



A structural informatics approach to mine kinase knowledge bases

Natasja Brooijmans¹, Dominick Mobilio¹, Gary Walker¹, Ramaswamy Nilakantan², Rajiah A. Denny³, Eric Feyfant³, David Diller⁴, Jack Bikker¹ and Christine Humblet⁴

¹ Computational Chemistry and Cheminformatics, Wyeth Research, 401 North Middletown Road, Pearl River, NY 10965, USA

² 95 Birch Street, Closter, NJ 07624, USA

³ Computational Chemistry and Cheminformatics, Wyeth Research, 200 Cambridgepark Drive, Cambridge, MA 02140, USA

⁴ Computational Chemistry and Cheminformatics, Wyeth Research, 865 Ridge Road, Monmouth Junction, NJ 08852, USA

In this paper, we describe a combination of structural informatics approaches developed to mine data extracted from existing structure knowledge bases (Protein Data Bank and the GVK database) with a focus on kinase ATP-binding site data. In contrast to existing systems that retrieve and analyze protein structures, our techniques are centered on a database of ligand-bound geometries in relation to residues lining the binding site and transparent access to ligand-based SAR data. We illustrate the systems in the context of the Abelson kinase and related inhibitor structures.

Introduction

Structure-based design approaches have become an integral component of drug design, leading to many successes in drug optimization, as illustrated by a plethora of medicinal chemistry publications in recent issues of the *Journal of Medicinal Chemistry*. The underlying computational capabilities have progressed considerably in the past few decades, leading to an impressive arsenal of modeling tools to empower structure-based design. Examples include methods of aligning and comparing protein sequences guided by the 3D structure details, techniques for assembling homology-based models, approaches to characterizing and comparing binding sites, and docking methods to predict protein–ligand complex structures. Aided by the large number of protein structures now available in the Protein Data Bank (PDB) [1,2], as well as industrial proprietary databases, computational chemists can leap into a wealth of structural information, based on either the protein or the ligand data.

Although structures can be readily retrieved from the PDB through several different queries, it is generally not straightforward to perform more in-depth analyses on a large scale across all protein–ligand complexes in the PDB. In addition, links to bioactivity data or protein-family-specific information are not available. With these limitations in mind, additional databases have been developed that use the information from the PDB but add additional layers of

features. Three different approaches have been used to build a variety of protein-structure-focused databases. The first approach focuses on protein–ligand interactions. Several of these databases – for example, Relibase [3] and Credo [4] – enable distance queries between the ligand and protein, which will aid in gaining a further understanding of driving forces of molecular recognition. Other protein–ligand-based databases focus on linking ligand structures to information on the protein domain they bind to [5,6]. A second approach links binding affinities to protein–ligand complex structures [7–10], which can give insights into structure–activity relationships (SARs). The third approach focuses on structures from a particular protein family and enhances the structural information available from the PDB with specific annotations that enable querying for specific features unique to the family [11–14], enabling more in-depth structural analyses across the family.

No databases have appeared in the literature that connect all three approaches to enhancing the information in the PDB, although the SARfari database (originally developed at Inpharmatica Inc and now available through EMBL-EBI) connects the second and third approach by linking activity data on kinase inhibitors to their protein structures and sequences [15]. Sequence alignments and overlays of kinase catalytic domain structures are also available. Rather than integrating only data on the co-crystallized ligands, binding data are available for all published inhibitors from selected journals through SARfari (<http://www.sarfari.org/kinasesarfari>).

Corresponding author: Humblet, C. (christine.humblet@comcast.net)

In this paper, we describe a combination of structural informatics approaches that enable protein–ligand interaction queries and protein-family-specific queries for kinases and matrix metalloproteases. Examples of issues that can now be addressed include (i) which kinases present a cysteine residue in the active site, (ii) which inhibitors engage in a hydrogen-bond interaction with a specific residue in the hinge, (iii) what the known conformations and geometries of helix C are and which inhibitor structures appear to induce a particular conformation for that helix, (iv) what can be learned about protein flexibility in relation to known ligand-bound structures, and (v) how a particular inhibitor series compares with known inhibitor structures and if or when inhibitor-bound structures are available, what molecular design concepts can inspire the next design cycles, and so on.

In addition, we have enabled additional queries based on the ligand structures to retrieve structure–activity data through the GVK Bio database (<http://www.gvkbio.com/informatics.html>) and our corporate database, effectively combining all three approaches outlined above. Our proprietary techniques are centered on a database of ligand-bound geometries in relation to the residues lining the binding site and, thus, give transparent access to ligand-based SAR data. We illustrate the systems in the context of protein kinase knowledge bases (KKBs).

