, 1LKK, 1LKL, 1X27, 1BHH, 1BHF |
| | | LYN | | SH3: 1WA7, 1W1F |
| | | SRC | 1FMK, 2PTK, 2SRC, 1KSW, 1Y57, 1YOJ, 1YOL, 1YOM | SH3: 1FMK, 2PTK, 2SRC, 1KSW, 1Y57, 1NLO, 1NLP, 1PRL, 1PRM, 1QWE, 1QWF, 1RLP, 1RLQ, 1SRL, 1SRM<br>SH2: 1FMK, 2PTK, 2SRC, 1KSW, 1Y57, 1A07, 1A08, 1A09, 1A1A, 1A1B, 1A1C, 1A1E, 1F1W, 1F2F, 1O41-9, 1O4A-R, 1P13, 1SHD, 1HCS, 1HCT |
| | Syk | SYK | 1XBA, 1XBB, 1XBC | SH2: 1A81, 1CSY, 1CSZ |
| | | ZAP70 | 1U59 | SH2: 1M61 |
| | Tec | BTK | 1K2P | SH3: 1AWW, 1AWX, 1QLY<br>PH: 1B55, 1BTK, 1BWN |
| | | ITK | 1SM2, 1SNU, 1SNX | SH2: 1LUI, 1LUK, 1LUM, 1LUN, 2ETZ, 2EU0<br>SH3: 1AWJ |
| RTK | EGFR | EGFR | 1M14, 1M17, 1XKK | ECD: 1IVO, 1MOX, 1YY9<br>JM: 1Z9I |
| | | ERBB2 | | ECD: 2A91, 1S78 |
| | | ERBB4 | | ECD: 2AHX |
| | Eph | EPHA1 | | FN3: 1X5A |
| | | EPHA2 | 1MQB | |
| | | EPHA4 | | SAM: 1BOX |
| | | EPHA8 | | FN3: 1X5L |
| | | EPHB2 | 1JPA | SAM: 1SGG, 1B4F, 1F0M<br>LBD: 1NUK, 1SHW, 1KGY |
| | FGFR | FGFR1 | 1AGW, 1FGI, 1FGK, 2FGI | ECD: 1EVT, 1FQ9, 1CVS, 2CR3<br>JM: 1XR0 |
| | | FGFR2 | 1GJO, 1OEC | ECD: 1EV2, 1DJS, 2FDB, 1II4, 1IIL, 1NUN, 1E0O |
| | | FGFR3 | | ECD: 1RY7 |
| | InsR | IGF1R | 1JQH, 1K3A, 1M7N, 1P4O | ECD: 1IGR |
| | | INSR | 1IRK, 1IR3, 1GAG, 1I44, 1P14, 1RQQ, 2AUH, 2B4S | |
| | Met | MET | 1R0P, 1R1W | PSI: 1SSL<br>SEMA: 1SHY |
| | Musk | MUSK | 1LUF | |
| | PDGFR/ | FLT3 | 1RJB | |
| | VEGFR | KIT | 1PKG, 1T45, 1T46 | |
| | | FLT1 | | ECD: 1FLT, 1QTY, 1QSV, 1QSZ, 1RV6 |
| | | KDR | 1VR2, 1Y6A, 1Y6B, 1YWN | |
| | Tie | TIE2 | 1FVR | |
| | Trk | TRKA | | LBD: 1WWA, 1HE7, 1WWW |
| | | TRKB | | LBD: 1WWB, 1HCF |
| | | TRKC | | LBD: 1WWC |

Some structures of pharmaceutically interesting PTKs are not yet available in the PDB, but are filed in patent claims (e.g. Pyk2 WO 2004078923, EphA7 WO 2004081180, ErbB4 WO 2004066921, Ret WO 2005062795).

SH2, Src homology 2 domain; SH3, Src homology 3 domain; PH, Pleckstrin homology; JM, juxtamembrane region; LBD, ligand-binding domain; ECD, extracellular domain; FERM, 4.1-ezrin-radixin-moesin domain; FAT, focal adhesion targeting sequence; SAM, sterile alpha motif; FN3, fibronectin type III domain; SEMA, semaphorin homology domain; PSI, plexin-semaphorin-integrin domain.

entations, which is necessary for the catalytic function of the enzyme. A cleft between these two lobes accommodates ATP, the adenosine group of which forms hydrogen bonds with the hinge region. In tyrosine kinases, the substrate binds to a platform on the C-terminal lobe, with the hydroxyl group of the tyrosine to be phosphorylated pointing toward the $\gamma$-phosphate of ATP. This platform is partially formed by residues from the activation loop (A-loop), the conformation of which is stabilized by the phosphorylation of one or more tyrosine residues in most tyrosine kinases.

Residues important for catalysis or for the correct architecture of the catalytic site come from different parts of the structure known as the P-loop (glycine-rich or nucleotide-binding loop), $\beta$ strand 3, and helix C, all within the N-terminal lobe, and the catalytic loop ($\beta$ strands 6 and 7) and the A-loop (about 20 residues located between $\beta$ strand 8 and helix $\alpha$EF), both in the C-terminal lobe (Fig. 1). Many of these structural elements correspond to regions with highly conserved sequences (Fig. 1b). For example, the P-loop has a consensus sequence of GxGx$\Phi$G (where x is any amino acid and $\Phi$ is Phe or Tyr), and the catalytic loop contains strictly conserved Asp and Asn residues. The activation segment starts and ends with anchoring regions, which have a conserved DFG sequence at the N-terminal anchor (often called the Mg-binding loop) and an APE, ALE or SPE sequence at the C-terminal anchor [6]. The A-loop, which has a variable sequence, lies between these two anchor regions along with the P+1 loop at the C-terminal end whose sequence contributes to substrate specificity.

The reaction catalyzed by tyrosine kinases involves the transfer of the $\gamma$-phosphate of ATP to the hydroxyl of the substrate tyrosine and requires the presence of a divalent metal ion. Transfer is generally considered to proceed by direct attack of the hydroxyl oxygen of the substrate on the $\gamma$-phosphate, probably via a dissociative mechanism [7]. There is some debate about the exact role of neighboring protein residues: the aspartate from the DFG motif coordinates the essential metal ion, which may be necessary for the correct positioning of the terminal phosphate for transfer, or it may be necessary to mask the charge of the $\gamma$-phosphate in the reaction transition state, or even to limit the electrostatic repulsion for an incoming nucleophile. The role of the conserved aspartate in the catalytic loop may be to direct the hydroxyl of the substrate for attack on the $\gamma$-phosphate of ATP, but there have been discussions about whether it could serve as a base catalyst, or to repel the phosphoproduct and facilitate dissociation [7]. Structural studies provide the basis for the design and understanding of ongoing experiments to understand these and other aspects of kinase structure and function.

## Structural biology

To date, all protein kinase structures at atomic resolution have been determined by protein crystallography (Table 1), but the kinase domain is not beyond the reach of full-structure determination by nuclear magnetic resonance (NMR), and we will probably see NMR structures

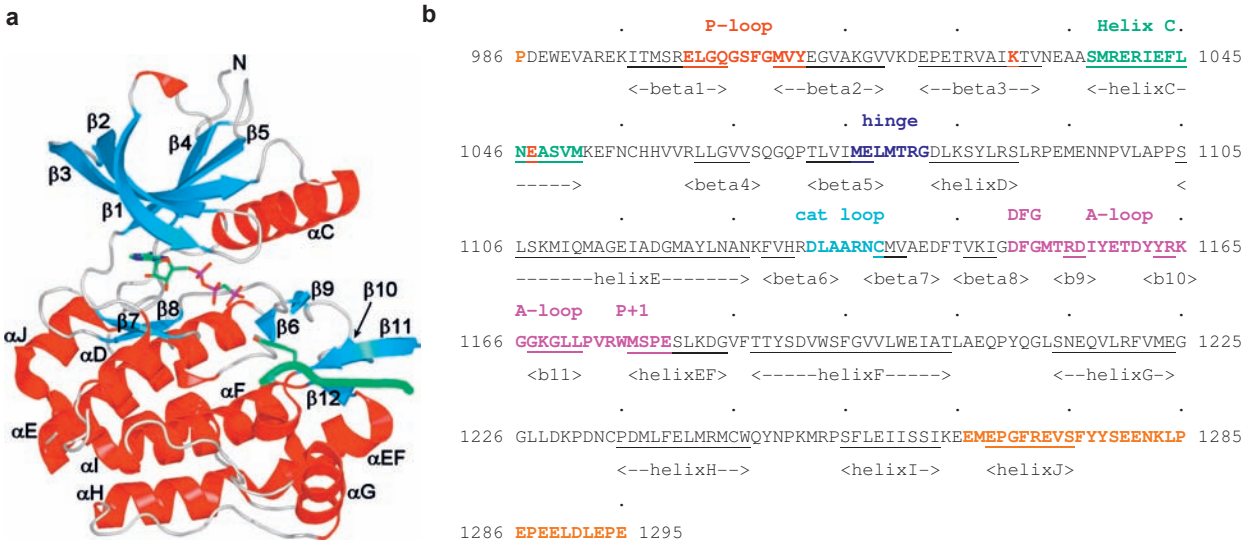


**Figure 1.** (*a*) Structure of active Igf1r kinase with bound ATP analog (AMP-PCP) and substrate (PDB entry 1K3A). Beta sheets are shown in blue and alpha helices in red. The substrate is green and the ATP analog has green carbons. There are no magnesium ions in this model [51]. This image, and those in Figures 2–6 have been prepared using the program Pymol (http://www.pymol.org/). (*b*) Sequence of the receptor tyrosine kinase (Igf1r), showing the structural elements mentioned in the text color-coded as follows: red, P-loop; green, helix C; blue, hinge region; cyan, catalytic loop; magenta, activation segment; orange, JM region and C-terminal tail. The lysine from $\beta$3 and the glutamate from helix C, which form a salt bridge in the active kinases, are also colored in red.

of kinases in the near future [8]. Both methods require the preparation of milligram amounts of highly pure and stable protein. For NMR structure determination, the homogeneity is not as critical as for crystallization, but complete isotopic labeling of the amino acids with $^{15}N$ and $^{13}C$ is required.

## Protein preparation

The first step in a protein kinase structure determination project is to try to define which part of the multidomain protein to prepare. The choice of the N and C termini of the construct for expression of the protein is often critical. Variations will have an effect on the expression levels and crystallizability due to effects on solubility, stability and activity [see ref. 9]. When there is high sequence homology with proteins of known structure defining the ends of the sequence that will form the globular domain is relatively straightforward. For less well defined domains such as kinases with lower sequence homology, or those with regions that might fold back and bind to the kinase domain, it is useful to express a longer or full-length version of the protein and subject it to limited proteolysis to help define the globular region [e.g. ref. 10]. In general it is important that there are no flexible regions that could interfere with crystallization, and for this reason it may even be necessary to remove long flexible loops from within the domain of interest [see refs. 11–14]. Hydrogen/deuterium exchange mass spectrometry (DXMS) has recently been shown to provide valuable information about flexible regions that can be used to help in construct design [15].

Problems in crystallization can arise when there is a mixture of similar, but not identical entities. Due to the fact that small differences can cause faults or even stop crystal growth, it can be more of a problem having highly similar impurities than completely different proteins in a crystallization mixture [16, 17]. A mixture of phosphorylation states can be tolerated [see e.g., ref. 18], but success in crystallization is exceptional. Phosphorylation frequently causes conformational changes in kinases, and a mixture of conformational states is not conducive to crystallization. During purification it is often possible to separate different phosphorylation states using ion-exchange chromatography, but the result is usually a much lower yield of protein. Site-directed mutants are often considered during construct design to avoid this type of problem. The preparation of constitutively activated protein by replacing a tyrosine with an aspartate, or uniformly unphosphorylated protein by replacing tyrosine with phenylalanine to avoid a mixture of phosphorylation states, are two tricks that have proven useful for the preparation of highly homogeneous kinases suitable for crystallization [19].

