Perspective

Knowledge Based Prediction of Ligand Binding Modes and Rational Inhibitor Design for Kinase Drug Discovery

Arup K. Ghose,*,[§] Torsten Herbertz,[§] Douglas A. Pippin, Joseph M. Salvino, and John P. Mallamo

Department of Medicinal Chemistry, Cephalon, Inc., 145 Brandywine Parkway, West Chester, Pennsylvania 19380

Introduction

Drug discovery is a complex and lengthy endeavor.[1] Many strategies exist to accelerate target-to-clinical candidate selection, as well as to provide the highest quality candidate. Several lead finding strategies include the use of accumulated information for ligands of previously executed discovery programs, such as pharmacophore modeling,[2,3] QSAR,[4] 3D-QSAR,[5,6] or the use of target protein structure in structure-aided drug design.[7] A common and successful strategy for ligand design is to employ a knowledge based approach, utilizing approved or clinically advanced compounds for rational ligand design. Many analyses have been carried out on late-stage clinical candidates and marketed drugs to gain insight into their distinguishing and favorable characteristics. An attempt is usually made to correlate in vivo performance with ligand structures and measured or calculated physicochemical properties. Terms such as leadlike and druglike have been used to describe beneficial structural features and physicochemical property ranges of these molecules.[8] The current article describes the use of Protein Data Bank (PDB) entries specifically for kinase inhibitor drug discovery.[9–13] The PDB contains information relating ligand binding in the context of a receptor[14] with protein kinases representing one of the largest known families of enzymes.[15] Kinases are responsible for the phosphorylation of tyrosine, threonine, and serine residues in proteins, part of the process that regulates many signal transduction pathways in cells,[16] thereby playing an important role in cell growth, metabolism, differentiation, and apoptosis. In the past 15 years the pharmaceutical industry has applied a large portion of their resources toward kinase inhibitor discovery, with the result that several kinase inhibitors have been approved as drugs since 2001. The commercial success of imatinib (*N*-(4-methyl-3-(4-(pyridin-3-yl)pyrimidin-2-ylamino)phenyl)-4-((4-methylpiperazin-1-l)methyl)benzamide) has provided a clear incentive for the continued pursuit of kinase inhibitor design.

Current State of Structure-Aided Ligand Design

Drug discovery targeting various kinases benefits from the availability of a large number of crystal structures of protein−ligand complexes and structure-aided ligand design tools of increasing sophistication. In an ideal world, one would be able to employ such tools and existing information to correctly predict the binding mode and affinity of a proposed ligand; however, the consensus is that automated docking programs and scoring functions perform far from ideal.[17,18] Available docking algorithms are currently more successful in predicting correct binding modes of ligand molecules than correctly estimating a binding affinity for the generated poses (which could be used to rank-order said ligands). In cases in which a ligand is docked back into the receptor from which it was extracted, the experimentally observed binding mode may often not receive the highest score, although it frequently can be found among the top scores. One major limitation in docking and scoring studies is the approximation of the receptor as a rigid protein structure, when in fact it is flexible and changes shape during complex formation. Although many approaches are aimed at correcting the known limitations of docking, including scaling of ligand and receptor atom radii,[19–21] induced-fit docking,[22] ensemble docking,[23,24] limited simulation,[7,25,26] and MM/PBSA.[27–29] However, the use of a generalized scoring functions, marginal protein plasticity,[30,31] and incomplete accounting of solvation

---

* To whom correspondence should be addressed. Phone: (610) 738-6882. Fax: (610) 422-0325. E-mail: aghose@cephalon.com.

§ These two authors contributed equally.
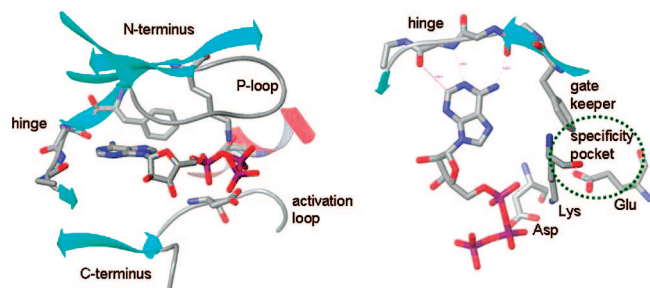
Figure 1. Typical picture of a kinase ATP binding site, illustrating structure and nomenclature. Left: ATP is bound in the cleft between the N-terminus and C-terminus. The P-loop (also known as the glycine-rich loop) forms the roof, and a C-terminal β-sheet covers the floor. Right: The hinge region provides the most important hydrogen bonds with the ATP site inhibitors. Gatekeeper and conserved lysine, glutamate (of helix C), and aspartate (of the DFG triplet) residues control access to specificity pocket (approximate location indicated by green dashed ellipsoid).

effects[32] have not yet overcome the disparity between prediction and experiment.[22]

If proteins change their structures in an apparently unpredictable way, the utility of structure-aided drug design appears to be limited. However, in reality the situation may not be that grave for kinases, as most parts of the ATP binding site show limited movement (except for the most flexible protein features such as the activation loop and to a lesser extent the P-loop; see Figure 1). As a consequence, through an analysis of the binding modes of different types of kinase ligands, one can anticipate the binding mode of a related ligand with "kinase inhibitor features". Such an analysis is extremely useful in deciding, for instance, where to place a hydrogen acceptor or donor atom, install a solubilizing group, or include a hydrophobic modification.

Kinase ATP Binding Site as a Target for Ligand Design

Most protein kinases have a common fold consisting of two lobes, the N-terminus lobe consisting of five antiparallel β-strands and one helix, and the C-terminus lobe, which is highly helical (Figure 1).

The ATP binding site is a narrow hydrophobic pocket located between the two lobes which are linked by a flexible hinge region. The hinge region usually has one hydrogen donor flanked by two hydrogen acceptors derived from the protein backbone. A schematic 2D representation of the ATP binding site is shown in Figure 2.

Kinases have an activation loop containing serine, threonine, or tyrosine residues, which may be phosphorylated. This activation loop occupies a part of the ATP binding site when these residues are not phosphorylated. After activating-phosphorylation by another (upstream) kinase, this hydrophilic, charged loop moves into the solvent to expose the ATP binding site, allowing ATP to bind and transfer its phosphate to a substrate that often may be a different (downstream) kinase. The N-terminal side of the activation loop consisting of a highly conserved triplet comprising aspartate, phenylalanine, and glycine residues, is the DFG motif. The aspartate residue (D) is catalytically involved in the phosphate transfer and typically engages in a salt bridge interaction with a conserved lysine residue. The top (roof) of the ATP binding site is formed by a loop, known as the glycine-rich loop for its highly conserved GXGXXG motif; this loop is alternatively designated the G-loop or P-loop. Clearly, the single most important residue in the ATP binding site is the gatekeeper (gk) residue. The peptide bonds
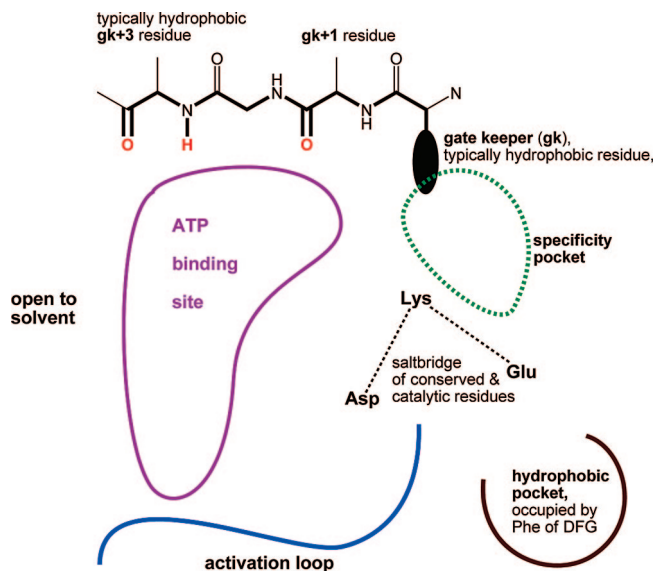
Figure 2. A 2D representation of the kinase binding site. The north side is occupied by the hinge. The northeast is occupied by the gatekeeper. The mobile activation loop occupies the southern area. The catalytic salt bridge occupies the southeast. The west side is the solvent channel. Except for the hinge region and the catalytic salt bridge regions, the kinase ATP site is rather hydrophobic.

forming the hydrogen bond acceptor motifs in the hinge region are thusly referenced as gk+1 and gk+3 relative to the position of the gatekeeper. The size/volume of the gatekeeper's side chain dictates access to the hydrophobic pocket located behind the gatekeeper, thereby defining the potential inhibitor selectivity of the ATP site. This, however, does not mean that targeting this site with ligand−protein interaction is the only way to enhance inhibitor selectivity.

An important form of a kinase is one in which the catalytic aspartate of the DFG triplet is in a rotated-out conformation, known as DFG-out.[33] In this conformation, the catalytically important aspartate (D) rotates out of the ATP binding pocket, accompanied by corresponding rotation of the adjacent phenylalanine (F) and glycine (G) residues. The Phe movement reveals a large hydrophobic pocket, which may be utilized by complementary groups in the ligand. Important drugs such as imatinib target this DFG-out conformation, emphasizing the potential importance of this conformational state in ligand design. Conversely, other inhibitors are being developed, or have been approved, that do not engage macromolecular features beyond the native, preformed ATP binding site.

An understanding of the ATP binding site compressed into 2D is very helpful, since medicinal chemists do and will continue to use simple sketched structures, either on paper or a computer, to conceive of and design new inhibitors. It is therefore very practical to introduce a simplified 2D schematic of the active site (see Figure 2) for the purpose of visualizing the binding poses of various kinase inhibitors.

The high degree of structural similarity of the ATP binding site and the presence of key interactions sites for ligand recognition lend itself to an informatics-driven analysis of protein−ligand interactions across the kinome.

Preparation of the Knowledge Base. Kinase Structure Queries and Extraction from the Protein Data Bank

Databases of aligned kinase ligand−receptor structures are available from commercial vendors (i.e., Eidogen-Sartenty [www.eidogen-sartenty.com]) or may be generated from primary

data as outlined in the workflow below. The "Advanced Search" query tools found on the "Search" tab of the Protein Data Bank (www.rcsb.org) Web page were used to locate kinase structures within the PDB. The query type "EC Number" under Biology & Chemistry was searched using "2.7" as query string (EC 2.7 is the "Transferring phosphorus-containing groups" subclass in the enzyme nomenclature classification scheme (http://www.chem.qmul.ac.uk/iubmb/enzyme/); classification of tyrosine, serine, and threonine protein kinases has changed over the years, and this broader query presumably combines all legacy classes (such as 2.7.1.37, 2.7.1−,...). This query yields greater than 2700 hits, all of which can be downloaded as individual pdb files. A PipeLine Pilot (www.accelrys.com) protocol employing PDB readers and text analysis tools searching for kinase-relevant text strings (such as "serine*kinase" or "tyrosine*kinase" or "threonine*kinase" or "protein*kinase" or "serine*threonine*kinase" or "2.7.1.37" or "2.7.1.−") eliminated most of those files not containing kinase domain structural coordinates. The list of the kinases and the corresponding PDB files studied in this work are shown in Table 1. There were 409 distinct structures studied that were derived from 82 different kinases.

Alignment of Protein Structures

All qualifying files from the previous step were read into Maestro (www.schroedinger.com) and examined visually to determine if the coordinate set contains a fully solved kinase domain. Entries were then prepared in the following manner: delete all water molecules, delete protein chains that are not part of the kinase domain containing chain, delete duplicate chains that may occur but have the same fold, delete protein features extending far beyond kinase domain (i.e., SH2 homology domains). Files in pdb format thus prepared were aligned into the reference frame of chain A of 1aq1.pdb using shell scripts and the command line version of *structalign* to align 10−15 files at a time (this process speeds up the alignment of 730+ protein structures requiring only about 10 min using four processors). Rotated pdb files were read into Maestro, and subsequently all protein features were deleted, bond orders were assigned for each ligand, and tautomeric forms were checked in the context of the original receptor hinge binding interactions and compared to the ligand structure database in the PDB. All ligands were exported into a SDF file for subsequent classification and inclusion into a substructure searchable ISIS database.

Structural Classification of Ligands

The initial structural classification of ligands was carried out with SARVision distributed by Altoris (www.chemapps.com). In this classification scheme the software starts from various common small structural moieties and determines the number of compounds with that substructure (scaffold). The number of hits defines the hierarchy of these scaffolds (Figure 3). The initial classification was further edited to remove duplicates that arose during Maximal Common Substructure (MCS) decomposition of fused ring systems or ring assemblies, e.g., purine is a sub-branch found on both pyrimidine and imidazole branches. Small scaffolds without an entire hinge binding motif were removed from the list. Once a scaffold was accepted for binding mode analysis, all hierarchically lower scaffolds were deleted, unless the lower hierarchy structures had different binding modes or some other special structural features.

Lastly, 2D structure representations of ligands in the kinase reference frame (orientation as in Figure 2) were copied from

the ISIS database into molecular structure fields in Tables 2–4. Where necessary, structures were corrected in ISIS/Draw to remove ambiguities in stereochemistry, ring, and bond orientation; functional groups in question usually involved a solubilizing group or groups engaging the specificity pocket and thus were distal to the hinge binding portion of the ligand.

Analysis of Ligands and Binding Modes

The kinase ligands studied here are broadly divided into three different classes: (i) pyrimidines; (ii) "miscellaneous"; (iii) DFG-out ligands. With ATP being a pyrimidine derivative, understandably there are a large number of kinase ligands in the PDB with a pyrimidine substructure; it is noteworthy that many approved drugs belonged to this class. The "miscellaneous" class comprises diverse chemotypes structurally unrelated to pyrimidines which nevertheless bind in similar fashion within the ATP binding pocket. DFG-out ligands as a class are diverse in structure and observed binding modes, which warrants a separate discussion.

Binding Modes of Different Pyrimidine-Related Ligand Classes

Kinase inhibitors with a pyrimidine substructure are ubiquitous in kinase inhibitor databases (www.eidogen-sertanty.com/products_kinasekb.html). This collection of all hinge binding pyrimidines is summarized in Table 2; a few important binding modes are discussed in Figure 4.