Databases overview

Protein relational database

To mine the PDB structures for ligand geometries and other features not available for searching at the Research Collaboratory for Structural Bioinformatics (RCSB) website [2], we included new,

enabling dimensions in the design of a protein relational database (PRDB) [16], as described in Fig. 1, which highlights the ability to retrieve crystal structures based on protein synonyms or details of the structure or through ligand substructure searches (similar to what is available through the PDB). However, structures can also be retrieved based on protein–ligand atom distances and protein residue triplet distances. Figure 1 also clearly illustrates that the KKB is merely an extension of PRDB; all PRDB querying capabilities are available, but several kinase-specific queries have been made possible as well. The figure further illustrates that because each ligand has a 1D ligand identifier assigned to it, additional data can be easily retrieved from internal and external databases, usually containing biological assay data.

PRDB was populated with data downloaded from the PDB and additional layers of data derived from a variety of geometric calculations, enabling the user to construct a wide range of powerful queries. For example, one can not only search for PDB files containing a known kinase inhibitor such as staurosporine (a query that can readily be done via the PDB web site) but also conduct much more specific searches by including ligand–protein interatomic distances and locate PDB files meeting the criteria (i) ligand <2.6 Å from any leucine atom, (ii) ligand <2.6 Å from leucine's nitrogen, (iii) any oxygen in ligand <2.6 Å from leucine's nitrogen and (iv) ligand's carbonyl oxygen <2.6 Å from leucine's nitrogen.

Protein kinase knowledge base

Drug discovery efforts against kinase targets have mainly focused on the catalytic (or ATP-binding) domain [17,18]. Although

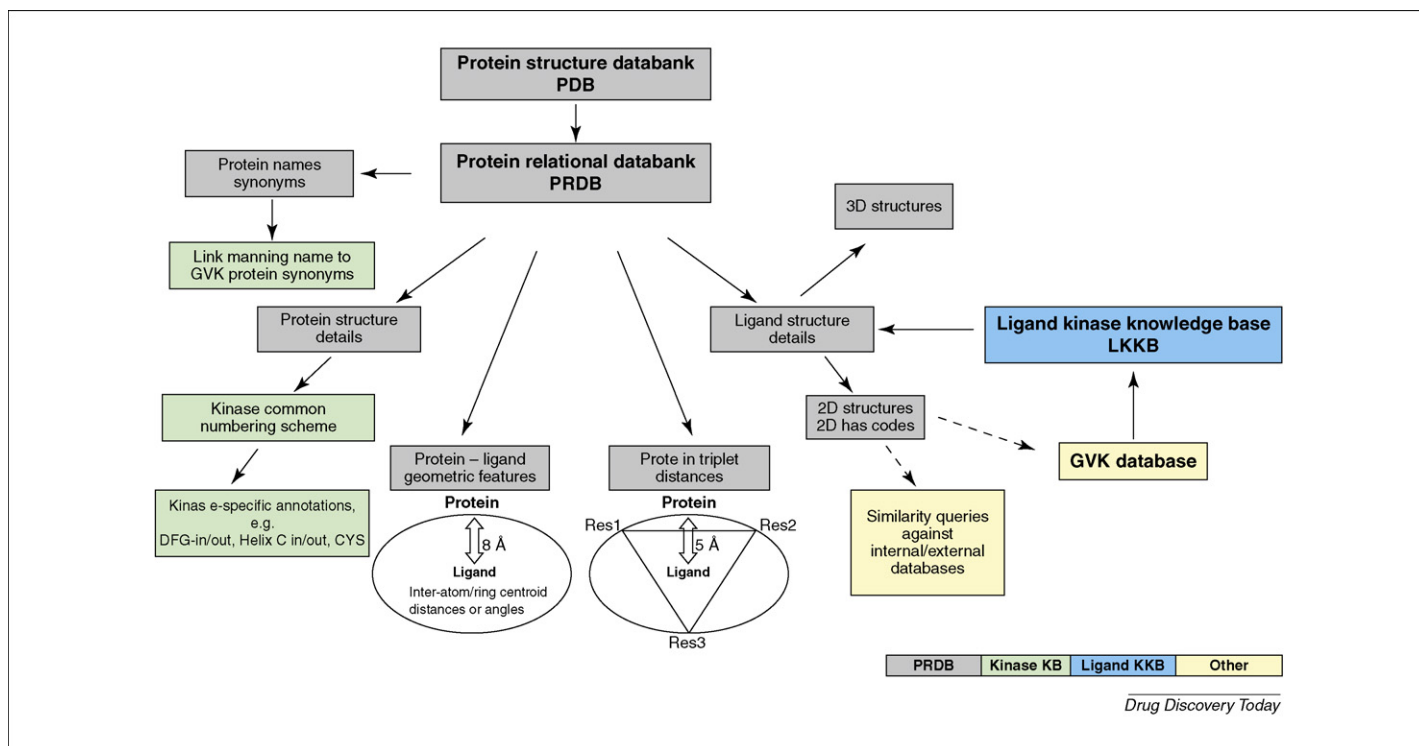


FIGURE 1

Schematic of PRDB, KKB and LKKB. Summary of features stored in PRDB (gray background), KKB (green background), LKKB (blue background) and other databases (yellow background) enabling a large number of queries through ADOpt. The dashed lines to the systems highlighted in yellow indicate that integration to these internal and external databases with binding and other biological data is through a different interface.

several different types of inhibitors have been identified, including inactive form inhibitors [19] and allosteric inhibitors [20–23], they all interact with the catalytic domain of the kinase. To enhance internal access to kinase crystal structures and enable structural comparisons of catalytic domain features, we extended the PRDB to include a structural KKB [16,24] that contains all catalytic domain structures deposited in the PDB (<http://www.rcsb.org>) [2].