Until now, for the purposes of structural biology, most tyrosine kinases have been expressed in baculovirus-infected insect cells. However, there are several exceptions, and the use of *Escherichia coli* is becoming more common [9, 20–23]. Coexpression with chaperones has been found to help produce soluble, active kinases [e.g. ref. 24], and the presence of phosphatases can also help, by reducing the amount of phosphorylation. This may be important if the active kinases are toxic to the cells, if the phosphorylated states are metastable, or if a mixture of phosphorylation states reduces potential yield [23, 25]. In many cases, the expression levels of the kinase are low, requiring the use of affinity tags to help enrich the sample during purification. A histidine tag is commonly used and is normally placed at the N terminus, followed by a specific cleavage site such as that for tobacco etch virus protease [e.g. ref. 26]. Cleavage of the tag is often necessary to avoid adding flexible regions that can interfere with the formation of crystals, but the histidine tag is short enough to be left on in some cases [9]. This also makes it a useful tag to use at the C terminus, where it is usually left uncleaved because the addition of a protease cleavage site would add as many extraneous amino acids as the tag itself.

In many cases, the isolated kinase domain can be unstable and may tend to precipitate out of solution during purification and concentration of the sample in preparation for crystallization. Addition of a ligand such as an inhibitor can help to stabilize the protein, making it possible to carry it through the purification process and reach the mg/ml concentrations required for crystallization [27, 28]. The same is true for expression of the kinase: in cases where the expression levels are low, addition of an inhibitor at micromolar concentrations to the expression media can influence both the phosphorylation state of the kinase and the expression levels. For example, in the case of Abl kinase expressed in baculovirus-infected insect cells, the addition of an inhibitor allowed the production of mostly unphosphorylated protein rather than a mixture of phosphorylation states, which resulted in a much greater yield of homogeneous protein, even though the expression levels were similar [27]. Other examples have shown a greatly increased level of kinase expression in the presence of an inhibitor [13, 29].

## Crystallography

Crystallization requires a highly homogeneous sample of the protein as a concentrated solution. Homogeneity concerns not only the presence of a single protein or complex, but also the electronic charge, conformational, cleavage and modification states. During the building of a crystal lattice, the addition of an inhomogeneous unit can ruin the crystal contacts required to continue crystal growth. The high concentration of protein is required to

allow the protein to come to a saturated state where it will hopefully form ordered aggregates, although the formation of disordered ones such as precipitates is usually more common. Much progress has been made in recent years to automate the process of crystallization, making it possible to screen for crystallization conditions using smaller amounts of protein while testing a larger number of potential crystallization conditions [30]. In addition, developments at synchrotron sources now allow the collection of data from much smaller crystals, software improvements allow more rapid processing and interpretation of the data, and structure determinations for homologous proteins such as the kinases are often routine. All of this has contributed to the rather rapid increase in kinase and other structural information generated over the past 10 years [31] (Table 1).

Protein crystals are extremely fragile since they are actually made up of protein molecules linked by just a few interactions such as hydrogen bonds, salt bridges or hydrophobic packing. The proteins are surrounded by solvent which on average comprises about 45% of the content of the crystal [32], but can vary between <30% to >80%. This makes the crystals a very good representation of how the protein looks in solution, since only a few surface residues are involved in crystal contacts and likely to have their conformation influenced by the crystal packing. In favorable cases where the binding site is empty and not involved in crystal contacts, already formed crystals can be soaked in compounds that bind to the protein to produce structures of complexes. This is a popular method to rapidly obtain structures for the support of drug discovery programs. However, kinases are very flexible proteins and tend to adapt to the binding of different inhibitors; it is therefore often necessary to form the complex before crystallization and then find conditions under which that complex will form crystals, because the shape and the surface properties of the protein will require a different packing arrangement.

Protein crystallography is limited to what will crystallize, and if the domains of a protein are separated by flexible linkers, or if their association is only transient or weak, it is usually difficult or even impossible to obtain crystals. However, multidomain proteins will crystallize if there is a very well defined globular and low-energy conformation of the protein, as is found for the Src family kinases [33–35]. The quality of protein packing in the crystal has an influence on the amount of information that can be obtained about the structure. Poorly packed or poorly ordered crystals diffract X-rays to only low resolution, while well-ordered crystals can diffract X-rays to high resolution (up to 1.5 Å or even better) and will provide accurate information about the locations of the non-hydrogen atoms in the crystal. For example, a resolution of 3.5 Å might allow the tracing of most of the path of the amino acid chain but will not give accurate information about ligand binding or side chain conformations, while 2.5 Å or better will allow location of water or other solvent molecules, and beyond 2.2 Å, it is often possible to see alternate conformations of parts of the structure [36].

**NMR spectroscopy**

NMR spectroscopy is highly complementary to crystallography as it can provide information about flexibility and dynamics in addition to solving solution structures of proteins and their ligand complexes [37]. NMR measurements of the dynamics of enzymes involved in signaling have shown that conformational flexibility on the microsecond to millisecond time scale is essential for their catalytic action [38, 39]. However, such measurements have not yet been published for the catalytic domains of kinases. One of the main reasons for this is the need for isotopically labeled protein, because these studies require NMR resonance assignments. The latter involves labeling the protein uniformly with $^{15}N$ and $^{13}C$, and also $^{2}H$ for proteins larger than 25 kDa, which allows sequential resonance assignment based on triple-resonance NMR experiments [40]. Labeling can be achieved by expressing the protein in *E. coli* in an isotopically enriched environment [41], but until recently, it has not been possible to express many of the tyrosine kinases in this system. Theoretically, labeling can also be done in eukaryotic expression systems such as baculovirus-infected insect cells; however, the enriched medium for protein isotope labeling is not yet commercially available and will probably be very expensive. Recent experiments with Abl kinase have shown that it is possible to perform selective amino acid labeling or complete uniform labeling in specially designed media in baculovirus-infected insect cells [8, 42]. Selective labeling allows all residues of a particular amino acid type in a protein to be labeled, which means that only those residues will be visible in the spectra. Strategies to perform resonance assignments and extract structural and dynamic information are being applied to Abl kinase [8, 43]. In contrast to the lack of NMR structures of the catalytic domain, many of the smaller modular domains of protein kinases are amenable to expression in *E. coli* and their solution structures have been determined (Table 1).

One of the strengths of protein NMR spectroscopy is the ability to detect ligand binding, which can be done robustly without having isotopically labeled protein [44]. These techniques are suitable for identifying weak binders with affinities in the millimolar or micromolar range, and higher-affinity ligands can be identified using ligand displacement studies [45]. Additional experiments with isotopically labeled protein, spin-labeled protein or spin-labeled ligands can be used to characterize the binding site or even determine the mode of binding [46]. Fragment-based screening and optimization of the fragment hits has been employed to discover novel, high-affinity

kinase inhibitors [47]. These binding studies have several advantages over the classical enzymatic inhibition assays for identification of novel inhibitors, such as the ability to target the inactive conformation of the kinase, the identification of unique allosteric binding sites that are likely to favor selectivity, the lack of need for potentially time-consuming assay design and the possibility to identify very small compounds which will by definition be weak binders, but provide an excellent basis for chemical optimization toward the design of new drugs [48].

## Kinase structure and flexibility

The structures of protein kinases in the active state are all very similar, despite the fact that they have different substrate specificities and different mechanisms for control [12, 49–53]. This similarity results from the constraints on the spatial arrangement of residues important for catalysis, so that the structural elements on which these amino acids reside tend to have the same relative orientation (Fig. 2). These elements include the conserved sequence at the N-terminal anchor of the A-loop (the DFG motif), containing the aspartate responsible for coordinating a magnesium ion which positions the phosphates of ATP for phosphotransfer, and a phenylalanine that packs under helix C and is therefore important for the correct positioning of this helix and the A-loop. The position of a lysine from strand $\beta$3, which is responsible for facilitating phosphoryl group transfer without influencing ATP binding, is often stabilized by the formation of a salt bridge with a glutamate residue from helix C. The extended $\beta$ sheet conformation of the P-loop ($\beta$ strands 1 and 2) also has a role in coordinating the phosphates of ATP. The $\beta$6 strand forms part of the catalytic loop, which contains an aspartate that orients the substrate tyrosine hydroxyl group and possibly facilitates the nucleophilic attack of the $\gamma$-phosphorous atom of ATP by abstracting a hydrogen from the tyrosine, and an asparagine that coordinates a second magnesium ion [7]. Interestingly, this loop is the only catalytic element that does not differ between the active and inactive states (Fig. 2).

The structure and sequence of the residues in the C-terminal anchor of the activation segment, which consists of the P+1 loop with a conserved PxxW sequence in tyrosine kinases and a short helix ($\alpha$EF) with a conserved glutamate residue, is important for substrate recognition [54]. In the few examples of structures of tyrosine kinases with substrates bound, the substrate peptide is found to bind in an extended conformation and forms anti-parallel $\beta$ sheet interactions with $\beta$11 of the A-loop [12, 50, 51]. The trypotophan from the P+1 loop of the kinase interacts with an arginine preceding the conserved asparagine in the catalytic loop (e.g. Igf1r with sequence AARN; see Fig. 1b), or its equivalent (e.g. Arg366 of Lck with the sequence RAAN), thus stabilizing the relative positions of the substrate-binding site and the catalytic loop. The P+1 loop proline forms a basis for packing with the tyrosine of the substrate and is therefore important for substrate recognition. The proline and the tryptophan of the P+1 loop of PTKs are usually replaced by threonine and tyrosine in Ser/Thr kinases. Residues determining specificity also come from $\alpha$EF and $\alpha$G. For example, amino acids in this region of Irk provide a hydrophobic pocket that is optimal for long hydrophobic side chains such as the methionine residues in the P+1 and P+3 positions of its substrate, where P0 is the site of phosphorylation [50]. Basic residues in the region (e.g. Lys1085 from $\alpha$D) seem to provide the appropriate electrostatic environment for acidic side chains at the P-1 position of the substrate in many tyrosine kinases. Thus, the specificity results from the complementarity between the shape, hydrophobicity and electrostatic potential of the protein and the few residues either side of the tyrosine phosphorylation site of the peptide.

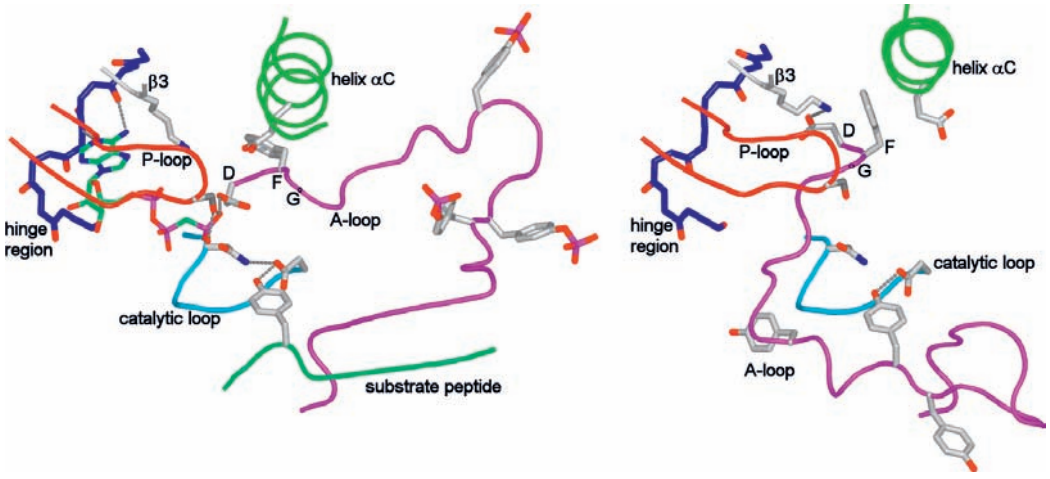


**Figure 2.** Details of the binding of an ATP analog in active Igf1r (PDB entry 1K3A) and the location of the same structural elements in inactive Igf1r (PDB entry 1M7N) showing residues important for the mechanism of action. Color-coding is as in Figure 1b.

In most tyrosine kinases, phosphorylation of a tyrosine in the A-loop stabilizes the active conformation through electrostatic interactions between the phosphate and basic residues of the A-loop and from the catalytic loop [54]. However, structures of phosphorylated kinases can still show some disorder in this region, and the binding of substrate may be necessary to stabilize the conformation required for activity [11]. In addition to the highly conserved site used to stabilize the active conformation, some kinases have secondary phosphorylation sites in the A-loop, each with different roles. For example, the insulin-like growth factor receptor kinase has three tyrosines which are phosphorylated in a stepwise fashion leading to incremental increases in activity [51]. The secondary sites are phosphorylated first and, rather than promoting interactions necessary for catalysis, are believed to contribute to activity by releasing the A-loop from the inactive conformation. In addition, the phosphorylation of these side chains results in their exposure to play a role in downstream signaling, as observed for insulin receptor kinase, where they are involved in the recruitment of the adaptor protein APS [55]. FGFR1 kinase is phosphorylated sequentially on five tyrosine residues [56]. Autophosphorylation in trans of Y653 in the A-loop enhances phosphorylation of the remaining sites (Y583 in the KID, Y463 in the JM region, Y585 in the KID and Y654 in the A-loop) resulting in stepwise increases in the catalytic activity. The work of Furdui and colleagues [56] also shows that phosphorylation of Y653 seems to be sufficient to trigger the recruitment and assembly of signaling proteins. Recruitment of the correct substrates before the kinase is fully activated may be a way of avoiding erroneous phosphorylation and signaling.

