Successful Computational Prediction of the Structure-Activity Relationship of a Potent JAK2 Inhibitor

Herbert R. Treutlein, Michelle Styles, and Jun Zeng

Abstract—JAK2 is a member of the JAK family of protein tyrosine kinases (PTK). JAK2 is an important intracellular mediator of cytokine signaling and its mutants are involved in haematological cancers and other chronic diseases. The specificity of inhibitors targeting the JAK2 PTK domain, particularly over other JAK kinases, is a critical design element of clinically relevant drugs. Due to the considerable structural similarity amongst the over 500 kinases in the human genome, designing a highly specific inhibitor of just one kinase is a nontrivial task. We have performed molecular docking studies of a high-affinity pan-JAK inhibitor (tetracyclic pyridone 2-tert-butyl-9-fluoro-3,6-dihydro-7H-benz[h]-imidaz[4,5-f]isoquinoline-7-one) into a JAK2 PTK structure using our in-house software ChemaPhore™. Lacking an X-ray structure, the studies were based initially on a homology model of JAK2. In order to improve the specificity and potency of the inhibitors, a series of modifications to the pan-JAK inhibitor were suggested and the new compounds were then synthesized and biologically tested. The results of the structure-activity relationship (SAR) obtained from wet screening are consistent with the proposed binding mode of the pan-JAK inhibitor. More importantly, recent X-ray crystal structures of JAK2 inhibitor complexes are consistent with our JAK2 homology model and confirm our predictions.

Index Terms—**Computational Drug Design, Jak Kinase, Target-based Design, Virtual Screening.**

I. Introduction

JAK kinases are a family of four multi-domain intracellular proteins which play a significant role in cellular signaling: JAK1, JAK2, JAK3 and Tyk2 [1-3]. JAK kinases consist of 7 domains, one of which is a functional tyrosine-kinase domain (JH1), i.e. an enzyme that transfers a phosphate group from adenosinetriphosphate (ATP) to a tyrosine side chain of a protein substrate. This phosphorylation is a major regulatory mechanism for intracellular signal transduction and makes protein kinases an attractive target for drug design. In particular, JAK2 and JAK3 have been shown to be of particular therapeutic interest. Cytopia's interests in JAK kinases are focused on the following areas

- *Myeloproliferative diseases (MPDs):* Diseases including polycythemia vera (PV) as well as essential thrombocythemia (ET) are relatively rare, but highly debilitating. JAK2, through spontaneous mutation (V617F), is thought to be responsible for approximately 90% of cases of PV.
- *Cancer:* JAK2 related pathways are known to play a role in the proliferation of cancer cells, such as haematological malignancies (e.g. acute lymphoblastic leukaemia) and lymphomas.
- *Cardiovascular diseases*: Pulmonary arterial hypertension and other cardiovascular diseases such as congestive heart failure.
- *Immunology and inflammatory diseases:* Many cytokine receptors, critical for the activity of immune cells, signal through JAK3. Unlike its family members, JAK3 is only expressed in lymphoid cells. JAK3 is an attractive target for the development of orally active, small molecule therapeutics for the prevention of transplant rejection amongst other diseases.

The focus of current kinase inhibitor design is to find a small molecule to replace ATP in its native binding pocket located in the kinase domain and thereby blocking the enzyme activity of the protein.

Our internal drug discovery engine, which includes computational drug design, medicinal chemistry, protein chemistry and biological screening functions, has successfully yielded a portfolio of potent and specific inhibitors of JAK kinases. These compounds are being further developed for application into the different therapeutic areas mentioned above.

II. Methods

A. Structure-based Drug Design and Virtual Screening

Commercial and academic software for Structure-Based Drug Design (SBDD) is now routinely used in the biotechno-

H. R. Treutlein is with Cytopia Research Pty. Ltd., 576 Swan Street, Richmond, Vic. 3121, Australia. (phone: 03-9208-4247; fax: 03-92084299; e-mail: herbert.treutlein@ cytopia.com.au).