The KKB was constructed from the human catalytic domain sequences collated by Manning *et al.*, which can be accessed through <http://www.kinase.com> [25]. The sequences were aligned and their associated PDB files were retrieved from the PDB. The matching chains were extracted from the parent PDB files and placed into a common numbering scheme. The 3D structures were then aligned and placed into a common reference frame. Finally, for each chain, two versions of PDB files were generated using the original and common residue numbers.

To help locate specific proteins, the name used by Manning *et al.* for each kinase was assigned to each of the PDB files generated. These names were linked to a list of synonyms from the GVK Bio Database, enabling easy access to the structures based on several different names or name fragments.

With the kinase chains individually searchable, we sought to add considerable value to the database by mapping important structural features onto the commonly numbered structures. The kinase-specific metadata generated includes automated structural classification schemes for Asp-Phe-Gly (DFG)-in or out and Helix C in or out conformations [24]. Reactive cysteines in the active site, as compiled by Gray and co-workers, were annotated as well [26]. Furthermore, hinge region residues, the gatekeeper residue, glycine-rich loop residues and DFG- and His-Arg-Asp (HRD)-motif residues were all annotated, in addition to active site residues.

The resulting KKB enhances the structural information present in the PDB in several ways, making the retrieval of structures of interest and structural comparison across the human kinome possible. For example, because of the use of different synonyms for protein names (e.g. Abl kinase, Abelson kinase and c-Abl) by different authors, retrieval of all structures for a particular protein target is not trivial and often requires a large amount of manual analysis of search results. In addition, in lead optimization, understanding of the structural features of the binding site of interest alone is not enough. To understand selectivity data, identify potential off-targets or enhance selectivity using structure-based design, easy access to binding site comparisons and the ability to identify kinase structures with particular residues at particular positions in the active site is important. The KKB enables these kinds of queries and makes structural analyses trivial.

Ligand kinase knowledge base

PRDB and KKB were augmented with our ligand kinase knowledge base (LKKB), extracted from the GVK database, which consists of structures for 338,000 kinase inhibitors with published biological data from either the literature or patents. These databases have been made available for searching and analysis through our desktop tool ADOpt (analyze, design and optimize) (unpublished data). In addition, the full GVK database can be mined with other data access tools.

Analyses

Abelson kinase crystal structures

We illustrate the utility of enhanced structural databases and various desktop visualization and analysis systems in the context of the Abelson kinase (Abl). The first kinase inhibitor approved for marketing, Gleevec, targets Abl. Many crystal structures have been determined, offering a broad coverage of ligand structure diversity, as well as a large variety of kinase-bound structural forms.

In structure-based design practices, an analysis of a known drug target can reasonably start with a review of ligands that have been crystallized with the target. In the RCSB PDB, a simple text-based query on the keyword 'Abl' results in 56 matches (July 2009). A more careful analysis of the structures retrieved reveals that only 27 hits contain the Abl catalytic domain, 2 contain the Abl2 catalytic domain, and 9 contain Abl SH2 and SH3 structures. The other 22 hits are of other kinase domains or of structures that are completely unrelated to Abl or kinases in general.