Adenosine/Adenine Analogues. The most frequent hinge binding scaffold found in the kinase inhibitor complexes within the PDB is adenosine. There are 89 kinase structures of 40 different kinases in the PDB complexed to adenosine derivatives. With only a few exceptions (1yxu, 2bzk), all adenosine derivatives have remarkably similar binding modes as exemplified by scaffold **1** in Table 2. Here, the NH_2 and N1 of the purine are involved in hinge binding with NH_2 to the gk+1 residue (Figure 2).[34] Only two exceptions occur, both in PIM1 kinase. As shown in PDB ID 1yxu, the NH_2 group of adenosine binds with the gk+3 residue and N7 of the purine is positioned toward the gatekeeper of the hinge (Table 2, **1a**). It is reasonable to encounter different binding modes for PIM1 bound ligands because the only hydrogen bond donor residue (gk+3, Figure 2) is missing with a proline in this position.[35] A second exception and distinct binding mode is observed for the PIM1 kinase structure PDB ID 2bzk (Table 2, **1b**), in which the NH_2 group forms a hydrogen bond with gk+1 and N7 is oriented toward the gk+3 residue. Finally, it is interesting that despite proline's missing hydrogen bond donor capability, there exists one PIM1 structure (1xr1) with the adenosine oriented in a "conventional" fashion.

9*H*-Purin-2-ylamine. There are 15 2-anilinopurine or 2-aminopurine structures bound to CDK2 (Table 2, **2**). All of these structures are very close chemical analogues with similar binding modes employing three hydrogen bonds at the hinge.[36] A large substituent in the 8-position flipped the ring, as is observed in one structure (Table 2, **2a**). The p38 structure bound to an anilinopurine is very different, since the imidazolo N9 hydrogen is replaced by an ethyl group. Here, the 2-anilino and N1 bind the hinge region (gk+3) with two hydrogen bonds (Table 2, **3**) with the anilino group on the solvent side.[37]

9-Methyl-*N*-phenylpurine-2,8-diamine. There are seven entries with this structural classification bound to six different kinases (Table 2, **4**). The ligands have the same binding mode where the aniline-NH and the N7 of purine are bound to the

Table 1. Different Kinases and the Corresponding PDB Files Studied in the Current Analysis (Organism is *H. sapiens* unless Otherwise Indicated)[a]

gene	organism	sequence description	PDB filename[b]
ABL		Abl tyrosine kinase	2g2h_dfgout; 2g2f; 2g2i, 2hz0_dfgout; 2hzi_dfgout
ABL1	*Mus musculus, Homo sapiens*	proto-oncogene tyrosine-protein kinase ABL	1fpu_dfgout; 1m52_dfgout; 1opj_dfgout; 1opk_dfgout; 1opl_dfgout; 2g1t; 2gqg, 2hiw_dfgout
ACK		activated Cdc42 Kinase 1	1u4d; 1u54
AKT2		Rac-β serine/threonine protein kinase	1o6k; 1o6l
AURKB	*Xenopus laevis*	Loc398457 protein; synonym: Aurora-B	2bfy
BRAF		B-Raf proto-oncogene serine/threonine-protein kinase	1uwh_dfgout; 1uwj_dfgout; 3c4c_dfgout
c-ABL		proto-oncogene tyrosine-protein kinase ABL1 (1B isoform)	1iep_dfgout; 2fo0_dfgout
c-FMS		colony-stimulating factor-1	2i0v_dfgout
CDK2		cyclin-dependent protein kinase 2; cell division protein kinase 2	1aq1; 1b38; 1b39; 1ckp; 1di8; 1dm2; 1e1v; 1e1x; 1e9h; 1fin; 1fq1; 1fvt; 1fvv; 1g5s; 1gih; 1gii; 1gij; 1gy3; 1gz8; 1h00; 1h01; 1h06; 1h07; 1h08; 1h0u; 1h0v; 1h0w; 1h1p; 1h1q; 1h1r; 1h1s; 1hck; 1jst; 1jsv; 1jvp; 1ke5; 1ke6; 1ke7; 1ke8; 1ke9; 1ogu; 1oi9; 1oiq; 1oir; 1oit; 1oiu; 1oiy; 1p2a; 1p5e; 1pf8; 1pkd; 1pxi; 1pxj; 1pxk; 1pxl; 1pxm; 1pxn; 1pxo; 1pxp; 1pye; 1qmz; 1r78; 1urw; 1v1k; 1vyw; 1vyz; 1w0x; 1w8c; 1wcc; 1y8y; 1y91; 2a4l; 2b52; 2b53; 2b54; 2b55; 2bhe; 2bhh; 2bkz; 2bpm; 2btr; 2bts; 2c4g; 2c5n; 2c5o; 2c5p; 2c5t; 2c5v; 2c5x; 2c5y; 2c6i; 2c6k; 2c6l; 2c6m; 2c6o; 2c6t; 2cch; 2cjm; 2clx; 2fvd; 2g9x; 2i40; 2iw6; 2iw8; 2iw9
CDK5		cell division protein kinase 5; synonym: τ protein kinase Ii, cyclin-dependent kinase 5	1ung; 1unh; 1unl
CDK6	*Saimiriine herpesvirus* 2	cell division protein kinase 6	1xo2; 2euf; 2f2c
CDK7		cell division protein kinase 7	1ua2
CHK1		serine/threonine- checkpoint protein kinase	1nvq; 1nvr; 1nvs; 1zlt; 1zys; 2br1; 2brb; 2brg; 2brh; 2brm; 2brn; 2bro; 2c3j; 2c3k; 2c3l; 2cgu; 2cgv; 2cgw; 2cgx
CHK2		serine/threonine-protein kinase Chk2	2cn5; 2cn8
CK17	*Schizosaccharomyces pombe*	casein kinase-1	2csn
CLK1		dual specificity protein kinase Clk1	1z57
CSK		C-terminal Src kinase	1byg
CSNK1	*Schizosaccharomyces pombe*	casein kinase-1	1csn; 1eh4
CSNK1G2		casein kinase 1 γ 2 isoform; synonym: Cki-γ 2	2c47
CSNK2A	*Zea mays*	protein kinase Ck2	1daw; 1day; 1ds5; 1f0q; 1j91; 1lp4; 1lpu; 1lr4; 1m2p; 1m2q; 1m2r; 1om1; 1zoe; 1zog; 1zoh
CSNK2A1		casein kinase Ii, α chain	1jwh; 1pjk; 1ymi
DAPK1		death-associated protein kinase	1ig1; 1jkk; 1jkl; 1p4f; 1wvx; 1wvy
EGFR		epidermal growth factor receptor	1m17; 1xkk; 2ito; 2ity
EPHA2		ephrin type-A receptor 2	1mqb
EPHB2	*Mus musculus*	neural kinase, Nuk = Eph/Elk/Eck family receptor-like tyrosine kinase	1jpa
FAK1		focal adhesion kinase 1	1mp8; 2tm
FGFR1		fibroblast growth factor receptor 1	1agw; 1fgi
FGFR2		fibroblast growth factor receptor 2	1oec
FYN		FYN kinase	2dq7
GCN2	*Saccharomyces cerevisiae*	serine/threonine-protein kinase Gcn2	1zy5; 1zyd
GRK6		G-protein-coupled receptor kinase 6	2acx
GSK3B		glycogen synthase kinase-3 Beta	1j1b; 1j1c; 1o9u; 1pyx; 1q3d; 1q3w; 1q41; 1q4l; 1q5k; 1r0e; 1uv5
HCK		hematopoetic cell kinase Hck; tyrosine-protein kinase Hck	1ad5; 1qcf; 2c0i; 2c0o; 2c0t; 2hk5
IGF1R		insulin-like growth factor 1 kinase	1jqh; 1k3a
INSR		insulin receptor, tyrosine kinase domain	1gag; 1i44; 1ir3; 1irk_dfgout; 1rqq
ITK		tyrosine-protein kinase Itk/Tsk	1sm2; 1snu
JAK2		Janus kinase 2	2b7a
JAK3		Janus kinase 3	1yvj
JNK3		C-JUN N-terminal kinase, also known as MAPK10	See MAPK10.
KDR		vascular endothelial growth factor receptor 2	1y6a; 1y6b; 1ywn_dfgout
KIT		proto-oncogene c-Kit kinase	1pkg; 1t46_dfgout
LCK		lymphocyte-specific protein tyrosine kinase	1qpc; 1qpd; 1qpe; 1qpj; 2ofw_dfgout
MAP2K1		dual specificity mitogen-activated protein kinase kinase 1	1s9j
MAP2K2		dual specificity mitogen-activated protein kinase kinase 2	1s9i
MAPK1	*Rattus norvegicus*	mitogen-activated protein kinase 1	1gol; 1pme; 1tvo; 3erk; 4erk
MAPK10		mitogen-activated protein kinase 10	1jnk; 1pmn; 1pmq; 1pmu; 1pmv; 2b1p
MAPK12		mitogen-activated protein kinase 12	1cm8

Table 1 Continued

gene	organism	sequence description	PDB filename[b]
MAPK14		mitogen-activated protein kinase 14, also know as p38	1a9u; 1bl6; 1bl7; 1bmk; 1di9; 1kv1_dfgout; 1kv2_dfgout; 1m7q; 1ouk; 1ouy; 1ove; 1oz1; 1w7h; 1w82_dfgout; 1w83_dfgout; 1w84; 1wbn_dfgout; 1wbo; 1wbs_dfgout; 1wbt_dfgout; 1wbv_dfgout; 1wbw; 1yqj; 1yw2; 1ywr; 1zyj; 1zz2; 1zzl; 2baj_dfgout; 2bak_dfgout; 2bal; 2baq; 2w; 2gfs; 2ghl; 2ghm; 2gtm; 2gtn; 2i0h
MAPK5		mitogen-activated protein kinase 5	2clq
MAPK8		mitogen-activated protein kinase 8 isoform 4	1uki; 2g01; 2gmx; 2h96
MAPKAPK2		Map kinase-activated protein kinase 2	1nxk; 1ny3
c-MET		hepatocyte growth factor receptor kinase	1r0p; 3c1x_dfgout
NEK2		NIMA-related kinase	2jav
p38		mitogen-activated protein kinase p38; also known as MAPK14.	See MAPK14.
PAK1		p21-activated kinase 1	2hy8
PAK4		p21-activated kinase 4	2cdz
PAK5		p21-activated kinase 5	2f57
PDPK1		3-phosphoinositide dependent protein kinase-1	1h1w; 1oky; 1okz; 1uu3; 1uu7; 1uu8; 1uu9; 1uvr; 1z5m; 2biy
PFPK5	*Plasmodium falciparum*	*Plasmodium falciparum* cdc2-related kinase	1v0o; 1v0p
PHK	*Oryctolagus cuniculus*	phosphorylase kinase	2phk
PHKB	*Oryctolagus cuniculus*	phosphorylase kinase β	1phk; 1ql6
PIM1		proto-oncogene serine/threonine-protein kinase Pim-1	1xr1; 1xws; 1yhs; 1yi3; 1yi4; 1yxt; 1yxu; 1yxv; 1yxx; 2bik; 2bil; 2bzh; 2bzi; 2bzj; 2bzk; 2c3i; 2j2i
PKA	*Mus musculus*	Camp-dependent protein kinase A	2erz; 2f7x; 2f7z; 2gnh; 2gni; 2gnj; 2gnl
PKI-alpha	*Bos taurus*	ρ-kinase	2gnf
PKNB	*Mycobacterium tuberculosis*	serine/threonine kinase from *Mycobacterium tuberculosis*	1mru; 1o6y; 2fum
PRKACA	*Bos taurus, Mus musculus*	Camp-dependent protein kinase, α-catalytic subunit; protein kinase, cAMP-dependent, catalytic, alpha	1atp; 1bkx; 1bx6; 1cdk; 1fmo; 1jbp; 1l3r; 1q24; 1q8t; 1q8u; 1q8w; 1rdq; 1re8; 1rej; 1rek; 1stc; 1sve; 1svg; 1svh; 1szm; 1u7e; 1veb; 1xh4; 1xh6; 1xh7; 1xh8; 1xha; 1ydr; 1yds; 1ydt; 2c1a; 2c1b
PRKCI		protein kinase C, ι	1zrz
PRKCQ		protein kinase C, θ type	1xjd
PRKR	*Saccharomyces cerevisiae*	interferon-induced, double-stranded RNA-activated protein kinase	2a19
RIO1	*Archaeoglobus fulgidus*	Rio1 (Right Open reading frame gene) kinase	1zp9; 1zp9c; 1ztf; 1zth; 1zthc
RIO2	*Archaeoglobus fulgidus* dsm 4304	Rio2 (Right Open reading frame gene) kinase	1tqm; 1tqp; 1zao; 1zar
ROCK I		Rho-Associated Protein Kinase 1	2eto
ROCK2	*Bos taurus*	ρ-associated protein kinase 2	2f2u
SKY1	*Saccharomyces cerevisiae*	serine/arginine-rich protein-specific kinase from budding yeast.	1q8y; 1q97; 1q99
SLK		STE20-like kinase	2j51
SPRK1		serine/arginine- rich protein specific kinase 1	1wbp
SRC Kinase		proto-oncogene tyrosine-protein kinase Src	1ksw; 1y57; 2df; 2dj
STK16		serine/threonine-protein kinase 16	2buj; 2buj
STK6		serine/threonine kinase 6 (Aurora A)	1mq4; 1muo; 1ol5; 1ol6; 1ol7; 2bmc; 2c6e_dfg
SYK		spleen tyrosine kinase	1xbb; 1xbc
TAK1		transforming growth factor β-activated kinase 1	2eva; 2j4o
TAO2	*Rattus norvegicus*	thousand-and-one amino acid kinase 2	1u5r; 2gcd
TGFBR1		transforming growth factor, β receptor 1	1py5; 1rw8; 1vjy
WEE1		Wee1-like protein kinase	1x8b
ZAP70		ζ-chain (TCR) associated protein kinase 70 kDa	1u59

[a] The kinase collection described in this paper was collected in the fourth quarter of 2006. However, we added a few more recent relevant kinase structures to this collection while preparing this manuscript. [b] Filenames were appended with "_dfgout" if the kinase structure was detected as DFG-out.

gk+3 residue. The anilino group occupies the solvent proximal side of the hinge.[38]

6-Benzylamino-9-alkylpurines. There are six ligands in this class bound to four different kinases. Five of these ligands have the same binding mode as is observed in the previous class of purines (Table 2, **5**), where the benzylamino and N7 are bound to the gk+3 residue with the benzyl group on the solvent side.[39] In the SRC kinase structure (Table 2, **5a**), the benzylamino group is located near the gatekeeper. This particular finding is likely due to the gatekeeper being replaced by Gly.[40]

2-Anilino-4-(3-pyridyl)pyrimidines. This is one of the more interesting classes of compounds in which the binding mode of

the ligand can change as a consequence of substitution on the phenyl ring of the anilino group or the kinase family that it is bound to. The first FDA approved kinase inhibitor, imatinib, belongs to this class. The first of the two binding modes involves binding of pyrimidine-N1 and the anilino group from the solvent side to the gk+3 residue (Table 2, **6**).[41] There are two structures, one with SRC and the other with SYK, with this binding mode.