M. Styles was with Cytopia Research until 2006. She is now with Pfizer GRD, Sandwich, Kent , UK (e-mail: michelle.styles@pfizer.com).

J. Zeng is with Cytopia Research Pty. Ltd., 576 Swan Street, Richmond, Vic. 3121, Australia (e-mail: jun.zeng@cytopia.com.au).

M. Chetty, S. Ahmad, A. Ngom, S. W. Teng (Eds.): PRIB2008, Supplementary Conf. Proc., pp.242-245, 2008.

logical and pharmaceutical industries. Most software developed to date tries to efficiently dock one compound at a time and find an optimal binding mode for each compound, independently of other information available. Thus the solutions provided address the more academic question of ab-initio binding mode and affinity prediction without existing knowledge of the particular chemistry and target involved, apart from the 3D coordinates. These methods are therefore not focused on the practical use of SBDD software industrial research. Many solutions also disregard a broad range of possible conformations of a protein target and available experimental data. The plethora of data from wet screening is often difficult to incorporate into the analysis of virtual screening results.

We have developed a set of software tools, called **ChemaPhore™**, to aid the design of potent and selective inhibitors based on virtual screening of 3D protein structures from X-ray, NMR and homology modelling; binding mode analysis; medicinal chemistry and wet screening. ChemaPhore™ employs a fragment-based approach to docking of compounds based on a method developed for designing peptide inhibitors of Ras-Raf interactions [16, 17]. ChemaPhore™ has been used successfully at Cytopia to design a number of highly potent and selective tyrosine kinase inhibitors. Benchmark tests based on publicly available data have confirmed the high standard of ChemaPhore™ (see Ref. [4] and Appendix).

B. *Homology Modelling*

All homology models presented in this paper were created with the **Modeler** software package [11] distributed by Accelrys (San Diego, CA, USA). The procedures employed followed the standard methods described in ref. [11].

The overall folding topology of protein kinase catalytic domains is highly conserved [15]. Homology modelling of unknown tyrosine kinase catalytic domains based on X-ray structures of related kinase catalytic domains is therefore highly feasible. Template structures chosen for the modelling of the Jak2 (and Jak3) kinases were the FGFR and LCK kinase domains (PDB [8] IDs: 1AGW [13], 1QPJ [14], resp.) The overall amino acid sequence identities and similarities for FGFR are 30% and 48%, resp. For LCK, the identities and similarities are 32% and 48%, resp. Since all tyrosine kinases bind ATP in a very similar way, the amino acids located in the ATP binding pockets are conserved at a significantly higher percentage, resulting in a high accuracy of the modelled ATP binding pocket.

III. RESULTS AND DISCUSSION

A. *Binding Mode of Compound 1*

In early 2002, a paper published by Thompson et al. [5] from Merck Research Laboratories described a JAK family-specific kinase inhibitor, a tetracyclic pyridone 2-tert-butyl-9-fluoro-3,6-dihydro-7H-benz[h]-imidaz[4,5-f]isoquinoline-7-one (referred to in this poster as "Compound 1", see Fig.1)

In 2002 no X-ray structure of any JAK kinase was available, however, for Cytopia's in-house screening and drug design purposes we had created homology models of the JAK2 and JAK3 kinases based on available kinase X-ray structures.

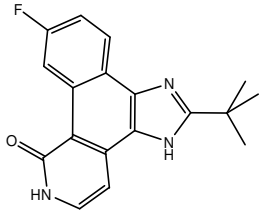

Figure 1: Chemical structure of "Compound 1": 2-tert-butyl-9-fluoro-3,6 -dihydro-7H-benz[h]-imidaz[4,5-f]isoquinoline-7-one

We used our homology models and our ChemaPhore™ docking software to predict the binding mode of Compound 1 inside the JAK2 ATP binding pocket. The predicted orientation (binding mode) of Compound 1 in the ATP binding site is shown in Fig.2. The model showed compound 1 sitting inside the ATP binding pocket of the Jak2 kinase domain. A tight fit is facilitated by the planarity of Compound 1. The compound showed good hydrogen bonding potential to the backbone of the "hinge region" of the ATP binding site (E930 and L932). The fluorine atom of compound 1 was located towards the opening of the pocket, while the pyridone ring was positioned towards the back pocket. The tert-butyl group of compound 1 was situated in a similar area as the pentose sugar of ATP.