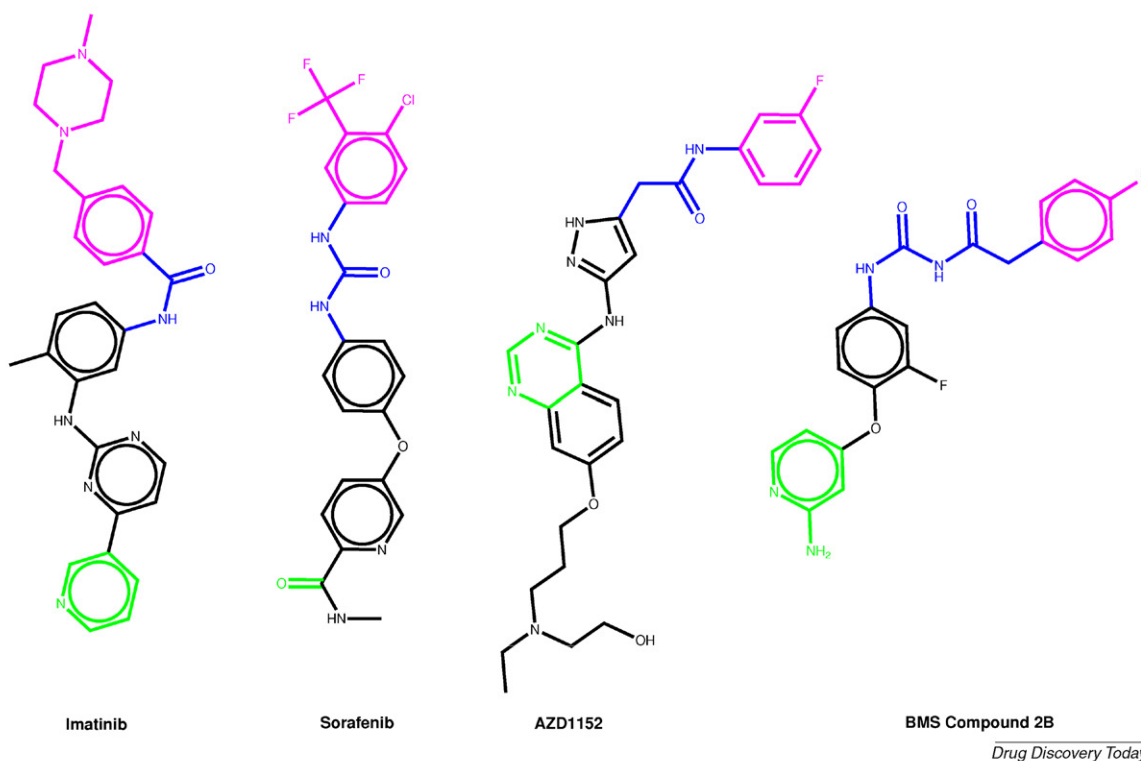
By contrast, in PRDB, a similar query with a Manning name of 'ABL' results in the retrieval of 56 chains from 27 parent PDB files of ligands bound to the kinase catalytic domain. The two Abl2 catalytic domains are correctly associated with the 'ARG' gene, which is the Manning name for Abl2. Furthermore, the readily available annotations of particular kinase-related structural features immediately highlight that 8 Abl structures were crystallized in a DFG-in conformation, whereas the remaining 19 were crystallized in a DFG-out conformation. The structures of the co-crystallized ligands are shown in Figure S1 (see Supplementary Data online).

DFG-in or out conformations

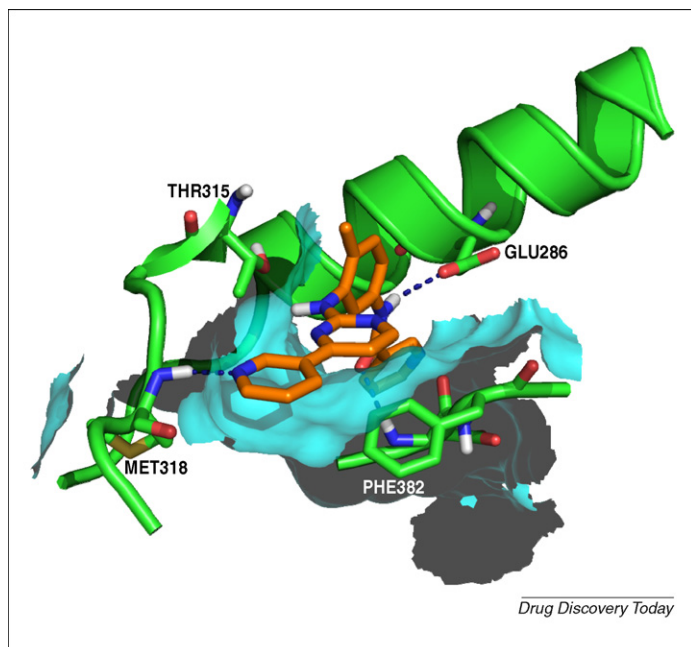
Upon inspection of this list of structures, it is observed that compounds of the same chemical series, the pyrrolopyridines (Figure S1, highlighted structures) can bind to both the DFG-in and the DFG-out form of Abl. The known pharmacophore of DFG-out compounds consists of two (hetero)aromatic rings connected by a 2-, 3- or 4-atom linker containing an amide or urea [19,26]. Known DFG-out inhibitors with the common pharmacophore include imatinib (1FPU.pdb [27]), sorafenib (1UWH.pdb [28]), AZD1152 (a hypothesized DFG-out inhibitor) [29] and Compound 3b (3CTH.pdb [30]) shown in Fig. 2. The terminal ring of the DFG-out headpiece accesses the large hydrophobic pocket that is created by the DFG-out conformation, as described in Fig. 3. This common pharmacophore has been used successfully to rationally design novel DFG-out inhibitors and identify kinases that can undergo the DFG-in or out conformational change [19,31].