The other binding mode is known as the DFG-out motif. In this motif the pyridine binds at the hinge region with only one hydrogen bond; the benzene ring occupies a newly formed hydrophobic pocket exposed by movement of the DFG triplet (Table 2, **6a**). These structures are discussed separately.
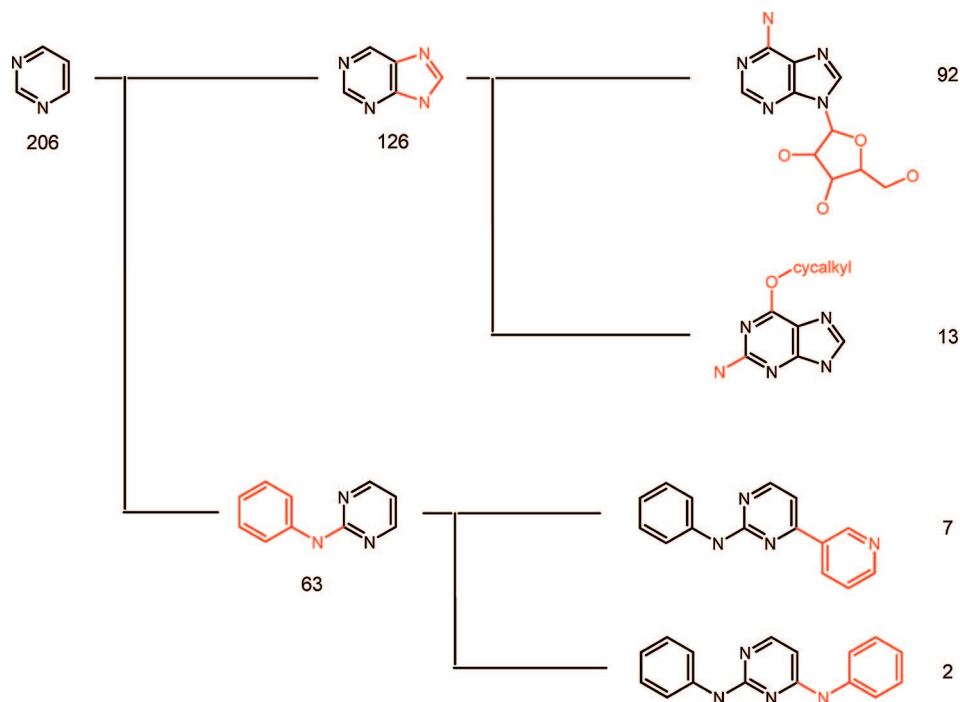
Figure 3. Hierarchical scaffold classification. Pyrimidine is the most common hinge binding scaffold with 206 hits from 409 kinase ligands in the PDB. Among these 206 pyrimidines, there are 126 purines among which 92 are adenosine derivatives.

Adenosine   2-Anilinopurine   2-Anilino-4-(3-pyridyl)-pyrimidine
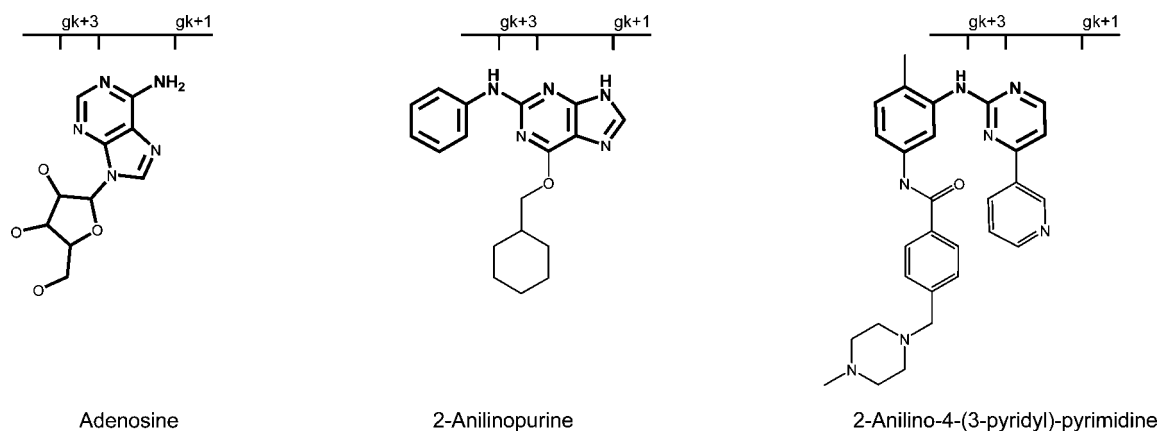
Figure 4. Depiction of relative ligand-hinge orientation of three representative pyrimidine containing kinase ligands. The hinge is depicted by the horizontal bar and three small vertical lines represent the one hydrogen donor flanked by two hydrogen acceptors.

2-Anilino-4-(5-thiazolyl or imidazolyl)pyrimidines. There are 11 structures with the same binding mode, and all are bound to CDK2. The pyrimidine N1 and aniline-NH coordinate with gk+3 of the hinge region (Table 2, **7**).[42]

6-Phenyl-8-methyl-2-phenylaminopyrido[6,5-*d*]pyrimidin-7-one. There are six structures: four with ABL1 and two with ABL. All have the same binding mode where the aniline-NH and pyrimidine-N1 are bound to the gk+3 residue (Table 2, **8**).[43] Interestingly enough, these are also special type DFG-out complex as discussed later.

2,4-Dianilinopyrimidines. There are two structures, both with CDK2 and both having a solubilizing group attached to the 2-anilino group. These compounds are bound with two hydrogen bonds at the hinge region: 2-anilino-NH and N1 of the pyrimidine. The reason these do not bind with three hydrogen bonds (N4 and two anilino-NH groups) may be explained by the presence of an NH with an aryl ring or a large group attached to it that cannot bind to the gk+1 residue without strong steric interference; thus, the 4-anilino group is positioned toward the activation loop (Table 2, **9**).[44]

4,6-Dianilinopyrimidines. There are five entries, all with CDK2. Like 2,4-dianilinopyrimidines, these ligands are bound with two hydrogen bonds at the hinge region. Which of the potentially equivalent anilino NH groups binds the hinge is governed by their respective substitution; the aniline group with the more hydrophilic substitution binds the hinge from the solvent side, directing the other aniline group toward activation loop (Table 2, **10**).[45]

N-Phenyl-[1,2,4]triazolo[5,1-*b*]pyrimidin-7-amine. There are seven entries, and all are bound to CDK2 with the same binding mode. There are two hydrogen bonds with the gk+3 residue: one from the anilino-NH on the solvent side and the other from N2 (H-acceptor) (Table 2, **11**).[46]

N-Phenylquinazolin-4-amine. This is another very important class of inhibitors, since two compounds of this class have been approved for the treatment of certain cancers, erlotinib (*N*-(3-ethynylphenyl)-6,7-bis(2-methoxyethoxy)quinazolin-4-amine) and gefitinib (*N*-(3-ethynylphenyl)-6,7-bis(2-methoxyethoxy)quinazolin-4-amine). There are five such structures with three kinases,

Table 2. Hinge Binding of Purine and Pyrimidine Analogues in PDB Kinase Ligands[a]
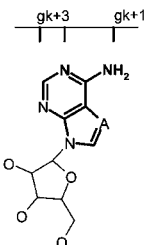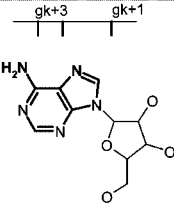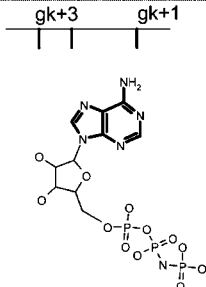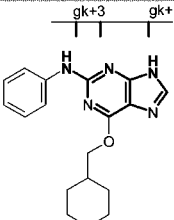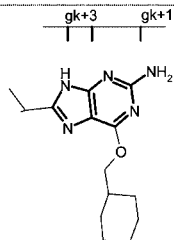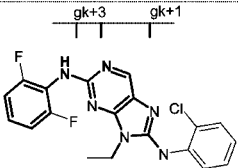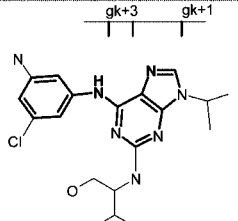
Scaffold	Most Common Binding Mode Illustrated in 2D[‡]	Frequency in PDB[§]	Unique Kinases Entries	Kinases (PDB ID)
1		89	40	ABL(2g2f, 2g2i), ABL1(2g1t), ACK(1u54), AKT2(1o6k, 1o6l), CDK2(1b38, 1b39, 1fin, 1fq1, 1gy3, 1hck, 1jst, 1qmz, 2cch, 2cjm), CDK7(1ua2), CHK2(2cn5), CSNK1(1csn), CSNK1G2(2c47), CSNK2A(1daw, 1ds5, 1lp4), CSNK2A1(1pjk, 1ymi), DAPK1(1ig1, 1jkk, 1jkl), EPHA2(1mqb), EPHB2(1jpa), FAK1(1mp8), GCN2(1zy5, 1zyd), GSK3B(1j1b, 1j1c, 1o9u, 1pyx), HCK(1ad5), IGF1R(1jqh, 1k3a), INSR(1gag, 1i44, 1ir3, 1rqq), KIT(1pkg), LCK(1qpc), MAP2K1(1s9j), MAP2K2(1s9i), MAPK1(1gol), MAPK10(1jnk), MAPK12(1cm8), MAPKAPK2(1ny3), PDPK1(1h1w, 2biy), PHKB(1phk, 1ql6), PIM1(1xr1, 1yi4, 1yxt), PKNB(1mru, 1o6y), PRKACA(1atp, 1bkx, 1cdk, 1fmo, 1jbp, 1l3r, 1q24, 1rdq, 1u7e), PRKR(2a19), RIO1(1zp9, 1ztf, 1zth), RIO2(1tqm, 1tqp, 1zao, 1zar), SKY1(1q8y, 1q97, 1q99), SPRK1(1wbp), STK6(1mq4, 1muo, 1ol5, 1ol6, 1ol7), TAO2(1u5r)
1a		1	1	PIM1(1yxu)
1b		1	1	PIM(2bzk)
2		15	1	CDK2(1e1v, 1gz8, 1h0u, 1h0v, 1h1p, 1h1q, 1h1r, 1h1s, 1oi9, 1oiu, 1oiy, 2c6o, 2g9x, 2iw8, 2iw9)
2a		1	1	CDK2(1w8c)
3		1	1	p38(2gtm)
4		7	6	CDK2(1ckp), CDK6(2f2c), PAK4(2cdz), PAK5(2f57), PFPK5(1v0p), SRCK(2bdf, 2bdj)

Table 2 Continued
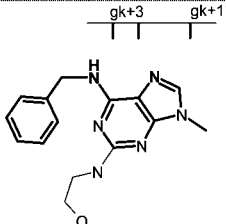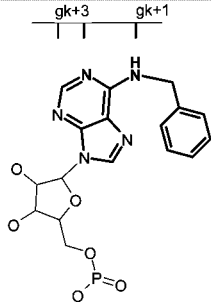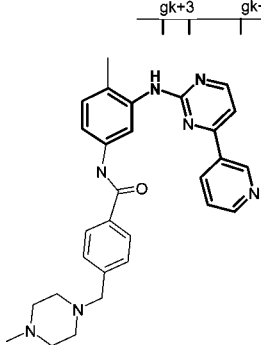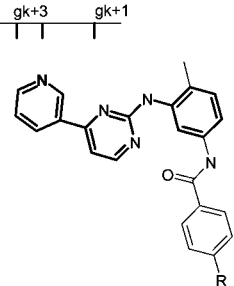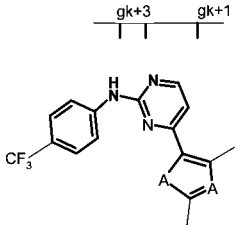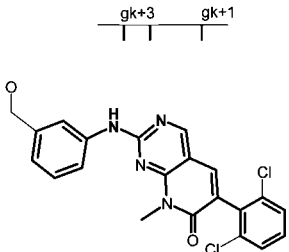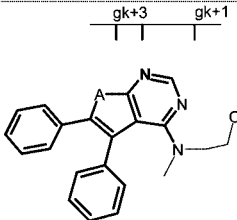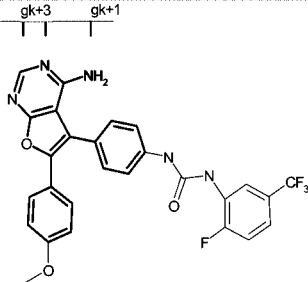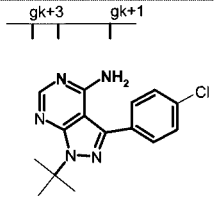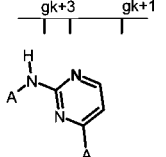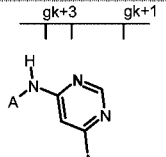
| Scaffold	Most Common Binding Mode Illustrated in 2D[‡]	Frequency in PDB[§]	Unique Kinases Entries	Kinases (PDB ID)
5		5	3	CDK2(1g5s, 1w0x, 2a4l), CDK5(1unl), MAPK1(4erk)
5a		1	1	SRCK(1ksw)
6		2	2	SRC(1y57), SYK(1xbb)
6a		5	3	ABL(2hyy), ABL1(1fpu, 1iep, 1opj), KIT(1t46)
7		11	1	CDK2(1oir, 1oit, 1pxl, 1pxm, 1pxn, 1pxp, 1urw, 2c5n, 2c5p, 2c5t, 2c5v)
8		6	2	ABL(2g2h_dfgout, 2hzi_dfgout), ABL1(1m52_dfgout, 1opk_dfgout, 1opl_dfgout, 2fo0)

Table 2 Continued

Scaffold	Most Common Binding Mode Illustrated in 2D‡	Frequency in PDB§	Unique Kinases Entries	Kinases (PDB ID)
9		2	1	CDK2(1h01, 1h08)
10		5	1	CDK2(1h00, 1h06, 1h07, 1v1k, 2iw6)
11		7	1	CDK2(1y8y, 1y91, 2c6i, 2c6k, 2c6l, 2c6m, 2c6t)
12		5	3	CDK2(1di8), EGFR(1m17, 1xkk), p38(1di9, 2bak),
13		3	2	CDK2(1pxi, 1pxj), MAPK1(3erk),
13a		4	2	CDK2(1pxo, 2c5o), p38(1bl7, 1bmk)
14		1	1	CDK2(2b53)