There are also examples of PTKs that do not require phosphorylation of tyrosines in the A-loop for activity [12, 57, 58]. In these kinases, the A-loop can adopt a similar conformation without the electrostatic stabilization contributed by phosphate. In the case of CSK, the positioning of the SH3, linker and SH2 domains influences the orientation of helix C, which stabilizes the conformation of the A-loop so that phosphorylation is not required. EGFR kinase does not require phosphorylation for activity, and mutagenesis of the A-loop tyrosine has no affect on activity [59]. Structural studies show that the tyrosine in the A-loop of EGFR superimposes very well with those of activated tyrosine kinase structures, but that a neighboring glutamate residue may play the role of a structural mimic for the phosphate [57].

Another critical element of kinase activity is the flexibility of the relative positioning of the N- and C-terminal lobes of the kinase [60]. Although the active kinase structures all show similar opening angles of the cleft between the two lobes, the many inactive kinase structures reveal a large range, from relatively closed compared with the active state to much more open conformations. The ability of the two lobes to rotate with respect to each other is thought to be normal and is a requirement for catalytic activity [61]. Restriction of this motion, for example by the binding of regulatory domains in the case of c-Src or Hck [26, 62], contributes to inactivation of the kinase. Restoration of the flexibility by mutating residues in the linker between the regulatory domains of Src family proteins to glycine would be expected to restore the relative flexibility, and was found to produce a more readily activated kinase [63].

## Inactive kinase conformations

In cells under resting conditions, most tyrosine kinases are held in an inactive state, which usually involves a conformation that disrupts the active arrangement of the catalytic residues or blocks the protein from binding cofactor or substrate. Unlike active kinases where the structures are very similar due to the constraint of catalytic activity, inactive kinase conformations are structurally diverse and structural biology has revealed several different mechanisms of self-regulation. In most cases, the position of helix C and or the A-loop is involved, and very often regions outside the kinase domain fold back to block the binding sites or cause conformational changes that render the kinase inactive. Many of these mechanisms are shared by kinases from distinct tyrosine kinase subgroups and also with non-tyrosine kinases.

The first crystal structure of insulin receptor kinase was determined using unphosphorylated protein and revealed an auto-inhibited state in which the A-loop blocks the binding of cofactor and substrate [5] (Fig. 3). In detail, the DFG motif at the N-terminal anchor of the A-loop adopts a conformation in which the phenylalanine and glycine are flipped over a distance of about 10 Å to block the ATP-binding site and the aspartate side chain is no longer appropriately oriented to coordinate the metal ion for correct positioning of ATP. The portion of the A-loop including Tyr1162 blocks the substrate binding site by imitating a substrate. Furthermore, the inactive structure has a more open conformation and both helix C and the P-loop from the N-terminal lobe are greatly shifted

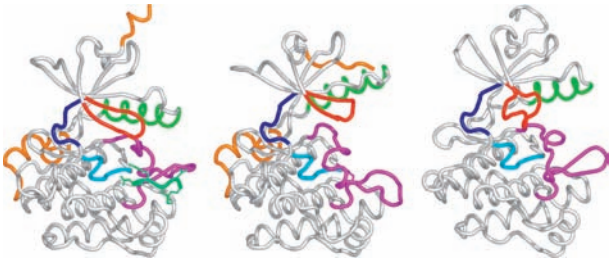

**Figure 3.** Comparison of the overall structures of activated insulin receptor kinase (PDB entry 1IR3, left) with: unphosphorylated Irk (PDB entry 1IRK, middle) and unphosphorylated Abl kinase (PDB entry 1IEP, right). Color scheme as in Figure 2.

compared with the conformations observed in the fully active state [50]. A similar inactive state is observed for the highly homologous insulin-like growth factor receptor 1 kinase (Igf1r) in the unphosphorylated state [64] (Fig. 2). Another structure of partially phosphorylated Igf1r [65] reveals an intermediate conformation between the activated and auto-inhibitory states, where most of the A-loop is disordered in the crystals. In the fully unphosphorylated state, the sequestering of the A-loop tyrosine in the substrate site (Tyr1136 in Igf1r) presumably protects it from phosphorylation and, therefore, activation. In the fully phosphorylated and thus fully active state, pTyr1137 stabilizes the active conformation of the A-loop, and pTyr1132 and pTyr1136 are exposed and may be the target of downstream signaling proteins [51].

The mode of self-regulation involving the blocking of the substrate site with the A-loop acting as a pseudo-substrate inhibitor is also used by other RTKs such as MuSK [66] and non-RTKs such as Abl kinase [67]. The inactive MuSK and Abl conformations also show a similar flip of the DFG motif (often referred to as the 'DFG-out conformation'), which creates a largely hydrophobic pocket at the back of the ATP-binding site in the case of Abl [27, 28, 67, 68]. In this case, the inactive conformation may be stabilized by the presence of an inhibitor that occupies the 'DFG-out' pocket. In the cases of MuSK and Abl, the C helix is not shifted as dramatically as in the inactive Irk kinase, and the Glu-Lys salt bridge that is observed in active kinases is retained. There are several structures of Abl in the inactive conformation and in many of these, the conformation of the P-loop is observed to adapt to the properties of whatever ligand is bound [27, 28, 67, 68]. Another type of inactive conformation of Abl kinase has been observed in several structures where, by forming a hydrogen bond with a backbone carbonyl oxygen, the aspartate side chain of the DFG motif has flipped over to take the place of the phenylalanine [68–70]. This kind of interaction can happen in the microenvironment of a protein when the electrostatic potential of the region results in protonation of a normally acidic side chain. However, it is not known whether this is an artifact of the environment of the protein in the crystals (e.g. relatively low pH), or whether it is a naturally self-regulating conformation of the protein. Kinases quite likely have more than one inactive state, because there are no mechanistic constraints on conformation like those required for the highly conserved active conformation.

Another common inactive kinase conformation is observed in the assembled inactive state of the Src kinases c-Src and Hck [26, 62]. In these kinases, the inactive state is characterized by a slight closure of the cleft between the N- and C-terminal lobes, and the adoption of an A-loop conformation that helps to fill the cleft and hold helix C away from the ATP-binding site, thus locating the important glutamate side chain more than 10 Å from its ideal position. It has been suggested that the pinning down of a tryptophan residue between helix C and the $\beta$ sheet of the N-terminal lobe might contribute to this 'C helix out' conformation. This has been confirmed in the case of Hck by mutation of this tryptophan, located at the N terminus of the kinase domain, to alanine, resulting in activation of the kinase [71]. A recent structure of c-Src kinase where the C-terminal tail tyrosine is not phosphorylated shows that the kinase domain can adopt an active conformation despite the fact that the A-loop is not phosphorylated, resulting from a release of the steric constraints on the position of the tryptophan [72]. Molecular dynamics calculations also suggest that the N-terminal end of the kinase domain acts as a conformational switch between the active and inactive states [73].

The structure of the fibroblast growth factor receptor 1 in the unphosphorylated state shows another means of self-regulation [74]. The P-loop folds down into the ATP-binding site and therefore blocks its binding. This loop is very flexible and often adapts to the shape of bound ligands [75], so it is not clear whether this would be sufficient to prevent the activation of the kinase. In addition to the inactive P-loop conformation, the A-loop adopts a unique inactive conformation that prevents the binding of substrate, and the kinase is held in a rather open state due to the interactions between residues from helix C and the DFG motif.

## The structural role of kinase regulatory domains

The inactive conformation of the kinase domains of the nRTKs c-Src and Hck have been described above, as has the importance of flexibility for their activity. The structures of the assembled inactive states of these two proteins show how the SH3 and SH2 domains lock the kinase into a rigid conformation by binding to the distal face of the kinase domain far from the ATP site (Fig. 4 [26, 33–35, 62]. The SH2 domain binds a phosphorylated tyrosine at the C terminus of the kinase domain and the SH3 domain binds to the linker between the SH2 domain and the kinase domain. This form of regulation has been described as an inducible snap-lock mechanism, where flexible linkers snap into rigidity when they bind their partners [63, 76]. The C-terminal phosphotyrosine behaves as a latch and the SH3-SH2 domains act as the lock. Any event which disrupts this assembled inactive state will lead to increased flexibility and a breaking of the snap-lock, thus making it easier for the A-loop tyrosine to be phosphorylated for full activation of the kinase. Once the kinase is activated, the regulatory domains become available to protein-protein interaction partners for downstream signaling, but the binding of a signaling partner can also activate the kinase [77]. However, there are multiple states in which the SH3 and SH2 domains are more or less accessible and this probably determines the prime

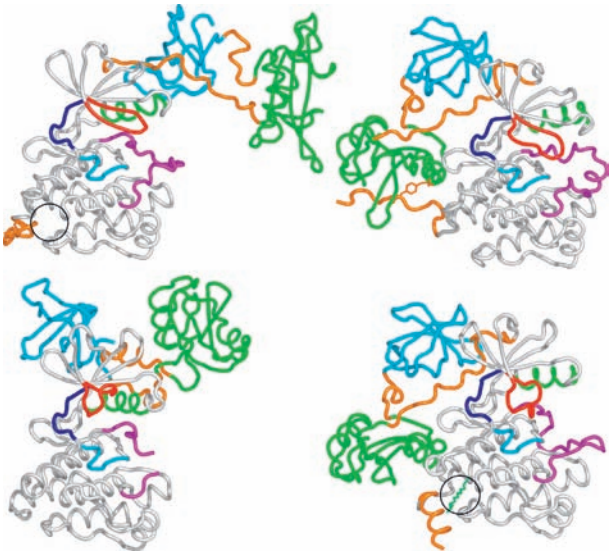

**Figure 4.** Comparison of active and inactive conformations of SH3 (blue) and SH2 (green) domain-containing structures. Top row, partially active c-Src (PDB entry 1Y57, left), assembled inactive c-Src (PDB entry 2SRC, right). Bottom row, inactive CSK (PDB entry 1K9A, left) and inactive Abl (PDB entry 1OPK, right). Black circles indicate the myristate (green) binding site in Abl kinase and the C-terminal leucine site in partially active c-Src.

signaling pathway. For example, the binding of HIV-1 Nef activates Hck by displacing the SH3 domain, without disrupting the SH2-phosphotyrosine latch [77]. This type of activation might favor signaling via interactions with the SH3 domain. In the case of unphosphorylated c-Src, the latch is disrupted, but the SH3-linker interaction is retained, which may favor signaling via interactions with the SH2 domain [72].

Another kinase using this method of self-regulation is Abl, which also has SH3 and SH2 domains. However, in the case of Abl, the latch is not a C-terminal phosphorylation site, but a myristoyl-binding site in the C-terminal lobe of the kinase [69, 78]. The unique N-terminal region of Abl kinase harbors a myristoylation site and has been shown to be involved in the negative regulation of the kinase [79]. The binding of myristoyl ester in the C-lobe pocket stabilizes a conformational change in helix I of the kinase domain, which allows the SH2 domain to make even closer interactions with the C-terminal lobe of the kinase domain than are observed for c-Src [69]. It is interesting to note that the SH3-SH2 domains of the c-Src and Abl assembled inactive states have very similar conformations, but the inactive conformations of the kinase domains are quite distinct (see above) [80]. This difference is coupled with the unique modes of intereaction between the SH2 and kinase domains, which allows the Abl kinase domain to come much closer to the SH2 domain than in c-Src, and this allows a more open structure for the kinase domain. Further studies of the Abl structure have recently revealed that the unique N-terminal region has extensive contacts, involving a phosphoserine-binding site, with the connector between the SH3 and SH2 domains, which serves to stabilize further this locked inactive conformation [81].