The pyrrolopyridine compounds found to bind to both the DFG-in and the DFG-out conformations of Abl do not adhere to this common pharmacophore and do not at all occupy the hydrophobic pocket generally occupied by DFG-out inhibitors, indicating that Abl can readily access the DFG-out conformation. The KKB was used to investigate this observation further and to see what other kinases have been crystallized in the DFG-out form with ligands that do not conform to the common DFG-out pharmacophore. All DFG-out structures could be readily retrieved through the ADOpt interface to the KKB, and a total of 30 kinases were found to have crystallized in the DFG-out form.

Subsequently, X-ray structures with ligands that adhere to the common pharmacophore were removed through the substructure searches shown in Figure S2. This left 31 unique crystal structures

**FIGURE 2**

Examples of known DFG-out inhibitors. The hinge interaction motif is shown in green, the DFG-out linker is shown in blue, and the hydrophobic group that interacts with the hydrophobic DFG-out pocket is shown in magenta.

**FIGURE 3**

Close-up of imatinib in complex with Abl (1FPU.pdb) highlighting the deep hydrophobic pocket present in DFG-out kinase structures. Hydrogen-bonding interactions with the hinge region and between the amide linker of the DFG-out headpiece and the backbone of the DFG-motif are shown. Thr315 is the gatekeeper residue.

covering 11 kinases, in addition to Abl, that contain non-traditional DFG-out compounds but were found to be in the DFG-out conformation nonetheless. Interestingly, several of the structures are bound to ATP analogs and the pan-kinase (active form) inhibitor staurosporine (Figure S3). This structural analysis thus indicates that several kinases can readily undergo the DFG-in or out conformational change without the need for a large hydrophobic group from an inhibitor to occupy this pocket.

Ligand activity data

Although access to structural data alone can clearly give rise to interesting observations and can answer questions that can subsequently be used in prospective structure-based drug design cycles, access to ligand activity data and the ability to connect it to structural data enables the investigation of even more sophisticated questions and hypotheses. Thus, a typical lead optimization project might start with the identification of potent and/or highly efficient known inhibitors to gain an understanding of the SAR of the target.

The LKKB was used to retrieve all known compounds that report binding data to ABL kinase. These were imported into ADOpt. Ligand efficiencies were calculated and two visually similar series were identified with ligand efficiencies higher than 0.35. The scaffolds of these two series are both 6-5 hetero-aromatic systems. Series 1 contains a pyrrolopyrimidine core, and series 2 contains a pyrazolopyrimidine core (Fig. 4).

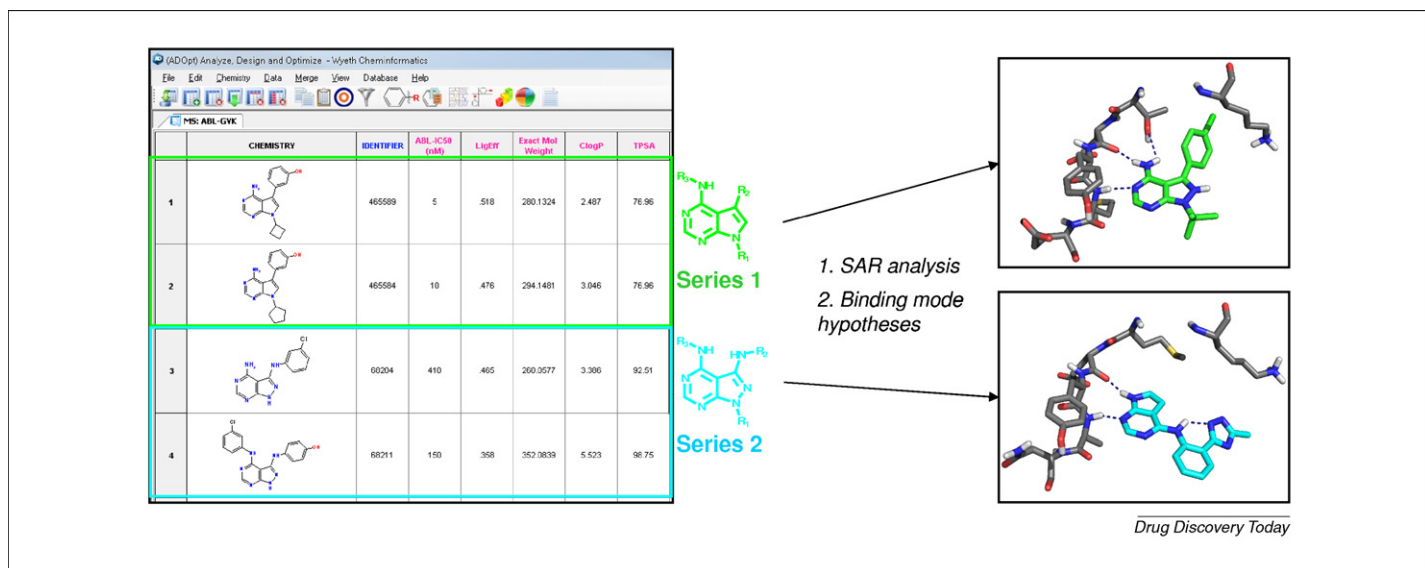
**FIGURE 4**

Diagram of SAR analysis with ADOpt to identify the most efficient Abl inhibitors. Series identification followed up by confirmation of the binding mode hypotheses based on an analysis of the available SAR from the GVK Medchem database. The top panel on the right-hand side shows the putative binding mode of series 1 (1QPE.pdb); the bottom panel shows the binding mode for series 2 (3CQU). Hydrogen-bonding interactions to the hinge region are shown, and a comparison of the two panels highlights the flipping of the 5-6 hetero-aromatic core.