Table 2 Continued

Scaffold	Most Common Binding Mode Illustrated in 2D[‡]	Frequency in PDB[§]	Unique Kinases Entries	Kinases (PDB ID)
15		6	1	CHK1(2br1, 2br2, 2brb, 2brg, 2brh, 2brm, 2brn),
15a		1	1	KDR(1ywn_dfgout)
16		4	2	HCK(1qcf, 2c0i, 2c0t), LCK(1qpe)
17		47	11	ABL(2g2h_dfgout), ABL1(1m52_dfgout, 1opk_dfgout, 1opl_dfgout, 2fo0), CDK2(1h01, 1h08, 1h1q, 1h1r, 1h1s, 1ogu, 1oi9, 1oiq, 1oir, 1oit, 1oiu, 1oiy, 1pxk, 1pxl, 1pxm, 1pxn, 1pxp, 1urw, 2c5n, 2c5p, 2c5t, 2c5v, 2c5x, 2c5y, 2c6o, 2fvd, 2g9x, 2iw8, 2iw9), CDK6(2euf), FGFR2(1oec), FGFR1(2fgi), FAK1(2etm), MAPK10(1pmn, 1pmq), p38(1ouk, 1ywr, 2ghl, 2ghm, 2gtm), PDPK1(1z5m), SRC(1y57), SYK(1xbb)
18		6	1	CDK2(1h00, 1h06, 1h07, 1jsv, 1v1k, 2iw6)

[a] (‡) Orientation of active site as illustrated in Figure 2; hinge region in the north, gatekeeper in the northeast, solvent in the west, catalytic salt bridge in the southeast and activation loop in the south. (§) Multiple chains of the same protein within a coordinate file were counted as one structure.
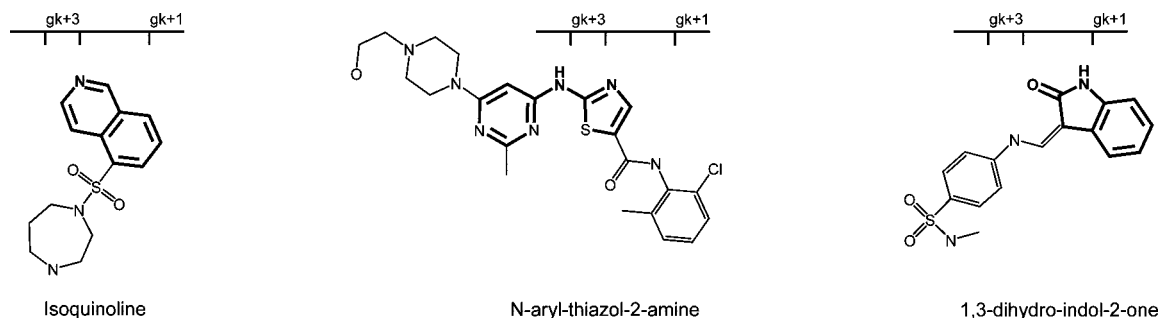
Figure 5. Depiction of relative ligand-hinge orientation of three instructive example non-pyrimidine-based kinase ligands.

CDKZ, EGFR and P38, all with the same binding mode. The compounds bind with only one hydrogen bond in the hinge region employing N1 of the quinazoline ring (Table 2, **12**).[47] The fused benzene ring is oriented toward the solvent side, with the *N*-phenyl toward the Asp-Lys salt bridge. Solubilizing substituents in position 6 or 7 on the quinazoline core are directed toward solvent.

5-(2-Amino-4-pyrimidinyl)imidazole or -thiazole. This is an important class of inhibitors for MAP kinases. As is expected,

the 2-aminopyrimidine moiety binds the hinge region; whenever a 4-phenyl group is attached to the imidazole or thiazole, it points toward the gatekeeper. We divided the binding mode of this class of compounds into two subclasses: (i) where the 2-amino group of the pyrimidine binds to the gk+1 residue (Table 2, **13**)[39] and (ii) where the 2-amino group of the pyrimidine binds to the gk+3 residue (Table 2, **13a**).[42] The hinge binding NH$_2$ group tends to be oriented toward the gatekeeper unless forced away toward solvent by bulky substitu-

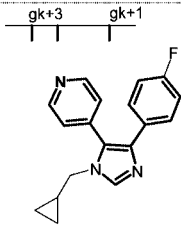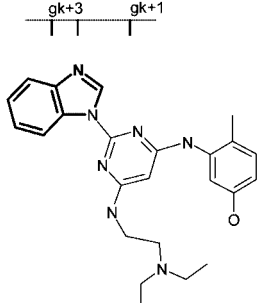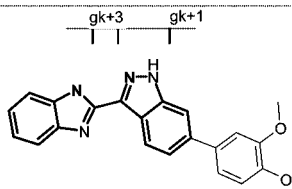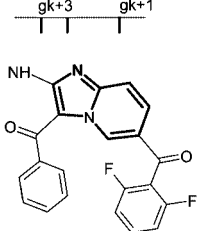**Table 3.** Hinge Binding of Miscellaneous Inhibitors[a]
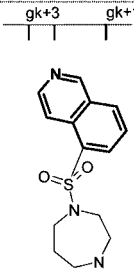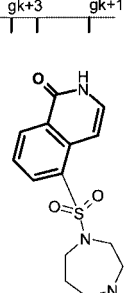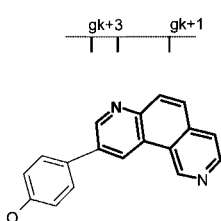
Scaffold	Most Common Binding Mode Illustrated in 2D[**]	Frequency in PDB[††]	Unique Kinases Entries	Kinases (PDB ID)
19		4	2	MAPK1(1pme), p38(1a9u, 1bl6, 2ewa_dfgout)
20		2	2	CDK2(2i40), HCK(2hk5)
21		2	1	CHK1(2c3k, 2c3l)
22		1	1	CDK2(1pye)
23		15	4	CK17(2csn), PKA(2gnh, 2gni, 2gnl), PRKACA(1q8u, 1q8w, 1ydr, 1yds, 1ydt, 2c1a,, 2c1b), ROCK(2esm, 2eto, 2f2u)
24		2	2	PKA(2erz), ROCK(2etk)
25		1	1	JNK3 (1pmu)

Table 3 Continued
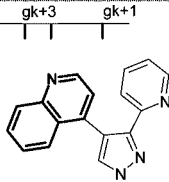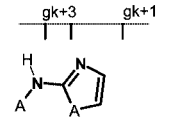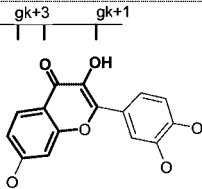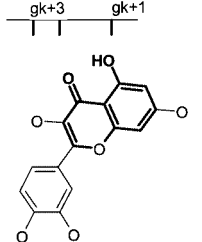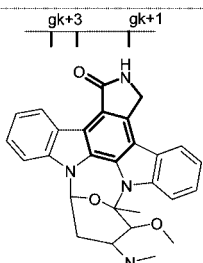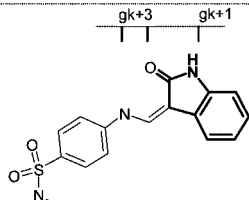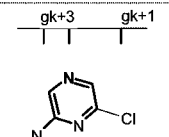
Scaffold	Most Common Binding Mode Illustrated in 2D**	Frequency in PDB ††	Unique Kinases Entries	Kinases (PDB ID)
26		1	1	TGFBR1(1py5)
27		5	3	ABL1(2gqg), CDK2 (2bts), GSK3β(1q5k), KDR(1y6a, 1y6b)
28		1	1	SRCK(1xo2),
28a		1	1	CDK6(2hck)
29		4	1	PRKACA(1bx6, 1re8, 1rej, 1rek)
30		33	22	CDK2(1aq1, 1pkd) CSK(1byg), CHK1(1nvr, 1nvs, 1nvq), DAP1(1wvx, 1wvy) FYN(2dq7), GSK3B(1q3d), ITK(1sm2, 1snu), JAK2(2b7a), JAK3(1yvj), LCK(1qpd,1qpj), MAPK5(2clq), MAPKAPK2(1nxk), c-MET(1r0p), PAK1(2hy8), PDPK1(2bzj,1oky, 1okz), PIM1(2bzh, 1yhs), PRKACA(1stc), PRKCQ(1xjd), STK16(2buj, 2buj), SYK(1xbc), TAO2(2gcd), WEE1(1x8b), ZAP70(1u59)
31		22	9	CDK2(1e9h, 1fvt, 1fvv, 1ke5, 1ke6, 1ke7, 1ke8, 1ke9, 1pf8, 1r78, 2bhe, 2bhh), CDK5(1unh), CHK1(2ayp), CSNK1(1eh4), FGFR1(1agw, 1fgi), GSK3B(1q41, 1uv5), NEK2(2jav), PFPK5(1v0o), PIM1(1yxx)
32		1	1	CDK2(1wcc)

Table 3 Continued
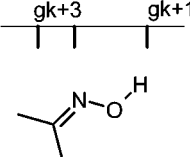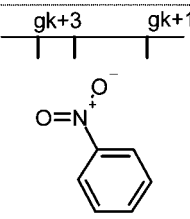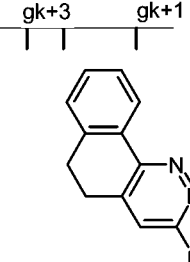
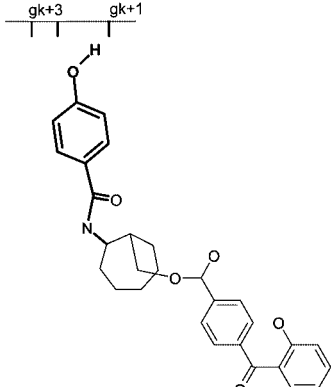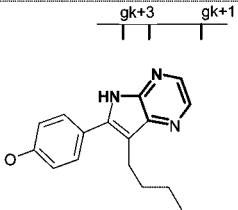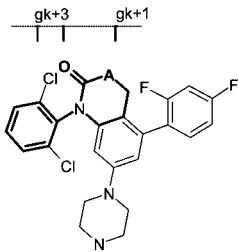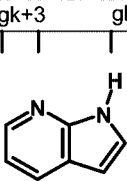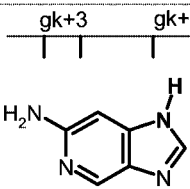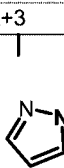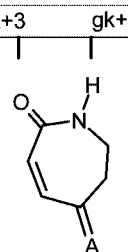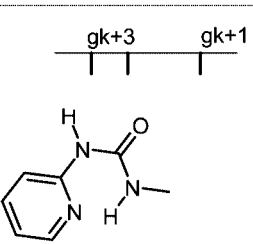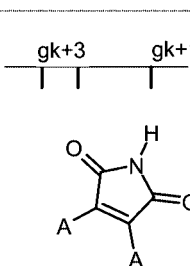| Scaffold	Most Common Binding Mode Illustrated in 2D[**]	Frequency in PDB[††]	Unique Kinases Entries	Kinases (PDB ID)
32a		1	1	CDK5(1ung)
33		2	1	MAP14(1ove, 1m7q)
34		1	1	CHK1(1zys)
35		1	1	CDK2(1h0w)
36		12	3	CDK2(1jvp, 1vyw, 1vyz, 2b52, 2b55, 2bkz, 2bpm, 2c4g), MAPK10(1pmv, 2b1p), MAPK8(1uki), STK6(2bmc)
37		6	5	ACK(1u4d), CDK2(1dm2), CHK1(1zlt, 2c3j), CHK2(2cn8), CLK(1z57), GSK3β(1q3w)
38		2	1	CDK2(1gii, 1gih)
39		11	5	GSK3B(1q4l, 1r0e), PDPK1(1uu3, 1uu7, 1uu8, 1uu9, 1uvr), PIM1(1xws, 2bik, 2bil, 2j2i), PRKACA(1szm), PRKCI(1zrz,),

Table 3 Continued

Scaffold	Most Common Binding Mode Illustrated in 2D**	Frequency in PDB ††	Unique Kinases Entries	Kinases (PDB ID)
40		25	7	ABL(2hzn_dfgout), MAPK1(1pme), p38(1a9u, 1bl6, 1oz1, 1w84, 1wbs_dfgout, 1wbt_dfgout, 1yqj, 2ewa_dfgout), PKA(2f7x, 2f7z, 2gnj), PKI-Alpha(2gnf), PRKACA(1q8t, 1sve, 1svg, 1svh, 1veb, 1xh4, 1xh6, 1xh7, 1xh8, 1xha), ROCK I(2etr)
41		1	1	CHK1(2cgu)
42		2	2	CHK1(2cgv), CSNK2A(1m2q)
43		1	1	DAPK1(1p4f)

[a] (**) Orientation of active site as illustrated in Figure 2; hinge region in the north, gatekeeper in the northeast, solvent in the west, catalytic salt bridge in the southeast and activation loop in the south. (††) Multiple chains of the same protein within a coordinate file were counted as one structure.

tion. What triggers binding of the NH$_2$ to gk+3 is not obvious, since in most other ligands with an analogous NH$_2$ group, it tends to bind to gk+1.

2-Substituted Quinazolines. Although quinazolines were already mentioned as a class, a slightly different ligand pose can be observed for a 2-substituted quinazoline bound to CDK2 (Table 2, **14**). Here, N1 of the quinazoline also binds through a single hydrogen bond to the hinge region; however, the ring orientation is substantially different from 2-unsubstituted quinazolines (vide supra).[48]

5,6-Di(phenyl)pyrrolo[3,2-*e*]pyrimidin-4-amine. Both pyrrolo and furano derivatives of this class (Table 2, **15**) bind the hinge with N1 of the pyrimidine ring unless the pyrrolo-N is substituted.[49] When the pyrrole-N is substituted, N-3 of the pyrimidine and the 4-NH$_2$ are involved in hinge binding (Table 2, **15a**).[50] There are six CHK1 kinase structures with the first binding mode and one KDR structure with the second binding mode. The second one is also a DFG-out ligand. It is interesting to speculate on the possible bindi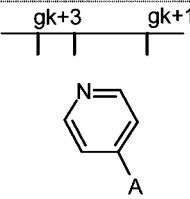ng modes for ligands containing two different available NH functions, pyrrole and 4-NH$_2$ pyrimidine in this scaffold, making two competing hinge-binding orientations conceivable, especially when the DFG-out moiety is missing.