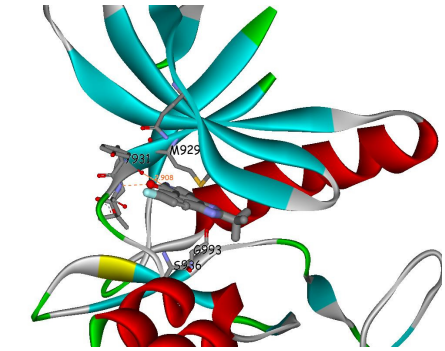

Figure 2: Structure of the complex formed by Jak2 kinase domain and compound 1. The protein structure is shown as ribbon model, colored according to its secondary structure: α-helix (red), ß-strand (cyan), turn (green), random (white). Selected protein residues involved in ligand binding are shown as stick models and labeled.

B. *SAR and Biochemical Proof of Binding Mode*

In order to validate the predicted binding mode, and lacking a detailed experimental structure of the complex formed by compound 1 and Jak2, we designed a number of chemical analogues of Compound 1 and measured their activity to Jak2. The analogues, shown in Fig.3, were designed to test properties of the binding mode that were deemed to be important contributors to the total affinity.

The structure affinity relationship (SAR) deduced from the experimental data showed that removing the fluorine atom from compound 1 or hydrogenation of a double bond in the

vicinity of M929 (the "gatekeeper" residue of the ATP binding pocket) showed minimal effect on binding affinity. Disrupting the planarity of the compound, however, disrupted activity significantly. These results are in agreement with the binding pose shown in Fig.2.

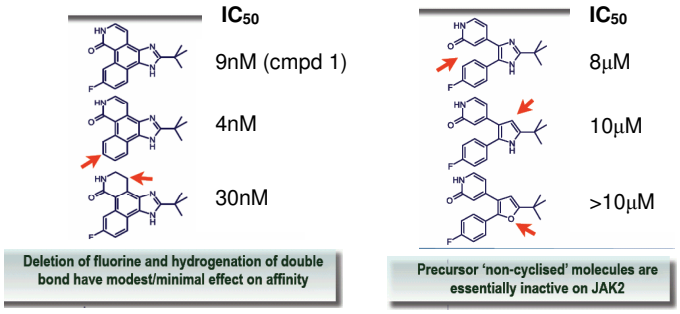

Figure 3: Jak2 activities of Compound 1 and analogues synthesized by Cytopia. Sown are the chemical structures and IC_{50} values measured. Modifications to Compound 1 are indicated by arrows.

C. Comparison of Predicted and Observed Binding Mode

More than two years after the calculations described above, i.e. in 2004, an X-ray structure of the complex formed between the JAK2 kinase domain and Compound 1 was solved by our collaborators at Monash University and subsequently published in 2006 [6] (PDB [8] ID: 2J90). The observed Compound 1 binding mode in the X-ray structure confirmed our earlier binding mode prediction. The orientation of the compound and the side chain contacts are identical and the RMS deviation between the overlaid compound structures is about 0.9 Å. The overlay is based on a best fit of Cα atoms of the protein in the complexes. The superposition of the ATP binding sites with bound Compound 1 of the X-ray (yellow) and the model (coloured by atom types) structures is shown in Fig.4.