An initial hypothesis might be that these two cores would bind in the same manner to the kinase domain. Using the R-group analysis capabilities of ADOpt, the SAR of both series was analyzed. This highlighted that series 1 is intolerant to substitution at R3; even methyl is detrimental to activity. This suggests the amino group is involved in hydrogen bonding to the enzyme, possibly the hinge region. Furthermore, R2 strongly prefers aromatic groups, suggesting an interaction with the kinase selectivity pocket. R1 is tolerant to a large variety of substituents, which might indicate its accessibility to solvent.

The same analysis for series 2 highlights a different pattern. In this case, R3 is tolerant to substitution and can even increase potency. In contrast to series 1, R1 in series 2 is not tolerant to substitution, suggesting that the $-NH$ group is involved in hydrogen bonding with the Abl hinge region. The SAR analysis thus strongly suggests that series 1 and series 2, although highly similar, have different binding modes in Abl.

To confirm this hypothesis, substructure searches were done against the KKB for 6-5 hetero-aromatic ring systems. The binding mode for series 1 is probably similar to that found in a Lyn kinase structure co-crystallized with a pyrazolopyrimidine inhibitor (1QPE [32]) (Fig. 4). This structure shows that R3 interacts with a backbone carbonyl group of the hinge region, R2 interacts with the specificity pocket, and R1 interacts with the ribose pocket. This is all consistent with the available SAR for series 1 with Abl.

The binding mode for series 2 is probably similar to that of a pyrrolopyrimidine structure found in AKT (3CQU [33]) (Fig. 4). In this structure, the $-NH$ (R1) and the adjacent pyrimidine nitrogen interact with the hinge region, R2 interacts with the selectivity pocket, and R3 interacts with the ribose pocket. Again, this is consistent with the available SAR.

The combination of SAR analysis and access to structural information has thus led to reasonable hypotheses for the binding modes of these two series. This information can be further used to

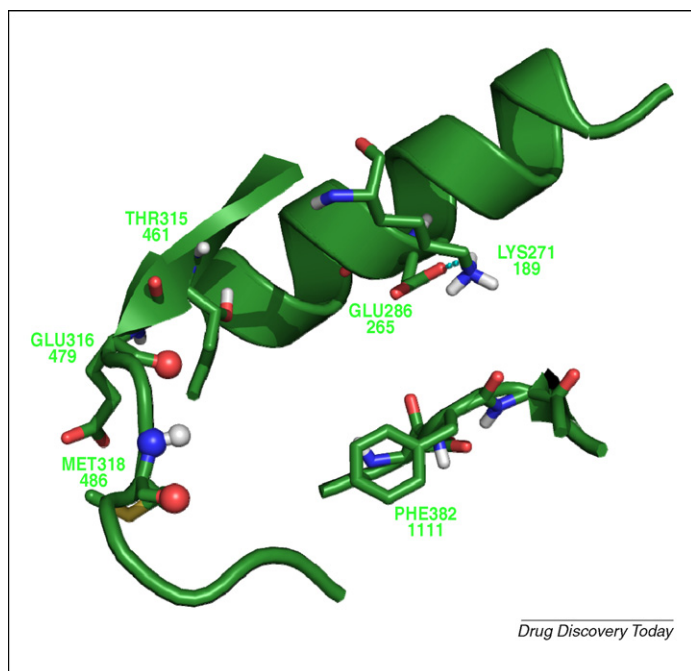
assess optimization potential in terms of potency or selectivity or to transfer the SAR to other (in-house) series.

Hinge region binding features

The molecular interaction with the hinge region of the kinases is another area of focus in structure-based design. The minimal requirement for ATP-competitive kinase inhibitors includes a single hydrogen-bonding interaction with the backbone NH of the central residue in the hinge region (Met318 in Abl) (Fig. 5). Often, an additional hydrogen-bonding interaction is observed for many kinase inhibitors, either with the backbone carbonyl of the same residue (Met318 in Abl) or with the backbone carbonyl of a more buried hinge region residue (Glu316 in Abl). By combining data from the KKB with activity data from the GVK database, the most efficient hinge region interaction motifs can be identified. This information can then be used in the prioritization of HTS hits and in the design of novel cores for hinge region binding.

To answer the question of whether one or two hydrogen-bonding interactions with the hinge give the most efficient inhibitors, ligand structures were retrieved that interact with the hinge region through the backbone NH of common residue 486 (Met318 in Abl) and optional interactions with the backbone carbonyl of common residue 486 or 479 (Glu316 in Abl) (Fig. 5).