The discovery of a similar pocket in the same location as the Abl myristate-binding site in the active conformation of c-Src kinase, where the C-terminal leucine is bound, suggests that c-Src may also be able to bind myristate. NMR experiments confirmed this hypothesis for the C-terminal tyrosine-phsophorylated form of the protein [72]. This suggests that c-Src may have an extra level of self-regulation where the binding of the myristoylated N terminus in this pocket sequesters the myristoyl group and prevents it from interacting with the membrane. The latter interaction is necessary for the activity of the protein [82]. This sequestering would also prevent the binding of the C terminus in the pocket, which would cover the C-terminal tyrosine and thus protect it from phosphorylation. The multiple levels of regulation observed for the Src family and Abl kinases suggest that these and other kinases are not regulated by a simple on/off molecular switch, but an intricate matrix of events that have many and varied points of access within the cell.

C-terminal Src kinase (CSK) also has N-terminal SH3 and SH2 domains, but no N-terminal acylation site and no A-loop or C-terminal tyrosine phosphorylation site. Biochemical studies show that the SH2 and SH3 domains are required for full catalytic activity [83]. It was not surprising, therefore, that the crystal structure of this protein revealed a novel arrangement of these domains with respect to the kinase domain [58]. The two SH domains lie on opposite sides of the N-terminal lobe and do not directly interact with each other. The SH3-linker-SH2-linker region seems to be responsible for stabilizing the active conformation of helix C, and thus the A-loop, because the structure of the kinase domain alone reveals catalytically incompetent positions of these two elements [20].

Most RTKs have a juxtamembrane (JM) region lying between the membrane and the kinase domain, which is involved in regulation. For example, the structure of Flt-3 [14] shows that the N-terminal part of the JM region forms a wedge which seems to stabilize the inactive conformation of the kinase domain by packing against helix C, the P-loop and the A-loop (Fig. 5). The structure suggests that any change in the length of the JM region will destabilize the binding to the kinase domain and lead to a loss of autoregulation. The length of the C-terminal part of the JM region is strictly conserved among members of the PDGF RTK family suggesting that they will all have a similar method of autoregulation. Internal tandem repeats are found in the JM region of many patients with acute myelogenous leukemia who show a high level of Flt-3 expression [84]. Mutations in the JM region of the stem cell factor receptor c-Kit also result in constitutive

activation and are responsible for a number of cancers. The C-terminal part of the JM region of c-Kit harbors two tyrosines that lead to kinase activation upon phosphorylation. The structure of the activated and the inactivated forms of c-Kit [12, 85] show that phosphorylation, and presumably mutation, prevents the JM region from adopting a regulatory state where it can block the normal path of the A-loop required for the active conformation (Fig. 5). Another kinase structure with part of the JM region present and likely to share a similar mechanism of regulation is that of EphB2 [21]. In this structure, the JM region stabilizes a distorted conformation of helix C and prevents the A-loop from adopting the active conformation. However, a structure of the isolated kinase domain of the highly homologous EphA2 shows the same distorted helix C conformation, so phosphorylation of the A-loop may be equally important as phosphorylation of the JM region to achieve the active conformation [30]. The insulin receptor tyrosine kinase family also has a JM region that contains a conserved tyrosine (Tyr984 in Irk). This tyrosine is found to stabilize an inactive position of helix C, by binding in a cleft between the N-terminal lobe $\beta$ sheet and the helix [86].

A similar JM-type mode of regulation seems to be found in the nRTKs. The sequence connecting the SH2 domains with the kinase domain is called 'interdomain B' in the Zap70 and Syk kinases and shows some homology to the EphB2 JM region. Mutation of conserved residues in this region revealed that interdomain B is involved in down-regulation of Zap70 [87]. Tyrosines 315 and 319, mutation of which has also been shown to impair signaling events [88], must be phosphorylated for full activation of the kinase.

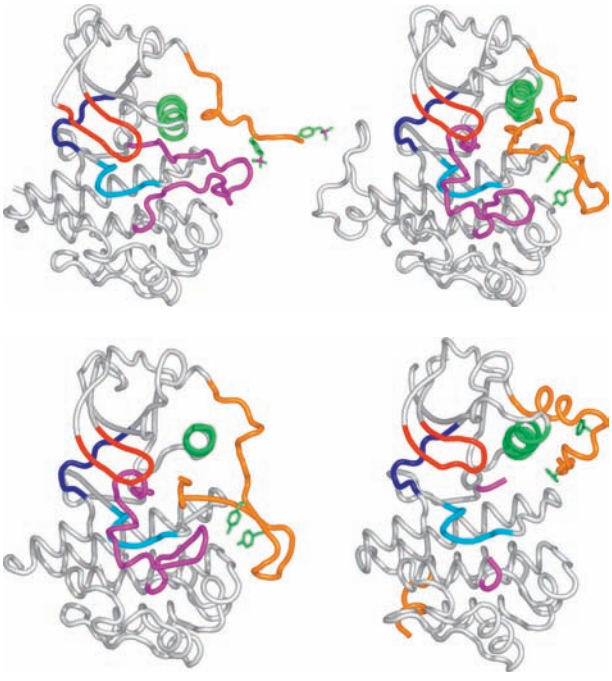

**Figure 5.** Comparison of the JM regions of cKit (active, 1PKG, top left; inactive, 1T45, top right), FLT3 (1RJB, bottom left) and EphB2 (1JPA, bottom right). The tyrosines which are phosphorylated on activation are shown in green.

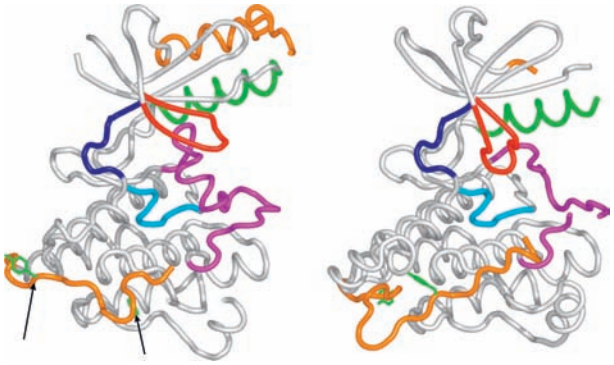

**Figure 6.** c-Met kinase (1R0P, left) and Tie2 kinase (1FVR, right) both have C-terminal regions (orange) that fold back onto the kinase domain. Activation and downstream signaling are promoted by phosphorylation of the residues indicated in green (and with arrows in c-Met).

Another regulatory element of the PTKs is the C-terminal tail of the kinase, which is a region of large variation in sequence and length. A structure of Tie2 containing the C-terminal tail shows that it folds back into a cleft that extends toward the substrate-binding site, possibly preventing the binding of substrate [18] (Fig. 6). Biochemical studies have confirmed the autoregulatory role of this region in Tie2, as its deletion leads to activation of the kinase [89]. Along with other RTKs, such as VEGFR and c-Met, Tie2 is activated by phosphorylation of the C-terminal tail, and this activation is important for the interactions with binding partners for downstream signaling [90–93]. In the structure of c-Met, the C-terminal tail has structural motifs that would be suitable for the binding of PTB and SH2 domains [19]. However, they are mostly buried and inaccessible for their binding partners. Presumably, as for Tie2, phosphorylation would expose the C-terminal tail and make the binding sites available for downstream signaling.

Kinases in the PDGFR/VEGFR subfamily, have a kinase insertion domain (KID) between helices $\alpha$-D and $\alpha$-E. The role of this structural element is currently unclear. In many structures, this domain had to be deleted from the construct in order to express a protein suitable for crystallization [see e.g. refs. 11, 13, 14]. It is postulated that this region provides binding sites for protein interaction partners [51].

## Beyond the kinase domain

Further understanding of the regulation and signaling of PTKs will require structures in context with their associated domains. For example, in the Jak kinase family, there is a pseudo kinase domain that is believed to interact

with and negatively regulate the tyrosine kinase [94, 95], but it is unclear in which way this interaction is formed. Other domains associated with a regulatory role include the FERM domains, also found in the nRTK family. A structure of the FAK kinase FERM domain suggests how the protein interactions within the FERM domains might be regulated by the binding of Src kinases which would stabilize an open active conformation [96]. The linker between the FERM and kinase domains masks a potential protein interaction site on the F3 lobe of the FERM domain. A PxxP motif and a tyrosine phosphorylation site in the linker can bind the SH3 and SH2 domains of Src kinase, which would release the linker, allowing the FERM domain to pursue its role in membrane localization and regulation of kinase activity.

Structures of isolated domains beg questions concerning how they interact with other domains and with other proteins in the cellular environment. Protein crystallography is capable of seeing these interactions as long as the partners can be trapped in a crystal. But, perhaps due to the transient nature of many of these events, and also the inherent flexibility of multidomain proteins, there are few examples available. There are exceptions of course, and many of these have appeared over the last few years, but the success rate is still too low to satisfy our thirst for structural information regarding protein-protein interactions. Some of the recent successes concern the insulin receptor tyrosine kinase. The structure of the tyrosine kinase domain in complex with PTP1B reveals that the high specificity of this phosphatase for IRK results from a non-catalytic mode of binding distant from the phosphorylated tyrosine in the A-loop, where significant sequence differences are found compared with even the closest homolog Igf1r [97]. In this structure, the phosphorylated A-loops are sequestered in an IRK-IRK dimer interface, so the complex may represent a mode of binding important for the selective recruitment of PTP1B to the insulin receptor. A previous structure of PTP1B with a phosphorylated A-loop peptide from IRK reveals the catalytic mode of interaction between the two proteins [98], but the lack of significant sequence or structural differences in the A-loop had left the selectivity unexplained.

A structure of IRK in complex with a peptide from the BPS region of the tissue-specific negative regulator Grb14 shows that the N-terminal part of the BPS region imitates substrate by binding in the substrate site of the kinase [99]. This structure provided the basis for a model to explain how the binding of the Grb14 SH2 domain to the phosphorylated A-loop and its dimerization could enhance this inhibitory effect. Binding to the adaptor protein APS also involves an SH2 domain, which interacts with pTyr1158 and pTyr1162 from the A-loop of IRK [55]. The recruitment of the APS protein is one of the first steps in downstream signaling of the insulin receptor. Again, dimerization of the SH2 domain is involved, but in this case, the form of dimerization is unique and involves conformational changes that would interfere with the normal extended $\beta$ sheet conformation observed in other SH2-pTyr-containing peptide complexes [55].

The flexibility of these modular domains is also observed in a ternary Fyn SH3-SAP-SLAM complex [100]. SAP, which consists of a single SH2 domain, is required for the recruitment and activation of the Src family kinase Fyn after SLAM ligation. Although the surface of the Fyn-SH3 domain recognized by SAP is different to that of the SH3-linker interaction in the assembled inactive state of Src kinases [26, 62], it does overlap and would therefore be expected to prevent the adoption of the inactive state. The novelty of the interaction between SAP and Fyn-SH3 is that it does not involve a polyproline structure and therefore reveals some diversity in the mode of interaction of the SH3 domain. The unconserved binding surfaces observed in the SH2 and SH3 domains may provide a mechanism for the different binding partners to stimulate different signaling pathways. Another domain of interest for activity in the nRTKs is the unique region of the Src family kinases. In isolation, this region is largely unstructured and for most members of the family its function is unclear. In Lck it mediates the association with the cytoplasmic tails of the T-cell receptors CD4 and CD8$\alpha$ [101]. NMR studies reveal that fragments of the unique region and of the CD receptor tails gain structure in each others' presence and in the presence of Zn. In the case of CD4, the interaction buries a dileucine motif that is necessary for internalization of CD4, which also requires phosphorylation of Ser408 in CD4 followed by dissociation of Lck. In the structure, the serine is exposed and readily accessible for phosphorylation, but the mechanism of the subsequent dissociation of Lck is not obvious [101].