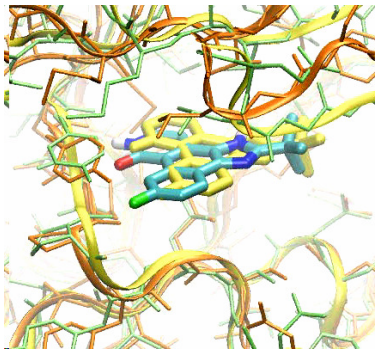

Figure 4: Overlay of the model and the X-ray structure [6] of the complex of Compound 1 and Jak2 kinase domain. Protein atoms are show in green (X-ray) and orange (model), the protein backbone is colored yellow (X-ray) and orange (model). Compound 1 is drawn as a yellow or atom type colored stick model for the X-ray or model structures, resp

IV. Conclusion

We have shown the suitability of homology models of kinase domains for structure-based drug design. We have successfully predicted the binding pose of a known kinase inhibitor in the ATP binding pocket of Jak2. The knowledge of the binding mode has been used to design inhibitors with improved enzyme selectivity (data not shown).

Appendix

CHEMAPHORE™ BENCHMARK

In order to validate our ChemaPhore™ software we have performed a number of in-house and public domain benchmark tests to compare the efficiency of computational (virtual) screening. Below are results of two of our benchmark tests.

During experimental high throughput screening (HTS) certain compounds (called "hits") will show an activity (e.g. IC_{50} < 10-50 μM) towards the protein target of interest. The hit rate is defined as the ratio of the number of hits and the total number of compounds in a library (x 100).

The aim of virtual screening is to find a small(er) subset of the original library which contains all or at least most of the hits. The hit rate of a computational method should then be larger than the experimental hit rate. A practical measure of how well a computational method performs is the "enrichment rate", defined as the ratio of the "computational hit rate" and the "experimental hit rate".

A. Finding Hits in a Random Library

For this test, we used the complete 2001 version of the Maybridge compound library wit 55,273 compounds, available from Maybridge, Trevillet, Tintagel, Cornwall PL34 0HW, UK. The aim was to find hits with IC_{50} < 50μM against the HCK kinase catalytic domain. We used the X-ray structure of HCK, PDB [8] ID 2hck [9] for virtual screening.

Selection	Library Size	Hit Rate [%]	Enrichment
HTS	55273	0.10 ‡	1.0
ChemaPhore™	147	8.84	88.4
In Detail:			
IC_{50} < 50μM	13	8.84	88.4
IC_{50} < 5 μM	6	4.08	40.8

Table 1: Library Enrichment by ChemaPhore™ at Different Activity Thresholds for the Random Maybridge library and HCK kinase.
‡ Assumes an expected HTS hit rate for a random library of 0.1%

Table 1 shows the results of the benchmark test. Due to the large number of compounds in the Maybridge library, it was not possible for us to screen all of the compounds experimentally. Therefore an experimental hit rate of 0.1% was assumed (see e.g. [7]). All 55273 compounds were, however, computationally screened. The 147 compounds with the best

score were selected, purchased and wet screened. As shown in Tab.1 we enriched the library computationally over 88-fold. The cost of the total Maybridge (2001) library would have been AUD 1.1 million to find potentially about 55 hits, however virtual screening using of ChemaPhore™ found 13 hits at a cost of AUD 3,000 for the 147 purchased compounds.

B. Diverse Library of Compounds Targeting Angiogenin

For this test that was published by Jenkins [7], the NCI Diversity set plus the ChemBridge DIVERSet was used (see ref. [7]) to result in a total library of about 18,000 highly diverse compounds. The protein target chosen was angiogenin (PDB [8] id: 1B1I [10]). In this case, experimental data for all 18111 compounds was made available in ref. [7]. The overall hit rate of the random library was found to be 0.07%.

Comparison with the published benchmark results are shown in Tab.2. Based on specific cutoffs for the internally used scoring systems of the virtual screening software, a number of library sub-sets were selected and the enrichment rate was calculated. The enrichment rate from ChemaPhore™ calculations was found to be in a similar range or higher than the results obtained from industry standard software [7].

Selection	Library Size	Hit Rate [%]	Enrichment
HTS	18111	0.07	1.0
ChemaPhore™			
1.0 x RMSD	671	1.04	15.7
2.0 x RMSD	239	1.67	25.3
From ref. [7] Table II			
DV/L 10%	1811	0.22	3.3
GOLD 2%	362	1.38	20.8
DV/L-GOLD 10%M	505	0.59	9.0

Table 2: Enrichment rates of HTS and VS at Different Scoring Thresholds. ChemaPhore™ thresholds are based on the average score and the root mean square deviation (RMSD) of the average score obtained for a library during virtual screening.