Structures that form two hinge region hydrogen bonds were separated into those that form both hydrogen bonds to the same residue (486) or to the two different hinge region residues (486 and 479). For ligands retrieved from the KKB through these searches, known activity data were compiled from the GVK databases. Similar to the SAR analysis data highlighted above, ligand efficiencies were calculated for the lowest IC_{50} value observed for each ligand. Only ligands with measured $IC_{50}s \leq 10 \mu M$ were retained. Average ligand efficiencies were calculated for molecular weight bins of 200–300, 300–400 and 400–500 for compounds forming a

**FIGURE 5**

Close-up of Abl-binding site showing the hinge region (left-hand side) and the atoms that can be involved in hydrogen-bonding interactions (sphere representation). Also shown are Helix C, the conserved Helix C acidic residue, the gatekeeper residue, the conserved catalytic lysine and the DFG-motif residues. The Abl residue names and numbers are shown in green with the corresponding residue numbers in the common numbering scheme shown in orange.

single hinge region hydrogen bond, two hydrogen bonds to the same residue or two hydrogen bonds to two separate residues.

The data are summarized in Table S1 and suggest that for fragment-type hits (200–300 MW), compounds that form a single hinge interaction are less efficient than compounds forming two hinge hydrogen bonds. The data also suggest that the additional hydrogen bond to the back of the binding site (common residue number 479) enhances potency more than the one to residue number 486. The same trend is seen for the drug-like molecules in the molecular weight range of 400–500. Somewhat surprisingly, at the lead-like stage of 300–400 MW, the cores are all equally efficient.

Analysis of all ligands that have a ligand efficiency ≥ 0.50 results in 19 ligands, of which only four form a single hydrogen bond, ten form two hydrogen bonds to residue 486 (backbone NH) and 479 (backbone carbonyl), and five form two hydrogen bonds with the same residue (486). This further strengthens the observations that two hydrogen-bond interactions with the hinge region result in more potent compounds than ones that only form a single hydrogen bond and that the more buried hydrogen-bonding

interaction that can be formed with common residue 479 is preferred over the interaction with the carbonyl of common residue 486 (Figure S4), although the analysis is complicated by the fact that the binding data originate from different laboratories and various kinases. Further research would be required to confirm these observations.

Concluding remarks

We have described how various structural informatics approaches can be applied to facilitate molecular analyses and drug design in the context of KKBs with a particular focus on Abl kinase. Whereas a variety of protein databases, such as PDB [1,2] and others [34,35], emphasize structural details primarily derived from crystallography results, the databases introduced here – PRDB and KKB – are particularly tailored to rapidly compare and contrast the geometric features pertinent to the ligands and their binding sites. Although these capabilities are shared with the Relibase system [3,36], PRDB and KKB add layers of detailed annotations derived from structural aspects specific to the protein kinase family. We have given a few examples to illustrate how this structural informatics-driven data mining can empower drug design objectives. In addition, we present additional layers derived from having integrated and readily available access to the combination of internal and external ligand SAR data repositories.

Although the protein kinase family is the first that has been integrated in our system, other protein target families can be added readily. Indeed, the geometric triplet component of PRDB, not illustrated in the Abl kinase analyses, provides an option to analyze the enzymatic catalytic triads. One can also envision implementations that are more suitable for analyzing particular structural motifs, such as helical bundles or protein–protein interfaces.

Although powerful, the current versions of PRDB and KKB are dependent on the availability of relevant crystal structures, and this can be limiting when attempting to assess target selectivity aspects. In its next generation, PRDB could benefit from a full integration of the structural information with the sequence information.

Acknowledgements

The authors thank our collaborators Kowticar, J. Mankala, S. Palli, S. Punyamantula, M. Thatipally and R. John at GVK Bio for their participation in the development of the structural informatics systems described in this manuscript. N.B. thanks Wayne Chang for initial implementation of algorithms to derive the KKB.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.drudis.2009.11.005.

References

- Berman, H.M. *et al.* (2000) The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242
- Deshpande, N. *et al.* (2005) The RCSB Protein Data Bank: a redesigned query system and relational database based on the mmCIF schema. *Nucleic Acids Res.* 33 (Database issue), D233–D237
- Hendlich, M. *et al.* (2003) Relibase: design and development of a database for comprehensive analysis of protein–ligand interactions. *J. Mol. Biol.* 326, 607–620
- Schreyer, A. and Blundell, T. (2009) CREDO: a protein–ligand interaction database for drug discovery. *Chem. Biol. Drug Des.* 73, 157–167
- Bashton, M. *et al.* (2008) PROCOGNATE: a cognate ligand domain mapping for enzymes. *Nucleic Acids Res.* 36 (Database issue), D618–D622
- Chalk, A.J. *et al.* (2004) PDBLIG: classification of small molecular protein binding in the Protein Data Bank. *J. Med. Chem.* 47, 3807–3816