Other examples of multidomain structures concern the extracellular domains of the RTKs. Structures of FGF-FGFR complexes reveal how FGF interacts with the extracellular binding domain in the inactive and active states and provide a model for the heparin-induced dimerization that leads to autophosphorylation of the intracellular kinase domains and activation [102–105]. The binding of heparin to the D2 Ig-like domain of FGFR and to FGF leads to formation of a stable 1:1:1 FGF:FGFR:heparin ternary complex. A second ternary complex can then be recruited, involving direct receptor-receptor interactions, but no direct FGF-FGF interactions, that is enhanced by the presence of the two heparin molecules which also contact the adjoining receptors [106]. Based upon structures which show the dimerization of the ligand alone, ligand-induced dimerization is observed or postulated for the VEGF, SCF and PDGF receptors, which is in contrast to the case of FGF [107]. Structures of the extracellular regions for Igf1r, EGFR, ErbB2 and ErbB3 show differing arrangements that provide insights into yet another model for ligand-induced oligomerization and activation

[see ref. 108 for a review]. The L1-CR1-L2 fragment of Egfr shows a mode of dimerization involving direct interaction of the receptors, but no interactions between the ligands bound to each fragment [109, 110]. A structure of the full extracellular domain of ErbB3 in the absence of ligand shows a completely different conformation, in which the dimerization arm of CR1 is buried against CR2 so that oligomerization cannot occur [111]. A similar arrangement has been found for Egfr [112]. This suggests that ligand binding induces a 130° rotation of the L1 and CR1 domains, which is essential for the correct positioning of the two receptor molecules to achieve dimerization and, thus, kinase activation [113]. ErbB2, which does not bind ligands and is known to form heteromeric complexes with ErbB1, ErbB3 or ErbB4 in order to provide amplification and diversification of signal output, has a structure in an active conformation [114, 115]. It is unable to form homodimers due to electrostatic repulsion.

The Trk receptors provide further examples of ligand-induced dimerization involving only direct ligand-ligand interactions [116, 117]. Binding of the homodimeric neurotrophins to the fifth Ig-like domain causes formation of a dimeric structure that separates the two receptor domains by about 23 Å. This distance is presumably necessary for the appropriate positioning of the intracellular tyrosine kinase domains for autophosphorylation, to trigger downstream signaling activities.

To further our understanding of these protein-protein interactions, other techniques and hybrid approaches are required. Protein crystallography can be complemented by the use of small-angle X-ray scattering (SAXS), which gives a low-resolution picture of the relative arrangement of protein domains in solution. For example, SAXS studies of multidomain constructs of the nRTKs Btk and Abl give some insight into their various states of activation [81, 118]. Other methods include electron microscopy, which revealed images of the arrangement of the multiple domains of the entire Igf1r [119]. Cryoelectron tomography is a developing technique which can visualize single macromolecules in their native cellular environment [120, 121], and recent developments using hydrogen/deuterium exchange mass spectrometry have proven valuable in the analysis of the domain interactions of Hck [122] and CSK [123, 124]. Recent advances in FTIR spectroscopy using $^{13}C/^{15}N$ uniformly labeled proteins have been useful in the investigation of BDNF-TrkB ligand-receptor interactions [125]. The coupling of site-directed mutagenesis experiments with molecular dynamics calculations for Hck and Abl provided a model for the role of the SH3-SH2 linker in regulation and activation [63, 69, 78]. Other computational methods, such as protein-protein docking using the structures of the individual domains, can often identify surfaces involved in interactions, even if the flexibility of the surfaces involved precludes accurate docking. These methods are regularly evaluated using blind trials with known structures [126]. This list of techniques is by no means complete as there are many other experimental and theoretical methods used in a hybrid fashion to address the challenge of understanding the roles of these proteins in a cellular context [127].

## Supporting drug discovery

One of the driving forces for the study of tyrosine kinases using structural biology is the use of the structural information to contribute to structure-based drug discovery [128]. In the early phases of discovery, fragment-based screening (FBS) for binders by NMR and/or X-ray [129] can complement or even replace high-throughput screening based on biochemical or cellular assays [130]. The method is based on the idea of incubating the protein with a mixture of molecular fragments of low molecular weight (usually between 100 and 250 Da), and detecting the binding of these fragments based on NMR spectra or the shape of the difference electron density in crystal structures. Due to the low functionality of the fragments (e.g. few hydrogen bond donors/acceptors), they are expected to have very low affinity in biochemical assays (micromolar to millimolar range) and would not normally be detected in a high-throughput screen unless the screen was designed to detect such compounds. As long as the solubility of the fragments is sufficient to achieve high concentrations, X-ray crystallography will show how they bind, and this provides an excellent basis for computational and medicinal chemistry to increase potency of binding to the target. This can be done by linking fragments that bind in different subpockets, or by the elaboration of the initial scaffold [129]. Examples of the applications of FBS by X-ray crystallography to kinases have shown that fragment hits can be optimized to provide inhibitors with nanomolar potency with just a few rounds of medicinal chemistry [131, 132]. A key attraction of the FBS method is the potential of finding novel chemical entities that would not have been identified by more established methods, such as high-throughput screening or even newer methods such as virtual screening. High-throughput screening is limited by the complexity of the library that is screened – if the chemical entity is surrounded by groups that prevent its binding to the protein, that type of compound will not come up in a high-throughput screen [133]. Virtual screening is limited by the flexibility of the protein target [134, 135], the difficulty of accounting for water molecules and the problems of generating a meaningful score for the different docking poses [136, 137]. However, protein crystallography can provide a good basis for these computational studies, by producing cocrystal structures of complexes with a wide variety of ligands that promote different conformational states. Kinases are flexible proteins and it has been shown that the selection

of an appropriate crystal structure is important for the success rate in correctly predicting the binding mode of kinase inhibitors [136].

In addition to the potential for discovery of new inhibitor scaffolds, structural biology also provides the means to discover or design compounds that bind to novel (non-ATP) binding sites on the protein and lead to inhibition or even stimulation of the kinase activity. For example, compounds predicted to bind to the myristate pocket on the C-terminal lobe of Abl kinase, which is about 20 Å from the ATP-binding site, show anti-proliferative activity to Bcr-Abl-transformed cells at a similar level to the marketed drug Glivec (imatinib), yet do not inhibit the kinase activity of the isolated protein [138]. A potential binding site for compounds to activate the insulin receptor, which could possibly be used in the treatment of non-insulin-dependent diabetes, has been proposed based on a comparison of the activated structure and an inactivated form where part of the JM region is present [86]. The comparison shows that a cavity is formed by the displacement of the JM region including Tyr984 and the shift of helix C into the active conformation. Compounds have also been designed to bind to the substrate-binding site of kinases, where good selectivity should be possible, but potency is more elusive. Based on the crystal structure of the IRK-substrate peptide complex and mechanistic studies, a potent and selective inhibitor could be built by linking ATP and a peptide, thus creating a bisubstrate analog [139]. The potency is greater than either substrate alone, since the binding of one part increases the local concentration of the other, creating a favorable entropy effect. The selectivity arises from the specific peptide sequence. However, the bisubstrate inhibitors are very large and do not have drug-like properties.

Structures of kinases in complex with their inhibitors are a useful basis for optimizing the potency, selectivity and physicochemical properties of these inhibitors. By analyzing the interactions between the protein and the compound, modifications can be made to improve the shape complementarity of the surfaces, to take advantage of potential electrostatic interactions, to identify regions that can be modified to improve solubility, or identify ways to modify groups to remove unwanted properties, such as metabolic instability, without losing potency.

There are many examples in the literature of the structure-based design of kinase inhibitors and some of these studies have succeeded in contributing to the discovery of compounds which are now in clinical trials [128, 140]. Some kinase inhibitors have reached the market: imatinib as a treatment for chronic myelogenous leukemia (CML) and gastronintestinal stromal tumours, erlotinib for metastatic non-small cell lung cancer [141]. While imatinib has good specificity, is very well tolerated and is very effective in the treatment of chronic-phase CML patients even after a period as long as 4 years [142], patients first treated in the late phases of the disease tend to develop resistance and relapse [143–145]. The reasons for relapse are multiple, but the major mechanism involves mutations in the kinase domain that reduce the binding affinity of imatinib [70, 146]. Resistance mutations have also been seen in the Egfr kinase domains of patients treated with gefitinib or erlotinib [147, 148]. Structural analysis and modeling of the Abl kinase mutations found in relapsed patients shows that they are distributed throughout the kinase domain, but that the most resistant mutations are located in the P-loop and at the gatekeeper position in the hinge region (T315I). The gatekeeper is a residue that separates the ATP-binding pocket from a cavity that is of variable size among kinases. The characteristics of the side chain in this position (size, shape and polarity) are often used in the design of inhibitors to gain some selectivity. The most prevalent mutations are T315I (gatekeeper), E255K (P-loop) and M351T (C-lobe). A crystal structure of the T315I mutant [unpublished data] shows that this isoleucine causes a steric clash and loss of a hydrogen bond with imatinib [68]. Modeling of the E255K mutant suggests that the loss of two hydrogen bonds would destabilize the inactive conformation of the P-loop leading to a shift of equilibrium toward the active conformation to which imatinib cannot bind [70]. A crystal structure of the M351T mutant shows that this mutation causes a slight shift of helix E in the C-terminal lobe but no obvious changes in the inhibitor-binding site [28]. In this case, the loss of affinity of imatinib, which is only mildly resistant compared with the previous two mutations, may be due to an increase in the entropy of the protein, which could also cause a shift in the equilibrium toward the active state. A structure of the imatinib-resistant mutant H396P also shows how this sequence change can favor the adoption of the active state due to the steric constraints of proline, which are compatible with the active conformation of the A-loop, but not that of the inactive conformation [149]. *In vitro* studies have shown that point mutations of regions within, and in contact with, the SH3 and SH2 regulatory domains also cause resistance to imatinib, although these cause only mild reductions of affinity and have not yet been detected in CML patients [150]. These mutations would also favor the transition toward an active state to which imatinib could not bind. A structural view of the causes for relapse supports the discovery of inhibitors that can overcome this resistance, such as nilotinib and dasatinib, which are now progressing well in clinical trials to treat patients who have relapsed under imatinib therapy [28, 151, 152]. These compounds are active against all but one, T315I, of the mutations isolated from imatinib-resistant patients [153, 154].

The fact that kinases are able to mutate to avoid inhibition is true for all of the inhibitors on the market or in clinical trials, but those that inhibit the inactive con-

formations seem to give rise to a slightly larger number of these mutations [155, 156]. This is probably because they interact with more residues that do not contact ATP, and they depend on a particular conformation that is not relevant for activity. However, these unique interactions and conformations are the likely reasons for the excellent selectivity profile observed for such inhibitors, and the result is often an excellent safety profile. Combinations of different inhibitors are likely to overcome the problems of resistance due to point mutations.

## Conclusion

Structural biology has accelerated our understanding of the structure and function of protein tyrosine kinases over the past 15 years, and is providing a valuable basis for the design of innovative experiments to discover more about these critical signaling proteins, and to maximize their potential as targets for drug therapy. However, we still have a lot to learn about the full range of interactions in the cell. Most of the structural biology to date has focused on the kinase domain alone. In addition to determining the structures of multidomain complexes and members of branches of the kinase family that have not yet been investigated, structural biology will have to continue to take advantage of the interaction with other disciplines in order to deepen our understanding of the cellular environment.

*Acknowledgements.* Many thanks to D. Fabbro, W. Jahnke, P. W. Manley, P. Ramage, C. Scheufler and H. Widmer for valuable feedback on the manuscript.