1-Methyl-3-phenylpyrazolo[4,5-*e*]pyrimidin-4-amine. This class has four structures in the PDB (Table 2, **16**). Three are of HCK and one is LCK. They display the same ligand binding mode and are quite similar to the KDR structure of the previous class (Table 2, **16a**) where N3 and the 4-NH$_2$ of the pyrimidine ring bind to the hinge.[51]

N-Substituted-4-substituted-pyrimidin-2-amine and N-Substituted-4-substituted-pyrimidin-2-amine. Although a few substructures that have already been discussed belong to these two more generic substructures (Table 2, **17** and **18**), these additional examples illustrate that there are small variations in these classes of compounds with the same binding mode.

Binding Modes of Other Hinge Binding Scaffold Types

In most of the purine- and pyrimidine-type ligands discussed in Table 2, hinge binding typically involves a hydrogen donor and a hydrogen acceptor separated by one or two atoms in a cis-like arrangement; however, many kinase inhibitors have been discovered that feature other chemotypes. Examples include ligands forming only one hinge interaction, hinge binders made up of ring systems other than pyrimidine, and molecules displaying less common hydrogen bonding moieties. A survey of these "miscellaneous" scaffolds is shown in Table 3, and a few representative binding modes are discussed in Figure 5.

Pyridine-4-(4-phenyl)-1*H*-imidazole-5-yl. These are inhibitors of MAP kinases. There are four such entries, three bound to p38 and one bound to MAPK1. The pyridine-N is bound to the gk+3 residue, and the phenyl group is pointed toward the gatekeeper (Table 3, **19**).[52] It is interesting to note that both p38 and MAPK1 have a small gatekeeper residue (Thr), which may facilitate binding of the phenyl group in that region.

Table 4. Structures and Binding Modes of Various Ligands in PDB, Binding the DFG-Out Conformation of Kinases[a]
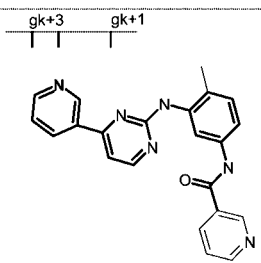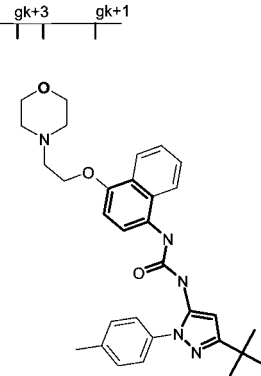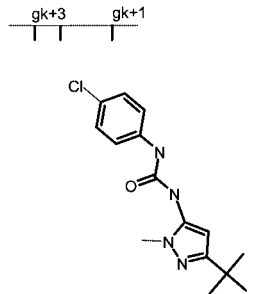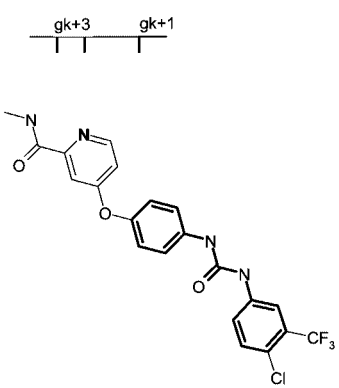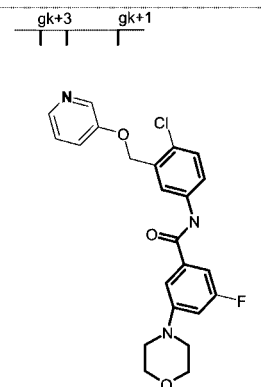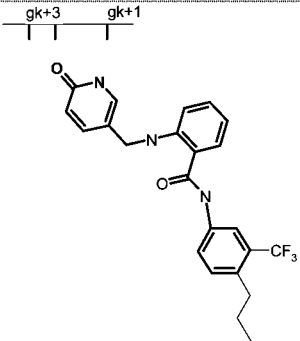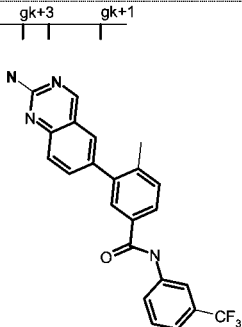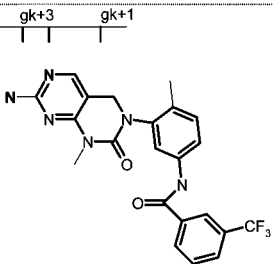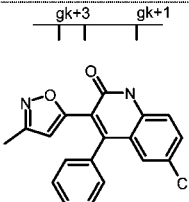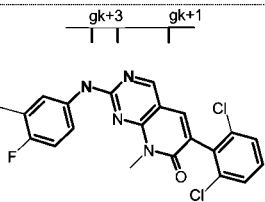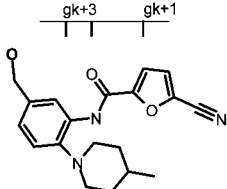
Scaffold	Most Common Binding Mode Illustrated in 2D[‡‡]	Frequency in PDB [§§]	Unique Kinases Entries	Kinases (PDB ID)
44		5	3	ABL(1fpu), c-ABL(1iep, 1opj, 2hyy), c-KIT(1t46)
45		2	1	p38(1kv2, 1wbn)
45a		3	1	p38(1kv1, 1w82, 2baj)
46		4	3	B-RAF(1uwh, 1uwj), VEGFR2(1ywn), c-ABL(2hzn)
47		5	1	MAP14(1w83, 1wbs, 1wbt, 1wbv, 2bak)

Table 4 Continued

Scaffold	Most Common Binding Mode Illustrated in 2D[‡‡]	Frequency in PDB[§§]	Unique Kinases Entries	Kinases (PDB ID)
48		1	1	ABL(2hz0)
49		2	1	LCK(2ofv, 2og8)
50		1	1	ABL(2hiw)
51		1	1	c-FMS(2i0v)
52		3	2	c-ABL(2fo0), ABL(2g2h, 2hzi)
53		2	1	c-FMS(2i0y, 2i1m)

[a] (‡‡) Orientation of active site as illustrated in Figure 2; hinge region in the north, gatekeeper in the northeast, solvent in the west, catalytic salt bridge in the southeast and activation loop in the south. (§§) Multiple chains of the same protein within a coordinate file were counted as one structure.

Benzimidazole. There are two distinct types of hinge binding benzimidazoles: one as a hydrogen acceptor using N3 (Table 3, **20**)[53] and another with its NH as the H-donor (Table 3, **21**).[54]

In both cases, the gk+3 residue is engaged and the benzene moiety of the fused ring points toward solvent. The hydrogen acceptor types bind with only one hydrogen bond and are

observed in two kinases HCK and CDK2. The hydrogen donor types have a pyrazole as the other hinge binding element and were observed in two CHK1 kinase structures.

Imidazo[1,2-*a*]pyridine. There is only one such structure (Table 3, **22**) with CDK2. Here, the imidazole N3 is hydrogen-bonded to gk+3 along with the NH$_2$ group adjacent to it. This structure is somewhat unusual, since the free aromatic amine is on the solvent side.[55]

Isoquinolines. There are several isoquinoline structures (Table 3, **23**) bound to four kinases (CK17, PKA, PRKACA, and ROCK) in the PDB. All of these ligands are related to fasudil (5-(1,4-diazepan-1-ylsulfonyl)isoquinoline), the first kinase drug approved in Japan. All share the same binding mode with one hinge hydrogen bond to the isoquinoline N.[56]

Isoquinolin-1(2*H*)-one. There are two isoquinolinone structures (Table 3, **24**) bound to two kinases ROCK and PKA. Both structures have the same binding mode where the NH and C=O are bound to the gk+1 and gk+3 residues, respectively.[57]

[2,7]Phenanthroline. There is one [2,7]phenanthroline structure (Table 3, **25**) bound to JNK3 kinase in the PDB. The ligand binds to the hinge region with one hydrogen bond between the N7 of the phenantholine system and NH of Met146 (the gk+3 residue) with the *p*-phenol substituent oriented toward solvent. A second hydrogen bond occurs between the N2 of the phenanthroline and Gly75 NH of the activation loop.

Quinoline. There is only one hinge-binding quinoline structure (Table 3, **26**) bound to TGFBR1 kinase. The binding mode is similar to the quinazoline structures (Table 3, **12**), one hydrogen bond in the hinge region and the fused benzene ring on the solvent side. It is interesting that pyrazole with two possible hydrogen bonds does not bind to the hinge.

N-Arylthiazol/oxazol-2-amine. There are two structures of *N*-phenyloxazol-2-amine (Table 3, **27**) bound to KDR. Both complexes have the anilinooxazole ring system bound with two hydrogen bonds with Cys917, the gk+3 residue. Three related thiazole analogues (one being dasatinib, *N*-(2-chloro-6-methylphenyl)-2-(6-(4-(2-hydroxyethyl)piperazin-1-yl)-2-methylpyrimidin-4-ylamino)thiazole-5-carboxamide) bind in similar fashion to ABL1 and CDK2 and GSK3β.

3-Hydroxypyran-4-one. There are two 3-hydroxypyran-4-one inhibitors bound in two kinases SRCK and CDK6 (Table 3, **28/28a**).[58] The chromone ligands are displayed in two distinct binding modes. In SRCK, the carbonyl and the hydroxyl of the chromone core bind to the gk+1 and gk+3 residues with the phenoxy substituent placed into the specificity pocket. In CDK6, the chromone has an extra hydroxyl group on the fused benzene ring. This hydroxyl group is bound to the gk+1 residue, which places the phenoxy ring toward the solvent channel.

4-Hydroxy-*N*-methylbenzamide. There are four structures of 4-hydroxy-*N*-methylbenzamide (Table 3, **29**) bound to PRKACA. All structures have the terminal phenol O−H in a bidentate interaction with the hinge gk+1 and gk+3 residues.

2,3-Dihydroisoindol-1-one. There are 32 cocrystal structures incorporating a 2,3-dihydroisoindol-1-one (Table 3, **30**) with 21 different kinases in the PDB. The bound ligands contain this basic substructure which is found in the natural product staurosporine. The binding modes found in all cases have the lactam carbonyl interacting with gk+3 and the lactam NH oriented toward gk+1.[59]

1,3-Dihydroindol-2-one. There are 23 structures of 1,3-dihydroindol-2-one (Table 3, **31**) bound to eight kinases in the PDB. Except for two, all have the indolone lactam group bound to the gk+1 and gk+3 residues.[60] There is one structure of CSNK1 kinase that shows the NH and carbonyl of the lactam

pointed toward the hinge but has the molecule flipped 180° so the carbonyl is facing the gatekeeper region while the N(H) is close to the gk+3 NH. It is quite possible that this structure is bound in a different tautomeric form. The other flipped structure came from PIM kinase, which does not have the hydrogen donor function of the hinge (gk+3 is proline) and displays a flipped orientation even for ATP.

Pyrazin-2-ylamine. There are two structures of a pyrazin-2-ylamine (Table 3, **32**) and (Table 3, **32a**) bound to two kinases, CDK2 and CDK5. The ligand bound to CDK2 is oriented with pyrazine N4 toward gk+3 of the hinge. The amino functionality is oriented out toward solvent. A bulky chlorine at the 6-position may be responsible for this binding mode. In the structure of the CDK5 bound ligand, the ligand is oriented with the aminopyrazine functionality toward the hinge region of the protein backbone. The N2-H and N1 of the ligand form hydrogen bonds with the gk+3 residue.

1-Phenyl-3,4-dihydro-1*H*-pyrimidin-2-one/1*H*-pyrid-2-one. There are two structures of 1-phenyl-3,4-dihydro-1*H*-pyrimidin-2-one/1*H*-pyrid-2-one (Table 3, **33**), both bound to p38 in the PDB. Both inhibitors are oriented with the carbonyl of the partially saturated ring and the unsubstituted amide toward the hinge binding with gk+1 and gk+3. The difluorobenzene ring is pointed toward the small gatekeeper (Thr) of p38. The specificity of these type of inhibitors for p38 kinase has been attributed to the hinge region Gly (gk+4) residue.[61]

1*H*-Pyrrolo[2,3-*b*]pyridine. There is only one kinase structure (CHK1) with a ligand of this type (Table 3, **34**). The binding mode is similar to the pyrimidine analogues, and the pyrrole NH is bound to gk+1 residue.

1*H*-Imidazo[5,4-*d*]pyridin-6-amine. There is only CDK2 structure with this type of ligand (Table 3, **35**). This is an interesting class of inhibitor, since it does not have the hydrogen hinge hydrogen acceptor observed in most kinase ligands. It also has a somewhat unusual binding mode, since the 2-aminopyridine could bind with the gk+1, gk+3 residues with two regular hydrogen bonds.[62]

1*H*-Pyrazole. Although a few more complex hinge binding pyrazoles were already discussed (Table 3, **21**), all hinge binding pyrazoles (Table 3, **36**) in which the 1H is bound to gk+1 are combined in this entry.[63]

5-Methylidene-6,7-dihydro-1*H*-azepin-2-one. There are six structures of this type (Table 3, **37**) with five kinases. All these structures have a fused five-member heterocyclic ring to the azepin ring and are related to hymenialdisine.[64]

3-Substituted-1-pyridin-2-ylurea. These intermolecular hydrogen-bonded 2-pyridylureas (Table 3, **38**) bind with two hydrogen bonds at the gk+3 residue of the hinge and the pyridyl group on the solvent side. There are two such structures, both with CDK2.[65]

3,4-Disubstitutedpyrrole-2,5-dione. There are 11 structures of this type of ligand (Table 3, **39**) with five different kinases. For most of these ligands, the two substituents are unsymmetrical indol-3-yl groups.[66]

4-Substituted Pyridine. There are 25 structures where a 4-substituted pyridine (Table 3, **40**) has only one hydrogen bond in the hinge region. A few of these examples were already included under scaffold (Table 3, **19**). Three of these complexes are also DFG-out type.