- 7 Hu, L. *et al.* (2005) Binding MOAD (Mother of All Databases). *Proteins* 60, 333–340
- 8 Liu, T. *et al.* (2007) BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Res.* 35 (Database issue), D198–D201
- 9 Li, L. *et al.* (2008) PDBcal: a comprehensive dataset for receptor–ligand interactions with three-dimensional structures and binding thermodynamics from isothermal titration calorimetry. *Chem. Biol. Drug Des.* 71, 529–532
- 10 Block, P. *et al.* (2006) AffinDB: a freely accessible database of affinities for protein–ligand complexes from the PDB. *Nucleic Acids Res.* 34 (Database issue), D522–D526
- 11 Rawlings, N.D. *et al.* (2008) MEROPS: the peptidase database. *Nucleic Acids Res.* 36 (Database issue), D320–D325
- 12 Allcorn, L.C. and Martin, A.C. (2002) SACS – self-maintaining database of antibody crystal structure information. *Bioinformatics* 18, 175–181
- 13 Jacobs, M.D. *et al.* (2008) Classifying protein kinase structures guides use of ligand-selectivity profiles to predict inactive conformations: structure of lck/imatinib complex. *Proteins* 70, 1451–1460
- 14 Lee, S. and Blundell, T.L. (2009) BIPA: a database for protein–nucleic acid interaction in 3D structures. *Bioinformatics* 25, 1559–1560
- 15 Houlton, S. (2008) Wellcome boost for open-access chemistry. *Nat. Rev. Drug Discov.* 7, 789–790
- 16 Mobilio, D. *et al.* A protein relational database and protein family knowledge bases to facilitate structure-based design analyses. Submitted
- 17 Huse, M. and Kuriyan, J. (2002) The conformational plasticity of protein kinases. *Cell* 109, 275–282
- 18 Bikker, J.A. *et al.* (2009) Kinase domain mutations in cancer: implications for small molecule drug design strategies. *J. Med. Chem.* 52, 1493–1509
- 19 Liu, Y. and Gray, N.S. (2006) Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* 2, 358–364
- 20 Akella, R. *et al.* (2008) Unique MAP kinase binding sites. *Biochim. Biophys. Acta* 1784, 48–55
- 21 Burke, J.R. *et al.* (2003) BMS-345541 is a highly selective inhibitor of I kappa B kinase that binds at an allosteric site of the enzyme and blocks NF-kappa B-dependent transcription in mice. *J. Biol. Chem.* 278, 1450–1456
- 22 Pargellis, C. *et al.* (2002) Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nat. Struct. Biol.* 9, 268–272
- 23 Ohren, J.F. *et al.* (2004) Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nat. Struct. Mol. Biol.* 11, 1192–1197
- 24 Brooijmans, N. *et al.* Development of a kinase structural database to enhance structure-based analyses and drug discovery. Submitted
- 25 Manning, G. *et al.* (2002) The protein kinase complement of the human genome. *Science* 298, 1912–1934
- 26 Zhang, J. *et al.* (2009) Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* 9, 28–39
- 27 Schindler, T. *et al.* (2000) Structural mechanism for STI-571 inhibition of Abelson tyrosine kinase. *Science* 289, 1938–1942
- 28 Wan, P.T.C. *et al.* (2004) Mechanism of activation of the RAF–ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* 116, 855–867
- 29 Tao, Y. *et al.* (2007) Enhancement of radiation response in p53-deficient cancer cells by the Aurora-B kinase inhibitor AZD1152. *Oncogene* 27, 3244–3255
- 30 Cai, Z.-W. *et al.* (2008) Discovery of orally active pyrrolopyridine- and aminopyridine-based Met kinase inhibitors. *Bioorg. Med. Chem. Lett.* 18, 3224–3229
- 31 Choi, Y. *et al.* (2009) Discovery and structural analysis of Eph receptor tyrosine kinase inhibitors. *Bioorg. Med. Chem. Lett.* 19, 4467–4470
- 32 Zhu, X. *et al.* (1999) Structural analysis of the lymphocyte-specific kinase Lck in complex with non-selective and Src family selective kinase inhibitors. *Structure* 7, 651–661
- 33 Lippa, B. *et al.* (2008) Synthesis and structure based optimization of novel Akt inhibitors. *Bioorg. Med. Chem. Lett.* 18, 3359–3363
- 34 Szabadka, Z. and Grolmusz, V. (2006) Building a structured PDB: the RS-PDB Database. In *Proceedings of the 28th IEEE EMBS Annual International Conference* pp. 5755–5758
- 35 Szabadka, Z. and Grolmusz, V. (2007) High throughput processing of the structural information in the protein data bank. *J. Mol. Graph. Model.* 25, 831–836
- 36 Gunther, J. *et al.* (2003) Utilising structural knowledge in drug design strategies: applications using Relibase. *J. Mol. Biol.* 326, 621–636