1 Ortutay, C., Valiaho, J., Stenberg, K. and Vihinen, M. (2005) KinMutBase: a registry of disease-causing mutations in protein kinase domains. Hum. Mutat. 25, 435–442.
2 Manning, G., White, D. B., Martinez, R., Hunter, T. and Sudarsanam, S. (2002) The protein kinase complement of the human genome. Science 298, 1912–1934.
3 Hanks, S. K. (2003) Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. Genome Biol. 4, 111–117.
4 Knighton, D. R., Zheng, J., Ten Eyck, L. F., Xuong, N.-H., Taylor, S. and Sowadski, J. M. (1991) Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science 253, 414–420.
5 Hubbard, S. R., Wei, L., Ellis, L. and Hendrickson, W. A. (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. Nature 372, 746–753.
6 Taylor, S., Knighton, D. R., Zheng, J., Sowadski, J. M., Gibbs, C. S. and Zoller, M. J. (1993) A template for the protein kinase family. Trends Biochem. Sci. 18, 84–89.
7 Adams, J. A. (2001) Kinetic and catalytic mechanisms of protein kinases. Chem. Rev. 101, 2271–2290.
8 Strauss, A., Bitsch, F., Fendrich, G., Graff, P., Knecht, R., Meyhack, B. and Jahnke, W. (2005) Efficient uniform isotope labeling of Abl kinase expressed in baculovirus-infected insect cells. J. Biomol. NMR 31, 343–349.
9 Jin, L., Pluskey, S., Petrella, E. C., Cantin, S. M., Gorga, J. C., Rynkiewicz, M. J., Pandey, P., Strickler, J. E., Babine, R. E., Weaver, D. T. and Seidl, K. J. (2005) The three-dimensional structure of the ZAP-70 kinase domain in complex with staurosporine: implications for the design of selective inhibitors. J. Biol. Chem. 279, 42818–42825.
10 Wei, L., Hubbard, S. R., Hendrickson, W. A. and Ellis, L. (1995) Expression, characterization, and crystallization of the catalytic core of the human insulin receptor protein-tyrosine kinase domain. J. Biol. Chem. 270, 8122–8130.
11 McTigue, M. A., Wickersham, J. A., Pinko, C., Showalter, R. E., Parast, C. V., Tempczyk-Russell, A., Gehring, M. R., Mroczkowski, B., Kan, C.-C., Villafranca, J. E. and Appelt, K. (1999) Crystal structure of the kinase domain of human vascular endothelial growth factor receptor 2; a key enzyme in angiogenesis. Structure 7, 319–330.
12 Mol, C. D., Lim, K. B., Sridhar, V., Zou, H., Chien, Y. E. T., Sang, B.-C., Nowakowski, J., Kassel, D. B., Cronin, C. N. and McRee, D. E. (2003) Structure of a c-Kit product complex reveals the basis for kinase transactivation. J. Biol. Chem. 278, 31461–31464.
13 Manley, P. W., Bold, G., Brüggen, J., Fendrich, G., Furet, P., Mestan, J., Schnell, C., Stolz, B., Meyer, T., Meyhack, B., Stark, W., Strauss, A. and Wood, J. (2004) Advances in the structural biology, design and clinical development of VEGF-R kinase inhibitors for the treatment of angiogenesis. Biochim. Biophys. Acta 1697, 17–27.
14 Griffith, J., Black, J., Faerman, C., Swenson, L., Wynn, M., Lu, F., Lippke, J. and Saxena, K. (2004) Structure of FLT3 autoinhibited by the juxtamembrane domain: implications for acute myeloid leukemia. Mol. Cell 13, 169–178.
15 Pantazatos, D., Kim, J. S., Klock, H. E., Stevens, R. C., Wilson, I. A., Lesley, S. A. and Woods, V. L. (2004) Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS. Proc. Natl. Acad. Sci. USA 101, 751–756.
16 Carter, C. W. Jr (1988) Cloning heterologous genes into, *E. coli* for enzyme production and crystal growth: problems of expression and microheterogeneity. J. Cryst. Growth 90, 168–179.
17 Plomp, M., McPherson, A. and Malkin, A. J. (2003) Repair of impurity-poisoned protein crystal surfaces. Proteins Struct. Funct. Genet. 50, 486–495.
18 Shewchuk, L. M., Hassell, A. M., Ellis, B., Holmes, W. D., Davis, R., Horne, E. L., Kadwell, S. H., McKee, D. D. and Moore, J. T. (2000) Structure of the Tie2 RTK domain: self-inhibition by the nucleotide binding loop, activation loop, and C-terminal tail. Structure 8, 1105–1113.
19 Schiering, N., Knapp, S., Marconi, M., Flocco, M. M., Cui, J., Perego, R., Rusconi, L. and Cristiani, C. (2003) Crystal structure of the tyrosine kinase domain of the hepatocyte growth factor receptor c-Met and its complex with the microbial alkaloid K-252a. Proc. Nat. Acad. Sci. USA 100, 12654–12659.
20 Lamers, M. B., Antson, A. A., Hubbard, R. E., Scott, R. K. and Williams, D. H. (1999) Structure of the protein tyrosine kinase domain of C-terminal Src kinase (CSK) in complex with staurosporine. J. Mol. Biol. 285, 713–725.
21 Wybenga-Groot, L. E., Baskin, B., Ong, S. H., Tong, J., Pawson, T. and Sicheri, F. (2001) Structural basis for autoinhibition of the EphB2 receptor tyrosine kinase by the unphosphorylated juxtamembrane region. Cell 106, 745–757.
22 Seeliger, M. A., Young, M., Henderson, M. N., Pellicena, P., King, D. S., Falick, A. M. and Kuriyan, J. (2005) High yield bacterial expression of active c-Abl and c-Src tyrosine kinases. Protein Sci. 14, 3135–3139.
23 Wang, W., Marimuthu, A., Tsai, J., Kumar, A., Krupka, H. I., Zhang, C., Powell, B., Suzuki, Y., Nguyen, H., Tabrizizad, M., Luu, C. and West, B. L. (2006) Structural characterization of autoinhibited c-Met kinase produced by coexpression in bacteria with phosphatase. Proc. Natl. Acad. Sci. USA 103, 3563–3568.

24 Caspers, P., Steiger, M. and Burn, P. (1994) Overproduction of bacterial chaperones improves the solubility of recombinant protein tyrosine kinases in *Escherichia coli*. Cell. Mol. Biol. 40, 635–644.
25 Atwell, S. and Feil, I. (2004) Methods of enhancing yields of soluble proteins. Patent WO 2004/009831 A1.
26 Schindler, T., Sicheri, F., Pico, A., Gazit, A., Levitzki, A. and Kuriyan, J. (1999) Crystal structure of Hck in complex with a Src family-selective tyrosine kinase inhibitor. Mol. Cell 3, 639–648.
27 Manley, P. W., Cowan-Jacob, S. W., Buchdunger, E., Fabbro, D., Fendrich, G., Furet, P., Meyer, T. and Zimmermann, J. (2002) Imatinib: a selective tyrosine kinase inhibitor. Eur. J. Cancer 38 (Suppl. 5), S19–S27.
28 Weisberg, E. L., Manley, P. W., Breitenstein, W., Brueggen, J., Cowan-Jacob, S. W., Ray, A., Huntly, B., Fabbro, D., Fendrich, G., Hall-Meyers, E., Kung, A. L., Mestan, J., Daley, G. Q., Callahan, L., Catley, L., Cavazza, C., Azam, M., Neuberg, D., Wright, R. D., Gilliland, G. D. and Griffin, J. D. (2005) Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl. Cancer Cell 7, 129–141.
29 Rummel, G., Stark, W., Strauss, A. and Wagner, J. (2004) Three-dimensional structure of the catalytic domain of protein kinase C theta, methods and use thereof. Patent WO 2004/076654 A1.
30 Nowakowski, J., Cronin, C. N., McRee, D. E., Knuth, M. W., Nelson, C. G., Pavletich, N. P., Rogers, J., Sang, B.-C., Scheibe, D. N., Swanson, R. V. and Thompson, D. A. (2002) Structures of the cancer-related Aurora-A, FAK, and EphA2 protein kinases from nanovolume crystallography. Structure 10, 1659–1667.
31 Dutta, S. and Berman, H. M. (2005) Large macromolecular complexes in the protein data bank: a status report. Structure 13, 381–388.
32 Matthews, B. (1968) Solvent content of protein crystals. J. Mol. Biol. 33, 491–497.
33 Xu, W., Harrison, S. C. and Eck, M. J. (1997) Three-dimensional structure of the tyrosine kinase c-Src. Nature 385, 595–602.
34 Sicheri, F., Moarefi, I. and Kuriyan, J. (1997) Crystal structure of the Src family tyrosine kinase Hck. Nature 385, 602–609.
35 Williams, J. C., Weijland, A., Gonfloni, S., Thompson, A., Courtneidge SA., Superti-Furga, G. and Wierenga, R. K. (1997) The 2.35 Å crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions. J. Mol. Biol. 274, 757–775.
36 Davis, A. M., Teague, S. J and Kleywegt, G. J. (2003) Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. Angew. Chem. Int. Ed. Engl. 42, 2718–2736.
37 Wüthrich, K. (1995) NMR – this other method for protein and nucleic acid structure determination. Acta Crystallogr. D51, 249–270.
38 Volkman, B. F., Lipson, D., Wemmer, D. E. and Kern, D. (2001) Two-state allosteric behavior in a single-domain signaling protein. Science 291, 2429–2433.
39 Eisenmesser, E. Z., Bosco, D. A., Akke, M. and Kern, D. (2002) Enzyme dynamics during catalysis. Science 295, 1520–1523.
40 Clore, G. M. and Gronenborn, A. M. (1998) NMR structure determination of proteins and protein complexes larger than 20 kDa. Curr. Opin. Chem. Biol. 2, 564–570.
41 Goto, N. K. and Kay, L. E. (2000) New developments in isotope labeling strategies for protein solution NMR spectroscopy. Curr. Opin. Struct. Biol. 10, 585–592.
42 Strauss, A., Bitsch, F., Cutting, B., Fendrich, G., Graff, P., Liebetanz, J., Zurini, M. and Jahnke, W. (2003) Amino-acid-type selective isotope labeling of proteins expressed in Baculovirus-infected insect cells useful for NMR studies. J. Biomol. NMR 26, 367–372.
43 Cutting, B., Strauss, A., Fendrich, G., Manley, P. W. and Jahnke, W. (2004) NMR resonance assignment of selectively labeled proteins by the use of paramagnetic ligands. J. Biomol. NMR 30, 205–210.
44 Jahnke, W. and Widmer, H. (2004) Protein NMR in biomedical research. Cell. Mol. Life Sci. 61, 580–599.
45 Jahnke, W., Floersheim, P., Ostermeier, C., Zhang, X., Hemmig, R., Hurth, K. and Uzunov, D. P. (2002) NMR reporter screening for the detection of high-affinity ligands. Angew. Chem. Int. Ed. Engl. 41, 3420–3423.
46 Homans, S. W. (2004) NMR spectroscopy tools for structure-aided drug design. Angew. Chem. Int. Ed. Engl. 43, 290–300.
47 Fejzo, J., Lepre, C. and Xie, X. (2003) Application of NMR screening in drug discovery. Curr. Topics Med. Chem. 3, 81–97.
48 Gill, A. (2004) New lead generation strategies for protein kinase inhibitors – fragment based screening approaches. Mini-Rev. Med. Chem. 4, 301–311.
49 Yamaguchi, H. and Hendrickson, W. (1996) Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. Nature 384, 484–489.
50 Hubbard, S. R. (1997) Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. EMBO J. 16, 5573–5581.
51 Favelyukis, S., Till, J. H., Hubbard, S. R. and Miller, W. T. (2001) Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. Nature Struct. Biol. 8, 1058–1063.
52 Breitenlechner, C. B., Kairies, N. A., Honold, K., Scheiblich, S., Koll, H., Greiter, E., Koch, S., Schäfer, W., Huber, R. and Engh, R. A. (2005) Crystal structures of active Src kinase domain complexes. J. Mol. Biol. 353, 222–231.
53 Lucet, I. S., Fantino, E., Styles, M., Bamert, R., Patel, O., Broughton, S. E., Walter, M., Burns, C. J., Treutlein, H., Wilks, A. F. and Rossjohn, J. (2006) The structural basis of Janus kinase 2 inhibition by a potent and specific pan-Janus kinase inhibitor. Blood 107, 176–183.
54 Nolen, B., Taylor, S. and Ghosh, G. (2004) Regulation of protein kinases: controlling activity through activation segment conformation. Mol. Cell 15, 661–675.
55 Hu, J. and Liu, J. (2003) Structural basis for the recruitment of the adaptor protein APS to the activated insulin receptor. Mol. Cell 12, 1379–1389.
56 Furdui, C. M., Lew, E. D., Schlessinger, J. and Anderson, K. S. (2006) Autophosphrylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. Mol. Cell 21, 711–717.
57 Stamos, J., Sliwkowski, M. X. and Eigenbrot, C. (2002) Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J. Biol. Chem. 277, 46265–46272.
58 Ogawa, A., Takayama, Y., Sakai, H., Chong, K. T., Takeuchi, S., Nakagawa, A., Nada, S., Okada, M. and Tsukihara, T. (2002) Structure of the carboxyl-terminal Src kinase, Csk. J. Biol. Chem. 277, 14351–14354.
59 Gotoh, N., Tojo, A., Hino, M., Yazaki, Y. and Shibuya, M. (1992) A highly conserved tyrosine residue at codon 845 within the kinase domain is not required for the transforming activity of human epidermal growth factor receptor. Biochem. Biophys. Res. Commun. 186, 768–774.
60 Li, F., Gangal, M., Juliano, C., Gorfain, E., Taylor, S. and Johnson, D. A. (2002) Evidence for internal entropy contribution to phosphoryl transfer: a study of domain closure, backbone flexibility, and the catalytic cycle of cAMP-dependent protein kinase. J. Mol. Biol. 315, 459–469.
61 Huse, M. and Kuriyan, J. (2002) The conformational plasticity of protein kinases. Cell 109, 275–282.
62 Xu, W., Doshi, A., Lei, M., Eck, M. J. and Harrison, S. C. (1999) Crystal structures of c-Src reveal features of its autoinhibitory mechanism. Mol. Cell 3, 629–638.