ACKNOWLEDGMENT

X-ray crystal structures were solved by Protein Crystallography Unit, Dept. of Biochemistry, Monash University in collaboration with Cytopia Research Pty Ltd. Compound synthesis by J. Feutrill, Cytopia Research Pty. Ltd.

REFERENCES

[1] Wilks AF. "Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction*," Proc Natl Acad Sci USA*, vol 86, no.5, pp.1603-7, March 1989.

[2] Wilks, AF, et al. "Two novel protein-tyrosine kinases, each with a second phosphotransferase-related catalytic domain, define a new class of protein kinase," *Mol Cell Biol.,* vol 11, no. 4., pp. 2057-65, April 1991.

[3] Harpur AG, et al. "JAK2, a third member of the JAK family of protein tyrosine kinases*," Oncogene*, vol. 7, no. 7, pp. 1347-53, July 1992.

[4] ChemaPhore™: Tool for Pragmatic Computer-Aided Structure-Based Drug Design; presentation at Drug Discovery Technology Meeting in Boston, August 2005.

[5] Thompson JE. "Photochemical preparation of a pyridone containing tetracycle: a Jak protein kinase inhibitor.," *Bioorg. Med. Chem. Lett,.* vol 12, pp. 1219-23, 2002.

[6] Lucet IS et al. "The structural basis of Janus kinase 2 inhibition by a potent and specific pan-Janus kinase inhibitor.," *Blood*, vol. 107, pp. 176-83, 2006.

[7] J. L. Jenkins et al., "Virtual screening to enrich hit lists from high-throughput screening: a case study on small-molecule inhibitors of angiogenin.," *Proteins: Struct. Funct. Gen.*, vol. 50,pp. 81-93, 2003.

[8] H.M.Berman, et al. "The Protein Data Bank," *Nucleic Acids Res*., vol 28, pp. 235-242, 2000.

[9] Sicheri, F., Moarefi, I., Kuriyan, J., "Crystal structure of the Src family tyrosine kinase Hck," *Nature* vol 385, pp. 602-609, 1997.

[10] Leonidas, D.D., et al., "Refined crystal structures of native human angiogenin and two active site variants: implications for the unique functional properties of an enzyme involved in neovascularisation during tumour growth," *J.Mol.Biol*., vol. 285, pp. 1209-1233, 1999.

[11] A. Šali and T. L. Blundell, "Comparative protein modelling by satisfaction of spatial restraints," *J. Mol. Biol.,* vol. 234, pp. 779-815, 1993.

[12] Campbell McInnes*, and Peter M. Fischer, "Strategies for the Design of Potent and Selective Kinase Inhibitors," *Current Pharmaceutical Design*, vol. 11, pp. 1845-1863, 2005.

[13] Mohammadi, M., et al., "Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors,"*Science* , vol. 276**,** pp. 955-960 , 1997.

[14] Zhu, X., et al. "Structural analysis of the lymphocyte-specific kinase Lck in complex with non-selective and Src family selective kinase inhibitors," *Structure Fold.Des.,* vol. 7**,** pp. 651-661, 1999.

[15] Manning, G., et al. "The protein kinase complement of the human genome," *Science*, 2002, vol. 298, p. 1912, 2002.

[16] Zeng J., Treutlein H.R.*,* "A method for computational combinatorial peptide deign of inhibitors of Ras." *Protein Eng,.* vol. 12, pp. 457-468, 1999.

[17] Zeng J., Nhen T.V., Zorzet A., Catimel B., Nice E., Maruta H., Tony Burgess A.W., and Treutlein H.R. "Computational Combinatorial Design of Short Peptides that Inhibit Ras-Raf Interaction *in vitro*." *Protein Eng*., vol. 14, pp. 39-45, 2001.