Rare Hinge Binding Moieties. Among the last three scaffolds (Table 3, **41−43**) are a few examples of DFG-in ligands
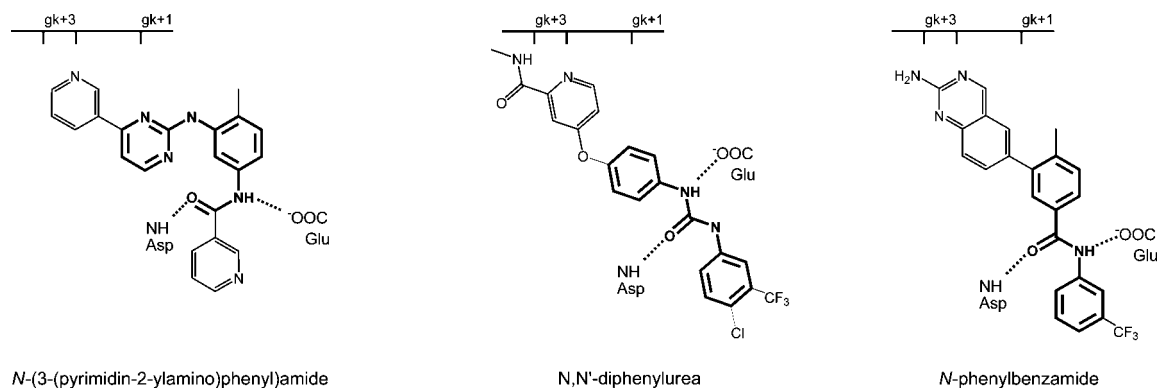
Figure 6. Depiction of relative ligand-hinge orientation of three commonly encountered classes of DFG-out kinase ligands. NH-Asp refers to the backbone NH of the Asp of the DFG triplet, whereas the ⁻OOC-Glu denotes the carboxy side chain of the conserved Glu residue of helix C.

with unusual hinge binding elements like oxime, nitro, or no hydrogen bonds at all.

Structures and Binding Modes of DFG-Out Ligands

DFG-out ligands have distinct structures and binding modes that warrant an independent discussion of their structural classification, their binding modes, and a separate summary in Table 4.[67] Ligand–protein complexes of DFG-out ligands usually display a number of noteworthy differences from those discussed in Tables 2 and 3.

(i) Within the kinase structure, the DFG-triplet changes its conformation. The catalytic aspartate (D) is moved out of the ATP pocket accompanied by phenylalanine (F), which vacates a hydrophobic pocket.

(ii) A hydrophobic group from the ligand can engage the hydrophobic pocket originally occupied by phenylalanine (F).

(iii) Discrete hydrogen bonds are observed between ligand and the aspartate-backbone NH, and a highly conserved glutamate side chain of helix C.

The Protein Data Bank contains several kinase–ligand complexes with a DFG-out conformation lacking a ligand moiety occupying the empty hydrophobic pocket. These ligands are also distinguished by an absence of hydrogen bonds with the aspartate-backbone NH or with the Glu side chain. Even though the proteins of these ligand–protein complexes resemble those of "DFG-out"-like apo structures, their ligand structures were included in the analysis of more conventional DFG-out ligands. The classification of DFG-out ligands focused on the unique structural features penetrating beyond the ATP binding pocket rather than those binding the hinge region; in fact, there are several ligands without a hinge binding element. Table 4 summarizes structural features and binding modes observed for DFG-out ligands; a few representative binding modes are discussed in Figure 6.

Pyrimidino-*m*-phenylenediamines (Imatinib-like Inhibitors). There are five kinase structures with this scaffold (Table 4, **44**): one ABL, three c-ABL, and one c-KIT. All of these complexes display the same binding mode in which the pyridine is hydrogen-bonded with the gk+3 NH.[68] The amide oxygen forms a hydrogen bond with the backbone aspartate NH of DFG, while amide NH is hydrogen-bound to the glutamate of helix C.

1-(5-*tert*-Butylpyrazol-3-yl)-3-phenylurea. There are five entries in the PDB that include this substructure (Table 4, **45/45a**). All five belong to the same kinase (p38). The first two structures (**45**) have a single hinge binding element, gk+3 NH,[69] and the other three structures (**45a**) do not show a hinge binding element in the crystal structure. All five structures have one

hydrogen bond with the backbone aspartate NH of DFG, a hydrogen bond with a Glu carboxylate of helix C, and occupy the hydrophobic pocket exposed in the DFG-out conformation.

***N,N′*-Diphenyluera.** There are four kinase structures with this scaffold (Table 4, **46**) belonging to three kinases: b-RAF (two), VEGFR2, and c-ABL. All have hydrogen bonds between the urea moiety on one hand and the backbone aspartate NH of DFG and a Glu carboxylate on the other.[68] These ligands also include a hinge binding element (gk+3 NH) attached to the interior phenyl group of the diphenylurea core. The exterior phenyl group occupies the DFG-out hydrophobic pocket.

***N*-Phenylbenzamide.** This scaffold is a substituted *N*-phenylbenzamide and may be considered as a hybrid of diphenylureas and imatinib-like structures (Table 4, **47–50**). If one considers the orientation of the amide moiety in imatinib to be typical, then some of the amides in this class could be considered reverse amides. All have a hinge binding element; the amide of the ligand is hydrogen-bonded to both the backbone NH of aspartate (DFG triplet) and the glutamate of helix C, and a hydrophobic group engages the vacated Phe pocket.

Questionable DFG-Out Ligands. Six DFG-out kinase structures were identified: one for c-ABL, two for ABL, and three for c-FMS (Table 4, **51–53**). Surprisingly, none of these ligands have the common DFG-out hydrogen bond or hydrophobic pocket occupied. These unusual ligand–protein complexes may help to uncover structural features that contribute or are required for the formation of the DFG-out conformation.

DFG-Out Conformation and Possible Roles of Two Key Residues

In addition to the variable number of hinge hydrogen bonds an inhibitor can engage in, it can bind to different protein conformations and phosphorylation states. Targeting the DFG-out conformation of kinases has become an important kinase inhibitor design strategy.[70] One popular rationale often advanced is that DFG-out represents an inactive conformation of a kinase, which has been observed only for some select members of the kinase family. Furthermore, inactive conformations are thought to be structurally more diverse with respect to other kinases when compared to their activated counterparts and thus might offer a path for the design of more selective inhibitors. Only a few kinases have been experimentally observed in both DFG-in and DFG-out conformations, while for most others no DFG-out structures have been published, although DFG-in structures are known. Which factors contribute to the propensity or ability of kinases to "attain" this conformation is unclear and remains an intriguing question in kinase drug discovery. 80% of human kinases have the DFG-triplet as the starting point of the

Table 5. DFG-Out Kinases, Their Two Key Residues, and a Representative PDB ID

kinase gene	gatekeeper type	representative PDB ID	XDFG sequence	comments
ABL	Thr	2hiw	ADFG	regular[a]
ABL	Thr	2hzi	ADFG	DFG-in type ligand
ABL1	Thr	1fpu	GDFG	regular
Aurora A (STK6)	Leu	2c6e	ADFG	unusual[b]
BRAF	Thr	1uwh	GDFG	regular
BRAF	Thr	3c4c	GDFG	unusual
c-ABL	Thr	1iep	ADFG	regular
c-ABL	Thr	2fo0	ADFG	DFG-in type ligand
c-FMS	Thr	2i0v	GDFG	DFG-in type ligand
c-KIT	Thr	1t46	CDFG	regular
c-MET	Leu	3c1x	ADFG	regular
INSR	Met	1irk	GDFG	apo
KDR	Val	1ywn	CDFG	regular
LCK	Thr	2ofv	ADFG	regular
MAPK14	Thr	1kv1	LDFG	regular
SRC	Thr	1y57	ADFG	DFG-in structure with imatinib!

[a] These DFG-out inhibitors place a hydrophobic group into the pocket vacated by the DFG-Phe originally occupied in the DFG-in structure; the inhibitor forms hydrogen bonds with the backbone NH of DFG-Asp and Glu. [b] The inhibitor does not occupy the vacated DFG-Phe pocket as is observed in "regular" DFG-out structures; it also lacks the two hydrogen bonds with conserved Asp and Glu residues.

activation loop and have the helix C Glu necessary for the hydrogen bond observed with many DFG-out ligands. However, in reality only a few kinases so far yielded DFG-out structures. An analysis of DFG-out structure characteristics (Table 5) supports the following observations:

(1) DFG-out conformations of kinases exist in both the presence and absence of ligands.

(2) Most DFG-out complexes show discrete hydrogen bonding interactions between ligand and protein (vide supra); however, other examples of ligand-bound kinases without such contacts also exist.

These observations at least partially refute earlier hypotheses[70] concerning the higher relative energy of DFG-out conformations and their need to stabilize ligand−protein interactions. Furthermore, the currently available PDB data set suggests the possible involvement of two key residues in the DFG-in to DFG-out dynamics, the gatekeeper as well as the N-terminal residue preceding the DFG triplet. Threonine (Thr), with a relatively small side chain, is by far the most common gatekeeper residue among kinases found in any DFG-out conformation, whether it be with DFG-in type or "DFG-out" type ligands (Table 5). In fact, a small gatekeeper residue (Thr or Val) appears to be a requirement for a "regular" (prototypical) DFG-out conformation except for a recently released c-Met structure.[71] The apparent requirement for a small gatekeeper residue to achieve stereotypical DFG-out conformations and ligand interactions is further supported by imatinib resistance data of T315A (2.4-fold wild-type IC_{50}, no major loss of activity) and T315I (>30-fold wild-type IC_{50}) mutants in BCR-ABL.[72] Gatekeeper and DFG-out ligand are quite proximal in the binding site, and a larger gatekeeper residue is more likely to cause steric congestion during complex formation with a ligand entering the DFG-out hydrophobic pocket.

The case of imatinib binding to different kinases serves as a contrasting example: Met is the gatekeeper residue in SYK with which it binds as a standard (DFG-in) 2-aminopyrimidine hinge binding inhibitor (PDB 1xbb). This favors the requirement of the small gatekeeper residue. However, even with the smaller Thr in SRC kinase (PDB 1y57), a similar binding mode is observed, indicating that a small gatekeeper alone even with a ligand possessing the requisite functionality is not sufficient to capture a DFG-out conformation.

Another observation of DFG-out protein structures concerns the N-terminal amino acid residue preceding the DFG triplet. An examination of XDFG sequences in Table 5 points to a high frequency of small glycine or alanine residues (increased flexibility?) in DFG-out structures without ligand or with "unusual" DFG-out type ligands. Table 5 also suggests that no DFG-out examples are currently reported in the PDB, where both the gatekeeper and X are large residues.

Empirical Rules To Predict Hinge Binding Orientation

The preceding analyses supports the conclusion that most kinase inhibitors form two, some even three hydrogen bonds with the hinge, involving one hydrogen acceptor flanked by one or two hydrogen donors. This array on the ligand serves to complement the backbone NH and two backbone C=O pointed toward the ATP binding pocket. It appears straightforward to identify the moieties most likely to be involved in hinge binding on the ligand, but it may be less clear which of the two hydrogen acceptors from the kinase hinge region (gk+1 or gk+3) will be engaged by a ligand. Similarly, it may not be apparent why a ligand forms only one or two out of three possible hydrogen bonds. Many of these binding mode related questions may be addressed with the following observations:

(1) A primary aromatic amine adjacent to the nitrogen in pyridine, pyrimidine, and other heterocyclic aromatic rings will have a tendency to bind to the backbone C=O near the gatekeeper (gk+1 residue). A hydrogen bond between the gk+1 backbone C=O and ligand is energetically favored over one with its gk+3 counterpart. Its position deep in the pocket results in less competition from solvent molecules, and a hydrogen bond formed here avoids the expensive prospect of an immobilized water molecule between ligand and protein.

(2) Secondary aromatic amines with an adjacent hydrogen acceptor (or cyclic secondary amides) in a five-membered ring, as in 7-azaindoles and staurosporine analogues, will also bind to the backbone C=O near the gatekeeper. An adjacent bulky group may push this type of NH to the solvent side, gk+3 area.

(3) Most other secondary amines and amides with a bulky group such as aryl, heteroaryl, or benzyl will bind to the solvent side backbone C=O (gk+3 residue). An exception to this rule is observed in cases where a phenyl is part of a 5 + 6 fused bicyclic system (i.e., oxindoles, benzopyrazoles) or if a methylene homologue is constrained into a 6 + 6 fused system (i.e., a dihydrobenzonaphthyridine).

(4) Considering observations 2 and 3, 2,4-dianilinopyrimidines will bind with only two hydrogen bonds at the hinge region. The second potential hydrogen donor anilino group is positioned toward the activation loop.

Ligand Design Strategies and Available Tools

The foregoing observations and rules can be fed back during the ligand design process. To generate ideas for new analogues of an active lead or engage in lead/scaffold hopping activities, medicinal chemists often start with a structure drawing program (i.e., ISISDraw, Chemdraw, ChemSketch, MarvinSketch, JME Editor) and a known active compound and scaffold of interest. They align the scaffold of interest to the known active compound and follow the topology of the active compound to add/modify structural moieties in the chosen scaffold. The aligned structures compressed into 2D
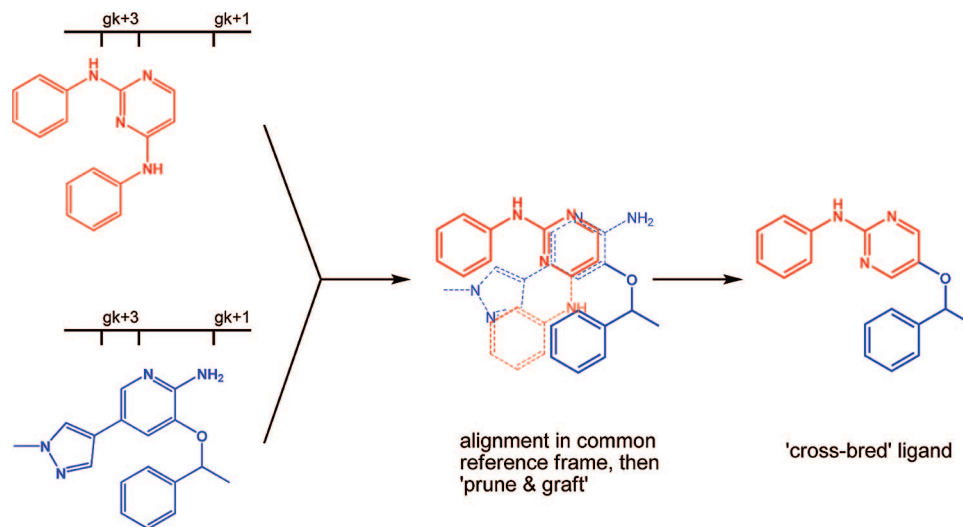
Figure 7. Basic concept of generating cross-bred ligands: (1) ligand−protein structure complexes have been aligned into a common reference frame; (2) the position of their hydrogen bond forming moieties relative to the gk+3 and gk+1 residues of the hinge orient the scaffolds such that fragmentation and reconnection points can be identified (either computationally or manually); (3) hybrid structures, containing portions from both scaffolds, are generated.

found in the preceding tables and some of the binding mode prediction rules may prove useful for this "back-of-the-envelope" approach.[73] This concept can also be employed in a more quantitative (subject to real 3D constraints, such as ligand bond lengths and angles as well as steric fit with the receptor) and automated fashion using the Ligand-Cross program developed by Eidogen-Sertanty (http://www.eidogen-sertanty.com/), which is closely related to the recently described BREED approach.[74] Using automated or manual docking, one can place a scaffold in the coordinate frame of a reference receptor, providing some tolerance for bond distance and bond angles. The program then generates all possible cross-bred or hybridized ligands by connecting structural moieties from the PDB ligands that are within preset limits. The basic concept of cross-ligand generation is illustrated in Figure 7.