63 Young, M. A., Gonfloni, S., Superti-Furga, G., Roux, B. and Kuriyan, J. (2001) Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation. Cell 105, 115–126.

64 Munshi, S., Kornienko, M., Hall, D. L., Reid, J. C., Waxman, L., Stirdivant, S. M., Darke, P. L. and Kuo, L. C. (2002) Crystal structure of the apo, unphosphorylated insulin-like growth factor-1 receptor kinase. J. Biol. Chem. 277, 38797–38802.

65 Pautsch, A., Zoephel, A., Ahorn, H., Spevac, W., Hauptmann, R. and Nar, H. (2001) Crystal structure of bisphosphorylated IGF-1 receptor kinase: insight into domain movements upon kinase activation. Structure 9, 955–965.

66 Till, J. H., Becerra, M., Watty, A., Lu, Y., Ma, Y., Neubert, T. A., Burden, S. J. and Hubbard, S. R. (2002) Crystal structure of the MuSK tyrosine kinase: insights into receptor autoregulation. Structure 10, 1187–1196.

67 Schindler, T., Bornmann, W., Pellicena, P., Miller, W. T., Clarkson, B. and Kuriyan, J. (2000) Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. Science 298, 1938–1942.

68 Nagar, B., Bornmann, W. G., Pellicena, P., Schindler, T., Veach, D. R., Miller, W. T., Clarkson, B. and Kuriyan, J. (2002) Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and Imatinib (STI-571). Cancer Res. 62, 4236–4243.

69 Nagar, B., Hantschel, O., Young, M. A., Scheffzek, K., Veach, D., Bornmann, W., Clarkson, B., Superti-Furga, G. and Kuriyan, J. (2003) Structural basis for the autoinhibition of c-Abl tyrosine kinase. Cell 112, 859–871.

70 Cowan-Jacob, S. W., Guez, V., Fendrich, G., Griffin, J. D., Fabbro, D., Furet, P., Liebetanz, J., Mestan, J. and Manley, P. W. (2004) Imatinib (STI571) resistance in chronic myelogenous leukemia: molecular basis of the underlying mechanisms and potential strategies for treatment. Mini-Rev. Med. Chem. 4, 285–299.

71 LaFevre-Bernt, M., Sicheri, F., Pico, A., Porter, M., Kuriyan, J. and Miller, W. T. (1998) Intramolecular regulatory interactions in the Src family kinase Hck probed by mutagenesis of a conserved tryptophan residue. J. Biol. Chem. 273, 32129–32134.

72 Cowan-Jacob, S. W., Fendrich, G., Manley, P. W., Jahnke, W., Fabbro, D., Liebetanz, J. and Meyer, T. (2005) The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. Structure 13, 861–871.

73 Banavali, N. K. and Roux, B. (2005) The N-terminal end of the catalytic domain of Src kinase Hck is a conformational switch implicated in long-range allosteric regulation. Structure 13, 1715–1723.

74 Mohammadi, M., Schlessinger, J. and Hubbard, S. R. (1996) Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism. Cell 86, 577–587.

75 Mohammadi, M., McMahon, G., Sun, L., Tang, C., Hirth, P., Yeh, B. K., Hubbard, S. R. and Schlessinger, J. (1997) Structure of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. Science 276, 955–960.

76 Harrison, S. (2003) Variation on a Src-like theme. Cell 112, 737–740.

77 Lerner, E. C. and Smithgall, T. E. (2002) SH3-dependent stimulation of Src-family kinase autophosphorylation without tail release from the SH2 domain *in vivo*. Nature Struct. Biol. 9, 365–369.

78 Hantschel, O., Nagar, B., Guettler, S., Kretzschmar, J., Dorey, K., Kuriyan, J. and Superti-Furga, G. (2003) A myristoyl/phosphotyrosine switch regulates c-Abl. Cell 112, 845–857.

79 Pluk, H., Dorey, K. and Superti-Furga, G. (2002) Autoinhibition of c-Abl. Cell 108, 247–259.

80 Karplus, M. and Kuriyan, J. (2005) Molecular dynamics and protein function. Proc. Natl. Acad. Sci. USA 102, 6679–6685.

81 Nagar, B., Hantschel, O., Seeliger, M., Davies, J. M., Weis, W. I., Superti-Furga, G. and Kuriyan, J. (2006) Organization of the SH3-SH2 unit in active and inactive forms of the c-Abl tyrosine kinase. Mol. Cell 21, 787–798.

82 Resh, M. D. (1994) Myristylation and palmitylation of Src family members: the fats of the matter. Cell 76, 411–413.

83 Sondhi, D. and Cole, P. A. (1999) Domain interactions in protein tyrosine kinase Csk. Biochemistry 38, 11147–11155.

84 Gilliland, D. G. and Griffin, J. D. (2002) The roles of FLT3 in hematopoiesis and leukemia. Blood 100, 1532–1542.

85 Mol, C. D., Dougan, D. R., Schneider, T. R., Skene, R. J., Kraus, M. L., Scheibe, D. N., Snell, G. P., Zou, H., Sang, B.-C. and Wilson, K. P. (2004) Structural basis for the autoinhibition and STI-571 inhibition of cKit tyrosine kinase. J. Biol. Chem. 279, 31655–31663.

86 Li, S., Covino, N. D., Stein, E. G., Till, J. H. and Hubbard, S. R. (2003) Structural and biochemical evidence for an autoinhibitory role for tyrosine 984 in the juxtamembrane region of the insulin receptor. J. Biol. Chem. 278, 26007–26014.

87 Brdicka, T., Kadlacek, T. A., Roose, J. P., Pastuszak, A. W. and Weiss, A. (2005) Intramolecular regulatory switch in ZAP-70: analogy with receptor tyrosine kinases. Mol. Cell. Biol. 25, 4924–4933.

88 Gelkop, S., Gish, G. D., Babichev, Y., Pawson, T. and Isakov, N. (2005) T cell activation-induced CrkII binding to the Zap70 protein tyrosine kinase is mediated by Lck-dependent phosphorylation of Zap70 tyrosine 315. J. Immunol. 175, 8123–8132.

89 Niu, X.-L., Peters, K. G. and Kontos, C. D. (2002) Deletion of the carboxyl terminus of Tie2 enhances kinase activity, signaling, and function. J. Biol. Chem. 277, 31768–31773.

90 Comoglio, P. M. (2001) Pathway specificity for Met signaling. Nat. Cell Biol. 3, E161–E162.

91 Meyer, R. D., Dayanir, V., Majnoun, F. and Rahimi, N. (2002) The presence of a single tyrosine residue at the carboxyl domain of vascular endothelial growth factor receptor-2/FLK-1 regulates its autophosphorylation and activation of signalling molecules. J. Biol. Chem. 277, 27081–27087.

92 Jones, N., Chen, S. H., Sturk, C., Master, Z., Tran, J., Kerbel, R. S. and Dumont, D. J. (2003) A unique autophosphorylation site in Tie2/Tek mediates Dok-R phosphotyrosine binding domain binding and function. Mol. Cell. Biol. 23, 2658–2668.

93 Yokoyama, N., Ischenko, I., Hayman, J. and Miller, W. T. (2005) The C terminus of RON tyrosine kinase plays an autoinhibitory role. J. Biol. Chem. 280, 8893–8900.

94 Chen, M., Cheng, A., Candotti, F., Zhou, Y.-J., Hymel, A., Fasth, A., Notarangelo, L. D. and O'Shea, J. J. (2000) Complex effects of naturally occurring mutations in the JAK3 pseudokinase domain: evidence for interactions between the kinase and pseudokinase domains. Mol. Cell. Biol. 20, 947–956.

95 Saharinen, P., Vihinen, M. and Silvennoinen, O. (2003) Autoinhibition of Jak2 tyrosine kinase is dependent on specific regions in its pseudokinase domain. Mol. Biol. Cell. 14, 1448–1459.

96 Ceccarelli, D. F., Song, H. K., Poy, F., Schaller, M. D. and Eck, M. J. (2006) Crystal structure of the FERM domain of focal adhesion kinase. J. Biol. Chem. 281, 252–259.

97 Li, S., Depetris, R. S., Barford, D., Chernoff, J. and Hubbard, S. (2005) Crystal structure of a complex between protein tyrosine phosphatase 1B and the insulin receptor tyrosine kinase. Structure 13, 1643–1651.

98 Salmeen, A., Andersen, J. N., Myers, M. P., Tonks, N. K. and Barford, D. (2000) Molecular basis for the dephosphorylation of the activation segment of the insulin receptor by protein tyrosine phosphatase 1B. Mol. Cell 6, 1401–1412.

99 Depetris, R. S., Hu, J., Gimpelevich, I., Holt, L. J., Daly, R. J. and Hubbard, S. R. (2005) Structural basis for inhibition of the insulin receptor by the adaptor protein Grb14. Mol. Cell 20, 325–333.

100 Chan, B., Lanyi, A., Song, H. K., Griesbach, J., Simarro-Grande, M., Poy, F., Howie, D., Sumegi, J., Terhorst, C. and Eck, M. J. (2003) SAP couples Fyn to SLAM immune receptors. Nat. Cell Biol. 5, 155–160.

101 Kim, P. W., Sun, Z. Y., Blacklow, S. C., Wagner, G. and Eck, M. J. (2003) A zinc clasp structure tethers Lck to T cell coreceptors CD4 and CD8. Science 301, 1725–1728.

102 Plotnikov, A., Hubbard, S. R., Schlessinger, J. and Mohammadi, M. (2000) Crystal structures of two FGF-FGFR complexes reveal the determinants of ligand-receptor specificity. Cell 101, 413–424.

103 Schlessinger, J., Plotnikov, A., Ibrahami, O. A., Eliseenkova, A. V., Yeh, B. K., Yayon, A., Linhardt, R. J. and Mohammadi, M. (2000) Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. Mol. Cell 6, 743–750.

104 Pellegrini, L., Burke, D. F., von Delft, F., Mulloy, B. and Blundell, T. L. (2000) Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. Nature 407, 1029–1034.

105 Stauber, D. J., DiGabriele, A. D. and Hendrickson, W. A. (2000) Structural interactions of fibroblast growth factor receptor with its ligands. Proc. Natl. Acad. Sci. USA 97, 49–54.

106 Mohammadi, M., Olsen, S. K. And Ibrahimi, O. A. (2005) Structural basis for fibroblast growth factor receptor activation. Cytokine Growth Factor Rev. 16, 107–137.

107 Jiang, X., Gurel, O., Mendiaz, E. A., Stearns, G. W., Clogston, C. L., Lu, H. S., Osslund, T. D., Syed, R. S., Langley, K. E. and Hendrickson, W. A. (2000) Structure of the active core of human stem cell factor and analysis of binding to its receptor Kit. EMBO J. 19, 3192–3203.

108 Ward, C. W. and Garrett, T. P. J. (2004) Structural relationships between the insulin receptor and epidermal growth factor receptor families and other proteins. Curr. Opin. Drug Discov. Dev. 7, 630–638.

109 Garrett, T. P. J., McKern, N. M., Lou, M., Elleman, T. C., Adams, T. E. Lovrecz, G. O., Zhu, H. J., Walker, F., Frenkel, M. J., Hoyne, P. A., Jorissen, R. N., Nice, E. C., Burgess, A. W. and Ward, C. W. (2002) Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor-$\alpha$. Cell 110, 763–773.

110 Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J. H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M. and Yokoyama, S. (2002) Crystal structure of the complex of human epidermal growth factor and receptor extracellular domain. Cell 110, 775–787..

111 Cho, H. S. and Leahy, D. J. (2002) Structure of the extracellular region of HER3 reveals an interdomain tether. Science 297, 1330–1333.

112 Fergusen, K. M., Berger, M. B., Mendrola, J. M., Cho, H. S., Leahy, D. J. and Lemmon, M. A. (2003) EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol. Cell 11, 507–517.

