Bioorganic & Medicinal Chemistry 23 (2015) 6520–6527

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Kinase hinge binding scaffolds and their hydrogen bond patterns

Li Xing [a,*], Jacquelyn Klug-Mcleod [b], Brajesh Rai [b], Elizabeth A. Lunney [c]

[a] Pfizer Worldwide Research and Development, 200 Cambridge Park Drive, Cambridge, MA 02140, United States
[b] Pfizer Worldwide Research and Development, 1 Eastern Point Road, Groton, CT 06340, United States
[c] Pfizer Worldwide Research and Development, 10777 Science Center Drive, San Diego, CA 92121, United States

ARTICLE INFO

Article history:
Received 21 May 2015
Revised 24 July 2015
Accepted 8 August 2015
Available online 29 August 2015

Keywords:
Protein kinase
Kinase hinge
Hydrogen bonds
Hinge scaffolds
Potency and selectivity
Crystal structure database

ABSTRACT

Protein kinases constitute a major class of intracellular signaling molecules, and describe some of the most prominent drug targets. Kinase inhibitors commonly employ small chemical scaffolds that form hydrogen bonds with the kinase hinge residues connecting the N- and C-terminal lobes of the catalytic domain. In general the satisfied hydrogen bonds are required for potent inhibition, therefore constituting a conserved feature in the majority of inhibitor-kinase interactions. From systematically analyzing the kinase scaffolds extracted from Pfizer crystal structure database (CSDb) we recognize that large number of kinase inhibitors of diverse chemical structures are derived from a relatively small number of common scaffolds. Depending on specific substitution patterns, scaffolds may demonstrate versatile binding capacities to interact with kinase hinge. Afforded by thousands of ligand–protein binary complexes, the hinge hydrogen bond patterns were analyzed with a focus on their three-dimensional configurations. Most of the compounds engage H6 NH for hinge recognition. Dual hydrogen bonds are commonly observed with additional recruitment of H4 CO upstream and/or H6 CO downstream. Triple hydrogen bonds accounts for small number of binary complexes. An unusual hydrogen bond with a non-canonical H5 conformation is observed, requiring a peptide bond flip by a glycine residue at the H6 position. Additional hydrogen bonds to kinase hinge do not necessarily correlate with an increase in potency; conversely they appear to compromise kinase selectivity. Such learnings could enhance the prospect of successful therapy design.

1. Introduction

The protein kinases encoded in the human genome are involved in regulating essentially all aspects of cellular processes, from metabolism and cell cycle progression to differentiation and apoptosis. Small molecule kinase inhibitors play promising roles in treating a number of diseases including inflammation, neurodegenerative disorders, cardiovascular diseases and cancer.[1] Hence they constitute a major component of the pipelines of drug discovery research and development in pharmaceutical industry. At the moment most of the approved kinase inhibitors for clinical use are in oncology, with Tofacitinib being the only marketed therapy for rheumatoid arthritis.

Protein kinase enzymes phosphorylate the hydroxyl-containing amino acids (e.g. serine, threonine or tyrosine) under sequence-specific contexts. Despite the large variety of cellular functions associated with protein kinases, their catalytic domains are structurally conserved. The approximately 300 amino acids of the kinase domain are structurally segregated into two distinct lobes. The smaller N-terminal lobe consists of mainly beta sheets and one conserved alpha helix (C-helix). The larger C-terminal lobe is mostly helical and contains the activation segment, which includes residues that in many kinases are phosphorylated upon kinase activation. The hinge region connects the two lobes. ATP binds in the folding cleft located between the two lobes forming conserved hydrogen bond interactions with the hinge. Protein kinases catalyze the transfer of the gamma phosphate group from ATP to the recipient substrate. Targeting the ATP binding site in drug design has led to discovery of numerous inhibitor templates that can bind to members of the kinase family.

Protein kinases themselves can be activated upon phosphorylation, often on the activation loop. X-ray crystal structures have elucidated that catalytically active kinases share almost identical conformations, whereas unactivated kinases have diverse structures. Two frequently observed inactive conformations expose additional pockets that are accessible to kinase inhibitors, commonly termed DFG-out and C-helix-out.[2,3] Partially occupying

Abbreviations: KD, kinase domain; CSDb, crystal structure database; HB, hydrogen bond; RMSD, root mean squared distance; P-loop, phosphate-binding loop; 3D, three-dimensional.

* Corresponding author. Tel.: +1 (617) 665 5369; fax: +1 (617) 665 5575.
E-mail address: li.xing@pfizer.com (L. Xing).

http://dx.doi.org/10.1016/j.bmc.2015.08.006
0968-0896/

the active site, the DFG-out conformation directly precludes binding of ATP. The C-helix-out state, on the other hand, can spatially accommodate ATP, but the catalytic machinery is disrupted. With the broken salt bridge between the conserved glutamate and lysine residues, the kinase is rendered inactive, incapable of turning over substrates. Based on surveys of published and Pfizer internal crystallographic data, the structures of over 200 distinct protein kinases have been resolved. Kinases with displaced C-helices or DFG-out conformation of the activation segment each account for approximately 30% of the total structures. Typically compounds containing certain chemical moieties help stabilize the inactive kinase conformation, for example, an urea and/or amide group extending toward the activation loop to facilitate DFG-out, or a hydrophobic aryl in the back pocket to secure the C-helix-out. Some kinases have been shown to naturally rest in the inactive states. For example, the C-helix rotates out in the crystal structure of the apo unphosphorylated BTK-KD.[4] In complex with a number of compact ligands well confined in the adenine pocket, the DFG motif and activation loop of cFMS are in the canonical inactive conformation.[5] Furthermore, a few kinases have shown unlimited versatility of adopting both DFG-out and C-helix-out capabilities, including SRC, ABL and CDK6.[6–13] Kinase structures with simultaneous C-helix-out and DFG-out is possibly a rare event given the redundant mechanism of kinase inactivation, but evidently not mutually exclusive as demonstrated by X-ray structures including EGFR, FAK, BMX and PDK1 kinases.[14–17]