Does an Optimal Number of Hinge Hydrogen Bonds Exist? The availability of three hydrogen bond forming moieties in the hinge region of kinases coupled with the observation that only one, two, or all three are engaged by various ligands poses the interesting question of the optimal number of hydrogen bonds for kinase drugs. Increasing the number of hydrogen bonds between a ligand and protein is sometimes considered a viable or even promising approach during lead affinity optimization (it is mentioned that attempts to add discrete hydrogen bonds between ligand and protein, to increase affinity or selectivity, have met with spectacular success as well as failure). However, the above analysis of kinase-bound ligand structures, particularly those aspects pertaining to approved drugs, directs one to a conflicting conclusion (Figure 8). Most of the approved kinase inhibitors form only one hydrogen bond with the hinge. No clear correlation between the number of hinge hydrogen bonds and ligand affinity emerges from the approved kinase inhibitor data set. The sole inhibitor example with three hinge hydrogen bonds (sunitinib, (*Z*)-*N*-(2-(diethylamino)ethyl)-5-((5-fluoro-2-oxoindolin-3-ylidene)methyl)-2,4-dimethyl-1*H*-pyrrole-3-carboxamide) does not achieve a higher kinase affinity than those that have only one observed hinge interaction. Furthermore, one of the two inhibitors with two hinge hydrogen bonds (dasatinib) attains the highest apparent affinity, whereas the other example (ruboxistaurin, 13-((dimethylamino)methyl)-10,11,14,15-tetrahydro-4,9:16,21-dimetheno-1*H*,13*H*-dibenzo(*e*,*k*)pyrrolo(3,4-

h)(1,4,13)oxadiazacyclohexadecene-1,3(2*H*)-dione) attains intermediate affinity. An analysis of hinge hydrogen bond number and ligand affinity is further complicated by the possible involvement of additional hydrogen bonds of the proximate gatekeeper residue with the inhibitor. In the case of the quinazoline-based inhibitors of EGFR, no additional ligand−threonine gatekeeper hydrogen bonds appear to be involved, whereas both ABL ligands, imatinib and dasatinib, form readily observed hydrogen bonds between inhibitor and the hydroxyl of threonine.

One caveat and a valid objection to this observation's statistical significance is the small number of examples in the data set. The success of molecules containing predominantly a single hydrogen bond (acceptor) functionality coordinating with the hinge stands out against what can be observed in contemporary kinase lead identification and optimization research. This stark contrast manifests itself in the number of published kinase inhibitors with two or more hydrogen bonds that have emerged from kinase-directed drug discovery research to appear in patents and scientific publications (http://www.eidogen-sertanty.com/products_kinasekb.html). The apparent success of kinase inhibitors with one hinge binding interaction over those with two or three is easily reconciled with established trends of reducing the overall number of hydrogen bond donors and acceptors (as well as the correlated polar surface area) to impart more favorable PK properties as a compound progresses from lead optimization through clinical development.[8]

Future Perspectives and Concluding Remarks

An understanding of inhibitor binding modes of many structurally diverse ligands bound to proteins of a discrete target family is undoubtedly very useful for ligand design. In this article we have only surveyed the binding modes of kinase inhibitors; however, the same approach may be similarly useful for other classes of proteins. The conclusion for kinase inhibitor molecules is that their binding modes can be categorized in terms of the ligand core structure and a few simple empirical rules. The extremely rapid growth of available kinase structures may require revision of these rules or addition of qualifiers for them to remain relevant for future ligands. Medicinal chemists may find the prediction rules devised herein helpful to quickly
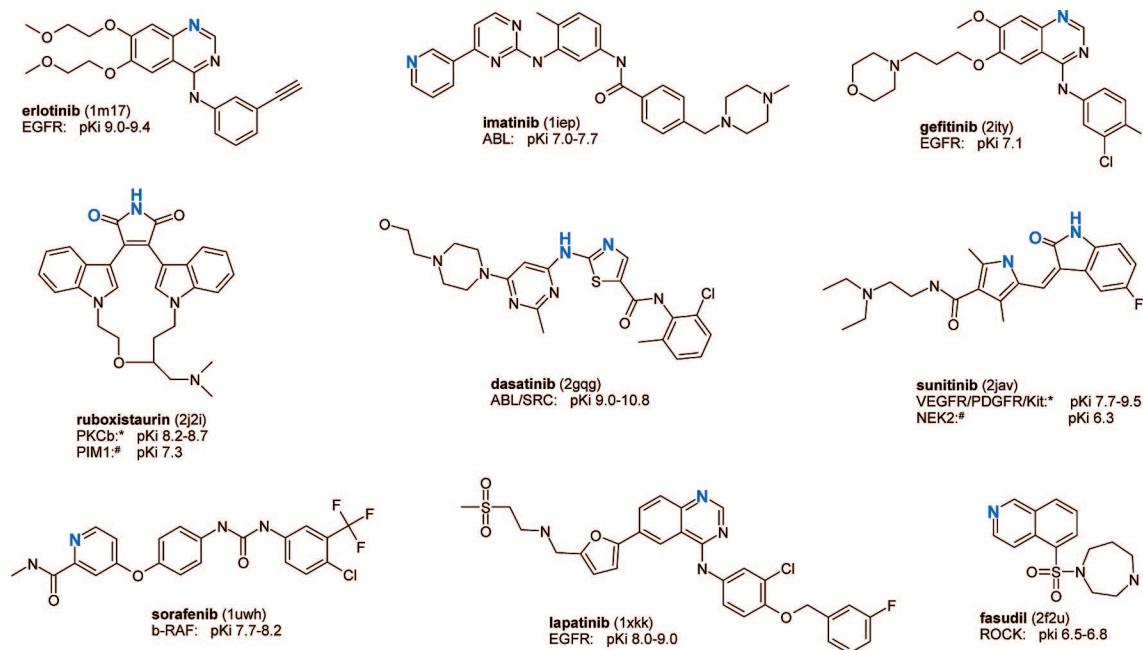
Figure 8. Number and location of heteroatoms in approved kinase inhibitors for which hydrogen bonds with the hinge can be observed crystallographically (PDB ID in parentheses). Hinge hydrogen bonding atoms are highlighted in blue. Approximate affinities (expressed as pK_i ranges computed from enzyme assay K_d and IC$_{50}$ data) for primary therapeutic target, molecular target it achieves highest activity against, and X-ray crystallography complexes are provided; their identities are the same for all but two examples. For ruboxistaurin and sunitinib, the ∗ designates therapeutic and highest activity kinase target whereas # indicates the kinase of the crystallographic complex.

formulate a binding mode hypothesis for a potential ligand. Likewise, a substructure searchable database combining a simplified representation of scaffolds with a 2D depiction of aligned 3D ligand structures may also prove useful. For computational chemists, such a database may help in picking the "correct" ligand pose one expects to observe experimentally from among the many generated by various ligand docking programs. A cross-breeding ligand hybridization methodology coupled with sophisticated ligand affinity prediction tools can be a useful hit/lead generation strategy; however, all of the aforementioned uses of a kinase knowledge base should be considered suitable only for idea generation rather than the routine prediction of high affinity ligands. Successful ligand design strategies also need to address the incomplete understanding of desirable vs undesirable off-target activities/selectivities, as well as inherently challenging physicochemical property profiles of kinase ligands. The predominant lipophilicity of these molecules is a consequence of a primarily hydrophobic binding pocket with a small polar hinge recognition site. A benchmark for physicochemical property profile for kinase inhibitors in drug discovery is still evolving, since only a limited number of kinase drugs have been approved thus far. Therefore, kinase inhibitor design should leverage the property profiles of approved oral drugs with those of approved kinase drugs (or in case of CNS kinase targets those of approved CNS drugs). Knowledge based drug discovery is an excellent way in ligand design to mate privileged structural components with exploratory moieties. Paradoxically, the advantages of knowledge based drug discovery depend critically on an increase in information that can only come from speculative approaches and serendipitous results.

Biographies

Arup K. Ghose received his Ph.D. degree in Medicinal and Computational Medicinal Chemistry from Jadavpur University, Calcutta, India, in 1978 under the guidance of Professor Arun Udaya De. In 1981 he moved to U.S. to join Professor Gordon M. Crippen as a postdoctoral fellow, initially at Texas A&M University and then at the University of Michigan. For the past 20 years he has been working in the pharmaceutical research industry at places like Sterling-Winthrop, Amgen, and Locus Pharmaceuticals. Currently he is a Senior Research Scientist and the Head of Computational Chemistry at Cephalon. He developed various computational methods for drug discovery, like AlogP, 3D-QSAR, pharmacophore modeling, characterization of druglike molecules, and protein structure based ligand design. He published over 60 papers and edited a book on computational methods in drug discovery.

Torsten Herbertz earned a Ph.D. in Physical Organic Chemistry from Rutgers University, New Brunswick, NJ, in 1998. Afterward, he joined the Discovery Department of Viropharma, Inc. as a Medicinal and Computational Chemist working on virology targets. In 2004, he worked as a Consultant with Adolor Corporation prior to joining Cephalon as a Research Scientist in Computational Chemistry in 2005. His main interests are better approximation of the dynamics and physics of ligand−protein interactions toward improved lead optimization and chemoinformatics-based sound lead identification and generation strategies.

Douglas A. Pippin earned a B.Sc. (honors) degree in Chemistry from Central Connecticut State University in 1996 and Ph.D. in Synthetic Organic Chemistry in 2001 from Dartmouth College under the guidance of Professor Peter A. Jacobi. In 2002, he began work at CGI Pharmaceuticals (formerly Cellular Genomics, Inc.) in Branford, CT, as a Senior Research Scientist where he led high throughput medicinal chemistry efforts in early lead identification and lead optimization of targeted kinase programs in oncology and autoimmune/anti-inflammatory disease. In 2005, Dr. Pippin joined Cephalon, Inc. where he is currently engaged in investigation of agents for treatment of cellular proliferation and CNS disorders.

Joseph M. Salvino earned his B.S. degree in Chemistry from Kutztown State College, Kutztown, PA, in 1980 and Ph.D. in Organic Chemistry from Brown University, RI, in 1987 under the guidance of Professor Paul G. Williard. He did postdoctoral work at the University of Pennsylvania from 1988 to 1989 under the supervision of Professors K. C. Nicolaou and Ralph Hirschmann. He started his professional career with Dupont Central Research and over the years worked at places like Sterling Winthrop and Rhone Poulenc Rorer with increasing responsibility. Currently he is the Director of the Hit-to-Lead Chemistry Department in Cephalon. His major interest in drug discovery research is in the areas of CNS and kinases using parallel chemistry. He has published over 75 papers and patents.

John P. Mallamo earned his B.S. degree in Chemistry and Mathematics from Colorado State University in 1977, where he worked three years in the laboratories of K. E. Debruin and A. I. Meyers as an undergraduate. In 1981 he completed his Ph.D. thesis studies at Johns Hopkins University under the guidance of Professor Gary H. Posner. He began his industrial career with Sterling-Winthrop Pharmaceuticals and progressed there with increasing responsibility. In early 1994 he joined the staff at Cephalon, Inc. where he is currently Vice President, World Wide Chemical Research and Development. His research interest includes pharmaceutical discovery in areas like oncology and CNS. He has published over 130 papers and patents.

Supporting Information Available: Additional chemical structures and information. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Ghose, A. K.; Viswanadhan, V. N. *Combinatorial Library Design and Evaluation: Principles, Software Tools and Applications in Drug Discovery*; Marcel Dekker, Inc.: New York, 2001.

(2) Guner, O. F. *Pharmacophore Perception, Development and the Use in Drug Design*; IUL Press: La Jolla, CA, 2000.

(3) Ghose, A. K.; Logan, M. E.; Treasurywala, A. M.; Wang, H.; Wahl, R. C.; et al. Determination of Pharmacophoric Geometry for Collagenase Inhibitors Using a Novel Computational Method and Its Verification Using Molecular Dynamics, NMR, and X-ray Crystallography. *J. Am. Chem. Soc.;* **1995**, *117*, 4671–4682.

(4) Hansch, C.; Leo, A.; Hoekman, D. *Monograph: Exploring the QSAR. Hydrophobic, Electronic and Steric Constants*; American Chemical Society: Washington, DC, 1995.

(5) Cramer, R. D. Topomer CoMFA: A Design Methodology for Rapid Lead Optimization. *J. Med. Chem.* **2003**, *46*, 374–388.

(6) Ghose, A. K.; Crippen, G. M.; Revankar, G. R.; McKernan, P. A.; Smee, D. F.; et al. Analysis of the in vitro antiviral activity of certain ribonucleosides against parainfluenza virus using a novel computer aided receptor modeling procedure. *J. Med. Chem.* **1989**, *32*, 746–756.

(7) Tominaga, Y.; Jorgensen, W. L. General model for estimation of the inhibition of protei kinases using Monte Carlo simulations. *J. Med. Chem.* **2004**, *47*, 2534–2549.

(8) Ghose, A. K.; Herbertz, T.; Salvino, J. M.; Mallamo, J. P. Knowledge-based chemoinformatic approaches to drug discovery. *Drug Discovery Today* **2006**, *11*, 1107–1114.

(9) Buijsman, R. Structural Aspects of Kinases and Their Inhibitors. *Chemogenomics in Drug Discovery: A Medicinal Chemistry Perspective*; Wiley-VCH Verlag GmbH: Weinheim, Germany, 2004.

(10) Klebl, B. M.; Muller, G. Second generation kinase inhibitors. *Expert Opin. Ther. Targets* **2005**, *9*, 975–993.

(11) Vulpetti, A.; Pavarello, P. An analysis of the binding modes of ATP-competitive CDK2 inhibitors as revealed by X-ray structures of protein−inhibitor complexes. *Curr. Med. Chem.: Anti-Cancer Agents* **2005**, *5*, 561–573.