113 Walker, F., Orchard, S. G., Jorisson, R. N., Hall, N. E., Zhang, H.-H., Hoyne, P. A., Adams, T. E., Johns, T. G., Ward, C., Garrett, T. P.J., Zhu, H. J., Nerrie, M., Scott, A. M., Nice, E. C. and Burgess, A. W. (2004) CR1/CR2 interactions modulate the functions of the cell surface epidermal growth factor receptor. J. Biol. Chem. 279, 22387–22398.

114 Cho, H. S., Mason, K., Ramyar, K. X., Stanley, A. M., Gabelli, S. B., Denney, D. W. Jr. and Leahy, D. J. (2003) Stucture of the extracellular region of HER2 alone and in complex with the herceptin Fab. Nature 421, 756–760.

115 Garrett, T. P. J., McKern, N. M., Lou, M., Elleman, T. C., Adams, T. E. Lovrecz, G. O., Kofler, M., Jorissen, R. N., Nice, E. C., Burgess, A. W. and Ward, C. W. (2003) The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors. Mol. Cell 11, 495–505.

116 Weismann, C., Ultsch, M. H., Bass, S. H. and de Vos, A. (1999) Crystal structure of nerve growth factor in complex with the ligand-binding domain of the TrkA receptor. Nature 401, 184–188.

117 Banfield, M. J., Naylor, R. L., Robertson, A. G. S., Allen, S. J., Dawbarn, D. and Brady, L. R. (2001) Specificity in Trk receptor : neutrophin interactions: the crystal structure of TrkB-d5 in complex with neurotrophin-4/5. Structure 9, 1191–1199.

118 Marquez, J. A., Smith, C. I. E., Petoukhov, M. V., Surdo, P. L., Mattsson, P. T., Knekt, M., Westlund, A., Scheffzek, K., Saraste, M. and Svergrun, D. I. (2003) Conformation of full-length Bruton tyrosine kinase (Btk) from synchrotron X-ray solution scattering. EMBO J 22, 4646–4624.

119 Luo, R. Z.-T., Benia, D. R., Fernandes, A., Yip, C. C. and Ottensmeyer, F. P. (1999) Quaternary structure of the insulin-insulin receptor complex. Science 285, 1077–1080.

120 Baumeister, W. (2004) Mapping molecular landscapes inside cells. Biol. Chem. 385, 865–872.

121 Banyay, M., Gilstring, F., Hauzenberger, E., Oefverstedt, L.-G., Eriksson, A. B., Krupp, J. J. and Larsson, O. (2004) Three-dimensional imaging of in situ specimens with low-dose electron tomography to analyze protein conformation. Assay Drug Dev. Tech. 2, 561–567.

122 Hochrein, J. M., Lerner, E. C., Sciavone, A. P., Smithgall, T. E. and Engen, J. R. (2005) An examination of dynamics cross-talk between SH2 and SH3 domains by hydrogen/deuterium exchange and mass spectrometry. Protein Sci. 15, 65–73.

123 Hamuro, Y., Wong, L., Shaffer, J., Kim, J. S., Stranz, D. D., Jennings, P. A., Woods V. L. Jr and Adams, J. A. (2002) Phosphorylation driven motions in the COOH-terminal Src kinase, Csk, revealed through enhanced hydrogen-deuterium exchange and mass spectrometry (DXMS). J. Mol. Biol. 323, 871–881.

124 Wong, L., Lieser, S. A., Miyashita, O., Miller, M., Tasken, K., Onuchic, J. N., Adams, J. A., Woods, V. L. and Jennings, P. A. (2005) Coupled motions in the SH2 and kinase domains of Csk control Src phosphorylation. J. Mol. Biol. 351, 131–143.

125 Li, T., Talvenheimo, J., Zeni, L., Rosenfeld, R., Stearns, G. and Arakawa, T. (2002) Changes in protein conformation and dynamics upon complex formation of brain-derived neurotrophic factor and its receptor: investigation by isotope-edited Fourier transform IR spectroscopy. Biopolymers 67, 10–19.

126 Mendez, R., Leplae, R., Lensink, M. F. and Wodak, S. (2005) Assessment of CAPRI predictions in rounds 3–5 shows progress in docking procedures. Proteins Struct. Funct. Bioinf. 60, 150–169.

127 Russel, R. B., Alber, F., Aloy, P., Davis, F. P., Korkin, D., Pichaud, M., Topf, M. and Sali, A. (2004) A structural perspective on protein-protein interactions. Curr. Opin. Struct. Biol. 14, 313–324.

128 Noble, M. E. M., Endicott, J. A. and Johnson, L. N. (2004) Protein kinase inhibitors: insights into drug design from structure. Science 303, 1800–1805.

129 Rees, D. C., Congreve, M., Murray, C. W. and Carr, R. (2004) Fragment-based lead discovery. Nature Rev. Drug Discov. 3, 660–672.

130 Von Ahsen, O. and Böhmer, U. (2005) High-throughput screening for kinase inhibitors. Chem. Bio. Chem. 6, 481–490.

131 Gill, A., Cleasby, A. and Jhoti, H. (2005) The discovery of novel protein kinase inhibitors by using fragment-based high-throughput crystallography. Chem. Bio. Chem. 6, 506–512.

132 Burley, S. K. (2005) Structure-guided discovery for protein kinases using fragment-based lead identification/lead optimization (abstract). ACS national meeting, MEDI-466.

133 Hann, M. M., Leach, A. R. and Harper, G. (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. J. Chem. Inf. Comput. Sci. 41, 856–864.

134 Teague, S. J. (2003) Implications of protein flexibility for drug discovery. Nature Rev. Drug Discov. 2, 527–541.

135 Cavasotto, C. N. and Abagayan, R. A. (2004) Protein flexibility in ligand docking and virtual screening to protein kinases. J. Mol. Biol. 337, 209–225.

136 Thomas, M. P., McInnes, C. and Fischer, P. M. (2006) Protein structures in virtual screening: a case study with CDK2. J. Med. Chem. 49, 92–104.

137 Stoermer, M. J. (2006) Current status of virtual screening as analysed by target class. Med. Chem. 2, 89–112.

138 Adrian, F. J., Ding, Q., Sim, T., Velentza, A., Sloan, C., Liu, Y., Zhang, G., Hur, W., Ding, S., Manley, P. W., Mestan, J., Fabbro, D. and Gray, N. S. (2006) Allosteric inhibitors of Bcr-abl-dependent cell proliferation. Nat. Chem. Biol. 2, 95–102.

139 Parang, K., Till, J. H., Ablooglu, A. J., Kohanski, R. A., Hubbard, S. R. and Cole, P. A. (2001) Mechanism-based design of a protein kinase inhibitor. Nat. Struct. Biol. 8, 37–41.

140 Keri, G., Orfi, L., Eros, D., Hegymegi-Barakonyi, B., Szantai-Kis, C., Horvath, Z., Waczek F, Marosfalvi, J., Szabadkai, I., Pato, J., Greff, Z., Hafenbradl, D., Daub, H., Muller, G., Klebl, B. and Ullrich, A. (2006) Signal transduction therapy with rationally designed kinase inhibitors. Curr. Signal Transduct Ther. 1, 67–95.

141 Collins, I. and Workman, P. (2006) Design and development of signal transduction inhibitors for cancer treatment: experience and challenges with kinase targets. Curr. Signal Transduct Ther. 1, 13–23.

142 Hahn, E. A., Glendenning, A., Sorensen, M. V., Hudgens, S. A., Druker, B. J., Guilhot, F., Larson, R. A., O'Brien, S. G., Dobrez, D. G., Hensley, M. L. and Cella, D. (2003) Quality of life in patients with newly diagnosed chronic phase chronic myeloid leukemia on imatinib versus interferon alfa plus low-dose cytarabine: results from the IRIS study. J. Clin. Oncol. 21, 2138–2146.

143 Gorre, M. E,, Mohammed, M., Ellwood, K., Hsu, N., Paquette, R., Rao, P. N. and Sawyers, C. L. (2001) Clinical resistance to STI-571 cancer therapy caused by *BCR-ABL* gene mutation or amplification. Science 293, 876–880.

144 Shah, N. P., Nicoll, J. M., Nagar, B., Gorre, M. E., Paquette, R. L., Kuriyan, J. and Sawyers, C. L. (2002) Multiple *BCR-ABL* kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast chrisis chronic myeloid leukemia. Cancer Cell 2, 117–125.

145 Roche-Lestienne, C., Soenen-Cornu, V., Grardel-Duflos, N., Lai, J.-L., Phillippe, N., Facon, T., Fenaux, P. and Preudhomme, C. (2002) Several types of mutations of the *Abl* gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. Blood 100, 1014–1018.

146 Cools, J., Maertens, C. and Marynen, P. (2005) Resistance to tyrosine kinase inhibitors: calling on extra forces. Drug Resist. Updates 8, 119–129.

147 Kobayashi, S., Boggon, T. J., Dayaram, T., Janne, P. A., Kocher, O., Meyerson, M., Johnson, B. E., Eck, M. J., Tenen, D. G. and Halmos, B. (2005) EGFR mutation and resistance of non-small cell lung cancer to gefitinib. N. Engl. J. Med. 352, 786–792.

148 Pao, W., Miller, V. A., Politi, K. A., Riely, G. J., Somwar, R., Zakowski, M. F., Kris, M. G. and Varmus, H. (2005) Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. PLoS. Med. 2, 1–11.

149 Young, M. A. Shah, N. P., Chao, L. H., Seeliger, M., Milanov, Z. V., Biggs, W. H., Treiber, D. K., Patel, H. K., Zarrinkar, P. P., Lockhart, D. J., Sawyers, C. L. and Kuriyan, J. (2006) Structure of the kinase domain of an imatinib-resistant Abl mutant in complex with the Aurora kinase inhibitor VX-680. Cancer Res. 66, 1007–1014.

150 Azam, M., Latek, R. R. and Daley, G. Q. (2003) Mechanisms of autoinhibition and STI-571/imatinib resistance revealed by mutagenesis of *BCR-ABL*. Cell 112, 831–843..

151 Kantarjian, H. M., Giles, F., Wunderle, L., Bhalla, K., O'Brien, S., Wassman, B., Tanaka, C., Manley, P., Rae, P., Mietlowski, W., Bochinski, K., Hochhaus, A., Griffin, J. D., Hoelzer, D., Albitar, M., Dugan, M., Cortes, J., Alland, L. and Ottmann, O. G. (2006) Nilotinib in imatinib-resistant CML and Philadelphia chromosome-positive ALL. N. Engl. J. Med. 354, 2542–2551.

152 Talpaz, M., Shah, N. P., Kantarjian, H., Donato, N., Nicoll, J., Paquette, R., Cortes, J., O'Brien, S., Nicaise, C., Bleickardt, E., Blackwood-Chirchir, M. A., Iyer, V., Chen, T.-T., Huang, F., Decillis, A. P. and Sawyers, C. L. (2005) Dasatinib in imatinib-resistant Philadelphia chromosome-positive leukemias. N. Engl. J. Med. 354, 2531–2541.

153 O'Hare, T., Walters, D. K., Stoffregen, E. P., Taiping, J., Manley, P. W., Mestan, J., Cowan-Jacob, S. W., Lee, F. Y., Heinrich, M. C., Deininger, M. W.N. and Druker, B. J. (2005) *In vitro* activity of Bcr-Abl inhibitors AMN107 and BMS-354825 against clinically relevant imatinib-resistant Abl kinase domain mutants. Cancer Res. 65, 4500–4505.

154 Weisberg, E., Manley, P. W., Mestan, J., Cowan-Jacob, S. W., Ray, A. and Griffin, J. (2006). AMN107 (nilotinib): a novel and selective inhibitor of BCR-ABL. Br. J. Cancer 94, 1765–1769.

155 Burgess, M. R., Skaggs, B. J., Shah, N. P., Lee, F. Y. and Sawyers, C. L. (2005) Comparative analysis of two clinically active BCR-ABL kinase inhibitors reveals the role of conformation-specific binding in resistance. Proc. Natl. Acad. Sci. USA 102, 3395–3400.

156 Ray, A., Cowan-Jacob, S. W., Manley, P. W., Mestan, J. and Griffin, J. D. (2006) Identification of Bcr/Abl point mutations conferring resistance to the Abl kinase inhibitor AMN107 by a random mutagenesis study. Blood (epub ahead of print).