For ATP competitive inhibitors, hydrogen bond interaction with the kinase hinge is usually indispensable for potent inhibition, regardless of the kinase conformational state that is targeted. In order to make better decisions to select, design and prioritize multiple chemotypes of choice for lead optimization, understanding the interaction patterns of the primary scaffolds that anchor the compounds via hinge hydrogen bonds is of paramount importance. Recently putative hinge-binding fragments were compiled from decomposing known kinase inhibitors mapping to kinase pharmacophore models.[18] We have previously reported the extraction of kinase hinge binding scaffolds from large body of X-ray structural data.[19] Herein we focus on the statistical analysis of the signature hydrogen bonds in terms of their three-dimensional (3D) properties. With an attempt to understand the relationship between the propensity of the key molecular interactions and their implication for biological activity as well as kinome specificity, the learnings presented here could provide guidance for novel compound design.

2. Methods

2.1. Crystal structure database (CSDb)

Kinase protein/ligand crystal complexes were retrieved from the in-house crystal structure database, a large collection of X-ray structure repository that contains both internally solved structures and those selected and imported from the Protein Data Bank.[20] A total of 3980 crystal structures of kinases in complex with small, drug-like ligands exemplifying a variety of ligand binding modes were pulled from the database. All of them are better than 3.5 Å in resolution, of which 85% are higher resolution structures of $\leqslant$2.5 Å. These data were then utilized to mine and extract the hinge interacting fragments. An important aspect of this database was that all structures from a given family were aligned onto a common frame of reference based on superposition of preselected residues within the ATP site, making the subsequent analysis of the data derived from this database particularly efficient.

The structural diversity of the compounds in the data set was analyzed. When subjected to a nearest-neighbor analysis, the compounds were found to cover different structural series of 670 clusters and 369 singletons, according to Ward's similarities (distance cutoff of 0.1) using CDK-Daylight fingerprints. The diverse representation of the chemical classes attests the data set is not biased by any specific structural template, and the trends identified are generally applicable.

2.2. Extraction of hinge scaffolds

The amino acids on kinase hinges are fully annotated from H1 to H13, with H3 corresponding to the gatekeeping residue.[19] A set of criteria was defined to guide hinge extraction as described previously.[19] Using the kinase annotation table, the ligand atoms that make hydrogen bonds with the hinge residues are identified. The rest of the molecules were then traversed along the connectivity map to complete the extraction of the core structures, including all ring fragments, the hetero-atoms and functional groups directly bonded to the rings containing cyano, nitro, nitroso, carbonyl, amide, urea, carboxylate, sulfone, sulfonamide, sulfoxide, sulfinic and sulfonic acids.[19] The rest of the structural substituents comprising atoms and bonds that were farther away from the hinge were then removed. Hydrogen atoms were added to fill the open valences. The geometric criteria of a hydrogen bond interaction X–H···Y–Z are: X···Y distance <3.4 Å, and both XHY and HYZ angles >90°. In the end, a total of 595 unique hinge scaffolds were identified from the nearly 4000 kinase inhibitors in CSDb.

2.3. Binding modes of hinge scaffolds

Each hinge binding scaffold, defined by a unique structural representation, could orient itself in many different ways in the ATP pocket. Frequently, each altered positioning presents distinct interactions with the kinase hinge. To account for such orientational and positional flexibility, each hinge scaffold is clustered into discrete representative conformations based on their in-place root-mean-squared-distance (RMSD) involving all heavy atoms. If the root-mean-squared-distance (RMSD) is less than or equal to 1 Å the positions are considered degenerative, otherwise a different binding conformation is recorded for the same hinge scaffold.

2.4. Kinase activity and selectivity data

Kinase activity data against targets denoted by X-ray experiments were obtained by querying against Pfizer screening database for the corresponding screens, tagged with IC_{50} as the endpoint type. Multiple IC_{50} measurements for the same target were retained only if they fell within the same order of magnitude and the values were averaged. The final dataset consisted of 1502 ligands with ranging activity values from 0.01 nM to 100 μM.

Kinase selectivity data were extracted from multiple screening panels measured both in-house as well as by commercial vendors, for compounds that appear in CSDb. For each selectivity panel, compounds were tested at either 1 μM and/or 10 μM concentrations at ATP K_m for the corresponding kinase. The Gini coefficient at each screening concentration were computed using the percent inhibition data. With a range between zero and one, the Gini coefficient has been commonly used to describe selectivity of kinase inhibitors since its introduction by Graczyk.[21]

3. Results and discussion

3.1. Hinge binding scaffolds

The protein kinases bind a common endogenous ligand. The ATP binding pocket is situated at the folding cleft of the N- and C-lobes, which are connected by the hinge region. Interactions are formed