(12) Perola, E. Minimizing false positives in kinase virtual screens. *Proteins: Struct., Funct., Bioinf.* **2006**, *64*, 422–435.

(13) McGregor, M. J. A pharmacophore map of small molecule protein kinase inhibitors. *J. Chem. Inf. Model.* **2007**, *47*, 2374–2382.

(14) Traxler, P.; Furet, P. Strategies toward the design of novel and selective protein tyrosine kinase inhibitors. *Pharmacol. Ther.* **1999**, *82*, 196–206.

(15) Liao, J. J.-L. Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. *J. Med. Chem.* **2007**, *50*, 409–424.

(16) Blume-Jensen, P.; Hunter, T. Oncogenic kinase signalling. *Drug Discovery Today* **2001**, *10*, 839–846.

(17) Leach, A. R.; Shoichet, B. K.; Peishoff, C. E. Prediction of protein−ligand interactions. Docking and scoring: successes and gaps. *J. Med. Chem.* **2006**, *49*, 5851–5855.

(18) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; et al. A critical assesment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.

(19) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.

(20) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; et al. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196.

(21) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; et al. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.

(22) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel procedure for modeling ligand receptor induced fit effects. *J. Med. Chem.* **2006**, *49*, 534–553.

(23) Bowman, A. L.; Nikolovska-Coleska, Z.; Zhong, H.; Wang, S.; Carlson, H. A. Small molecule inhibitors of the MDM2-p53 interaction discovered by ensemble-based receptor models. *J. Am. Chem. Soc.* **2007**, *129*, 12809–12814.

(24) Damm, K. L.; Carlson, H. A. Exploring experimental sources of multiple protein conformations in structure-based drug design. *J. Am. Chem. Soc.* **2007**, *129*, 8225–8235.

(25) Michel, J.; Verdonk, M. L.; Essex, J. W. Protein−ligand binding affinity predictions by implicit solvent simulations: a tool for lead optimization. *J. Med. Chem.* **2006**, *49*, 7427–7439.

(26) Michel, J.; Taylor, R. D.; Essex, J. W. Efficient generalized Born models for Monte Carlo simulations. *J. Chem. Theory Comput.* **2006**, *2*, 732–739.

(27) Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M. Validation and use of the MM-PBSA approach for drug discovery. *J. Med. Chem.* **2005**, *48*, 4040–4048.

(28) Lyne, P. D.; Lamb, M. L.; Saeh, J. C. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J. Med. Chem.* **2006**, *49*, 4805–4808.

(29) Weis, A.; Katebzadeh, K.; Soderhjelm, P.; Nilsson, I.; Ryde, U. Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field. *J. Med. Chem.* **2006**, *49*, 6596–6606.

(30) Alberts, I. L.; Todorov, N. P.; Dean, P. M. Receptor flexibility in de novo ligand design and docking. *J. Med. Chem.* **2005**, *48*, 6585–6596.

(31) Cavasotto, C. N.; Kovacs, J. A.; Abagyan, R. A. Representing receptor flexibility in ligand docking through relevant normal modes. *J. Am. Chem. Soc.* **2005**, *127*, 9632–9640.

(32) Maple, J. R.; Cao, Y.; Damm, W.; Halgren, T. A.; Kaminski, G. A.; et al. A polarizable force field and continuum solvation methodology for modeling of protein−ligand interactions. *J. Chem. Theory Comput.* **2005**, *1*, 694–715.

(33) Hantschel, O.; Superti-Furga, G. Regulation of the c-Abl and BCR-Abl tyrosine kinases. *Nat. Rev.* **2004**, *5*, 33–44.

(34) Brown, N. R.; Noble, M. E.; Lawrie, A. M.; Morris, M. C.; Tunnah, P.; et al. Effects of phosphorylation of threonine 160 on cyclin-dependent kinase 2 structure and activity. *J. Biol. Chem.* **1999**, *274*, 8746–8756.

(35) Kumar, A.; Mandiyan, V.; Suzuki, Y.; Zhang, C.; Rice, J.; et al. Crystal Structures of proto-oncogene kinase Pim1: a target of aberrant somatic hypermutations in diffuse large cell lymphoma. *J. Mol. Biol.* **2005**, *348*, 183–193.

(36) Rudberg, P. C.; Tholander, F.; Thunnissen, M. M.; Haeggstrom, J. Z. Leukotriene A4 hydrolase/aminopeptidase. Glutamate 271 is a catalytic residue with specific roles in two distinct enzyme mechanisms. *J. Biol. Chem.* **2002**, *277*, 1398–1404.

(37) Sabat, M.; Vanrens, J. C.; Clark, M. P.; Brugel, T. A.; Maier, J.; et al. The development of novel C-2, C-8, and N-9 trisubstituted purines as inhibitors of TNF-alpha production. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4360–4365.

(38) Eswaran, J.; Lee, W. H.; Debreczeni, J. E.; Filippakopoulos, P.; Turnbull, A.; et al. Crystal structures of the p21-activated kinases PAK4, PAK5, and PAK6 reveal catalytic domain plasticity of active group II PAKs. *Structure* **2007**, *15*, 201–213.

(39) Wang, Z.; Canagarajah, B. J.; Boehm, J. C.; Kassisa, S.; Cobb, M. H.; et al. Structural basis of inhibitor selectivity in MAP kinases. *Structure* **1998**, *6*, 1117–1128.

(40) Witucki, L. A.; Huang, X.; Shah, K.; Liu, Y.; Kyin, S.; et al. Mutant tyrosine kinases with unnatural nucleotide specificity retain the structure and phospho-acceptor specificity of the wild-type enzyme. *Chem. Biol.* **2002**, *9*, 25–33.

(41) Cowan-Jacob, S. W.; Fendrich, G.; Manley, P. W.; Jahnke, W.; Fabbro, D.; et al. The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **2005**, *13*, 861–871.

(42) Kontopidis, G.; McInnes, C.; Pandalaneni, S. R.; McNae, I.; Gibson, D.; et al. Differential binding of inhibitors to active and inactive CDK2 provides insights for drug design. *Chem. Biol.* **2006**, *13*, 201–211.

(43) Nagar, B.; Hantschel, O.; Seeliger, M.; Davies, J. M.; Weis, W. I.; et al. Organization of the SH3-SH2 unit in active and inactive forms of the c-Abl tyrosine kinase. *Mol. Cell* **2006**, *21*, 787–798.

(44) Beattie, J. F.; Breault, G. A.; Ellston, R. P.; Green, S.; Jewsbury, P. J.; et al. Cyclin-dependent kinase 4 inhibitors as a treatment for cancer. Part 1: identification and optimisation of substituted 4,6-bis anilino pyrimidines. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2955–2960.

(45) Pratt, D. J.; Bentley, J.; Jewsbury, P.; Boyle, F. T.; Endicott, J. A.; et al. Dissecting the determinants of cyclin-dependent kinase 2 and cyclin-dependent kinase 4 inhibitor selectivity. *J. Med. Chem.* **2006**, *49*, 5470–5477.

(46) Richardson, C. M.; Williamson, D. S.; Parratt, M. J.; Borgognoni, J.; Cansfield, A. D.; et al. Triazolo[1,5-*a*]pyrimidines as novel CDK2 inhibitors: protein structure-guided design and SAR. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1353–1357.

(47) Wood, E. R.; Truesdale, A. T.; McDonald, O. B.; Yuan, D.; Hassell, A. A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res.* **2004**, *64*, 6652–6659.

(48) Sielecki, T. M.; Johnson, T. L.; Liu, J.; Muckelbauer, J. K.; Grafstrom, R. H.; et al. Quinazolines as cyclin dependent kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1157–1160.

(49) Foloppe, N.; Fisher, L. M.; Howes, R.; Kierstan, P.; Potter, A.; et al. Structure-based design of novel Chk1 inhibitors: insights into hydrogen bonding and protein−ligand affinity. *J. Med. Chem.* **2005**, *48*, 4332–4345.

(50) Miyazaki, Y.; Matsunaga, S.; Tang, J.; Maeda, Y.; Nakano, M.; et al. Novel 4-amino-furo[2,3-*d*]pyrimidines as Tie-2 and VEGFR2 dual inhibitors. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 2203–2207.

(51) Burchat, A.; Borhani, D. W.; Calderwood, D. J.; Hirst, G. C.; Li, B.; et al. Discovery of A-770041, a src-family selective orally active lck inhibitor that prevents organ allograft rejection. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 118–122.

(52) Vogtherr, M.; Saxena, K.; Hoelder, S.; Grimme, S.; Betz, M.; et al. NMR characterization of kinase p38 dynamics in free and ligand-bound forms. *Angew. Chem., Int. Ed.* **2006**, *45*, 993–997.

(53) Bamborough, P.; Christopher, J. A.; Cutler, G. J.; Dickson, M. C.; Mellor, G. W.; et al. 5-(1*H*-Benzimidazol-1-yl)-3-alkoxy-2-thiophen-ecarbonitriles as potent, selective, inhibitors of IKK-epsilon kinase. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 6236–6240.

(54) Foloppe, N.; Fisher, L. M.; Francis, G.; Howes, R.; Kierstan, P.; et al. Identification of a buried pocket for potent and selective inhibition of Chk1: prediction and verification. *Bioorg. Med. Chem. Lett.* **2006**, *14*, 1792–1804.

(55) Hamdouchi, C.; Keyser, H.; Collins, E.; Jaramillo, C.; De Diego, J. E.; et al. The discovery of a new structural class of cyclin-dependent kinase inhibitors, aminoimidazo[1,2-*a*]pyridines. *Mol. Cancer Ther.* **2004**, *3*, 1–9.

(56) Bonn, S.; Herrero, S.; Breitenlechner, C. B.; Erlbruch, A.; Lehmann, W.; et al. Structural analysis of protein kinase A mutants with Rho-kinase inhibitor specificity. *J. Biol. Chem.* **2006**, *281*, 24818–24830.

(57) Jacobs, M.; Hayakawa, K.; Swenson, L.; Bellon, S.; Fleming, M.; et al. The Structure of dimeric ROCK I reveals the mechanism for ligand selectivity. *J. Biol. Chem.* **2006**, *281*, 260–268.

(58) Sicheri, F.; Moarefi, I.; Kuriyan, J. Crystal structure of the Src family tyrosine kinase Hck. *Nature* **1997**, *385*, 602–609.

(59) Zhou, T. J.; Sun, L. G.; Gao, Y.; Goldsmith, E. J. Crystal structure of the MAP3K TAO2 kinase domain bound by an inhibitor staurosporine. *Acta Biochim. Biophys. Sin.* **2006**, *38*, 385–392.

(60) Jautelat, R.; Brumby, T.; Schafer, M.; Briem, H.; Eisenbrand, G.; et al. From the insoluble dye indirubin towards highly active, soluble CDK2-inhibitors. *ChemBioChem* **2005**, *6*, 531–540.

(61) Fitzgerald, C. E.; Patel, S. B.; Becker, J. W.; Cameron, P. M.; Zaller, D.; et al. Structural basis for p38 MAP kinase quinazolinone and pyridol-pyrimidine inhibitor specificity. *Nat. Struct. Biol.* **2003**, *10*, 764–769.

(62) Gibson, A. E.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Curtin, N. J.; et al. Probing the ATP ribose-binding domain of cyclin-dependent kinases 1 and 2 with O6-substituted guanine derivatives. *J. Med. Chem.* **2002**, *45*, 3381–3393.

(63) Furet, P.; Meyer, T.; Strauss, A.; Raccuglia, S.; Rondeau, J.-M. Structure-based design and protein X-ray analysis of a protein kinase inhibitor. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 221–224.

(64) Meijer, L.; Thunnissen, A. M.; White, A. W.; Garneir, M.; Nikolic, M.; et al. Inhibition of cyclin-dependent kinases, gsk-3beta and ck1 by hymenialdisine, a marine sponge. *Chem. Biol.* **2000**, *7*, 51–63.

(65) Ikuta, M.; Kamata, K.; Fukasawa, K.; Honma, T.; Machida, T.; et al. Crystallographic approach to identification of cyclin-dependent kinase 4 (CDK4)-specific inhibitors by using CDK4 mimic CDK2 protein. *J. Biol. Chem.* **2001**, *276*, 27548–27554.

(66) Bertrand, J. A.; Thieffine, S.; Vulpetti, A.; Cristiani, C.; Valsasina, B.; et al. Structural characterization of the GSK-3*β* active site using selective and non-selective ATP-mimetic inhibitors. *J. Mol. Biol.* **2003**, *333*, 393–407.

(67) Bogoyevitch, M. A.; Fairlie, D. P. A new paradigm for protein kinase inhibition: blocking phosphorylation without directly targeting ATP binding. *Drug Discovery Today* **2007**, *12*, 622–633.

(68) Cowan-Jacob, S. W.; Fendrich, G.; Floersheimer, A.; Furet, P.; Liebetanz, J.; et al. Structural biology contributions to the discovery of drugs to treat chronic myelogenous leukaemia. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2007**, *63*, 80–93.

(69) Gill, A. L.; Frederickson, M.; Cleasby, A.; Woodhead, S. J.; Carr, M. G.; et al. Identification of novel p38alpha MAP kinase inhibitors using fragment-based lead generation. *J. Med. Chem.* **2005**, *48*, 414–426.

(70) Liu, Y.; Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* **2006**, *2*, 358–364.

(71) Schroeder, G. M.; Chen, X. T.; Williams, D. K.; Nirschl, D. S.; Cai, Z. W.; Wei, D.; Tokarski, J. S.; An, Y.; Sack, J.; Chen, Z.; Huynh, T.; Vaccaro, W.; Poss, M.; Wautlet, B.; Gullo-Brown, J.; Kellar, K.; Manne, V.; Hunt, J. T.; Wong, T. W. Identification of pyrrolo[2,1-*f*][1,2,4]triazine-based inhibitors of Met kinase. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1945–1951.

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

(73) Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. Adapting Structure-Based Drug Design in the Paradigm of Combinatorial Chemistry and High Throughput Screening: An Overview and New Examples with Important Caveats for Newcomers in Combinatorial Library Design Using Pharmacophore Models or Multiple Copy Simultaneous Search Fragments. *Rational Drug Design: Novel Methodology and Practical Applications*; American Chemical Society: Washington, DC, 1999.

(74) Pierce, A. C.; Rao, G.; Bemis, G. W. BREED: generating novel inhibitors through hybridization of known ligands. Application to CDK2, P38, and HIV protease. *J. Med. Chem.* **2004**, *47*, 2768–2775.