Journal of Computer-Aided Molecular Design, **16:** 831–840, 2002.
KLUWER/ESCOM

Protein Alpha Shape Similarity Analysis (PASSA): A new method for mapping protein binding sites. Application in the design of a selective inhibitor of Tyrosine kinase 2

Kristin Tøndel[a,*], Endre Anderssen[a] & Finn Drabløs[b]
[a]*Department of Chemistry, Norwegian University of Science and Technology, Sem Selands v. 14, N-7491 Trondheim, Norway;* [b]*SINTEF Unimed MR-center, N-7465 Trondheim, Norway*

Received 1 October 2002; Accepted 6 January 2003

Key words: 3D QSAR, alpha shapes, drug design, homology modelling, PLS, proteins, selectivity, tyrosine kinase inhibitors.

Summary

We have developed a method that we have called Protein Alpha Shape Similarity Analysis (PASSA), that identifies interaction sites that can be utilised to achieve selectivity towards a protein. We have shown that this method is able to identify residues of tyrosine kinases that interact with known selective inhibitors using the following test cases: Abelson (Abl) kinase in complex with STI-571 and Janus kinase 2 (Jak2) in complex with AG-490. The 3D structures of the tyrosine kinase domains of Tyrosine kinase 2 (Tyk2) and Jak2 have been predicted by homology modelling. Computational docking of AG-490 and a set of tyrphostins known not to inhibit Jak2 indicated that our homology models are able to separate inhibitors from non-inhibitors. PASSA has also been used to identify unique properties of Tyk2. According to our results, interactions with hydrogen acceptors and donors on the following residues can be utilised to achieve selectivity towards Tyk2: Y955, E1053, D1062 and S1063. These residues are placed close to non-conserved hydrophobic pockets. The PASSA results, together with results from Multiple Copy Simultaneous Search (MCSS) were used to suggest functional groups of a selective Tyk2 inhibitor.

Introduction

Protein kinases contribute to regulation and coordination of e.g. metabolism, gene expression, cell growth, cell motility, cell differentiation and cell division [1]. The Janus kinase (Jak) family of non-receptor tyrosine kinases consists of four known mammalian proteins (Jak1-3 and Tyk2) that play a critical role in initiating signalling cascades of a large number of cytokine receptors [2–5]. Tyrosine kinases are usually regulated by phosphorylation of tyrosine residues in the activation loop, located between the conserved DFG and APE motifs [6]. This tyrosine phosphorylation causes conformational changes in the activation loop, that allow ATP and protein substrates to access the active site [7].

*To whom correspondence should be addressed.
E-mail: kristito@phys.chem.ntnu.no

The Jaks catalyse phosphorylation of the Signal Transducers and Activators of Transcription (STAT) family of transcription factors [6]. After phosphorylation on tyrosine residues, the STAT molecules form homo- or heterodimers [8], which are translocated into the nucleus. The STAT proteins then bind to DNA, and activate gene transcription [2]. The Jak-STAT signalling cascade has been shown to contribute to growth and survival of e.g. human multiple myeloma cells [9], acute lymphoblastic leukaemia [10] and a variety of other malignancies [11,12]. This makes the Jaks potential targets for new cancer therapies. One way to block the function of the Jaks is to inhibit ATP binding. ATP competitive inhibitors are generally non-selective, but the development of inhibitors like STI-571 [13] shows that ATP binding sites can be used as targets for selective drugs. Since none of the Jaks have experimentally determined 3D structures at the

present time [14,15], we have made homology models [16] of the tyrosine kinase domains of Tyk2 and Jak2. A selective inhibitor of Jak2 has been reported [10]. We have therefore focused our design work on Tyk2, and used the model of Jak2 for method testing.

The quality of homology models is highly dependent on the choice of template structures. According to Chothia and Lesk [17], templates with sequence identity $> 50\%$ to the target proteins are likely to give reasonable models. If the sequence identity drops to 20%, there will be large structural differences. However, the active sites of distantly related proteins can have very similar geometries [18,19]. A weakness of using structures predicted by homology modelling as basis for the design of selective drugs is that to achieve selectivity one has to utilise variable regions of the proteins. These are the regions predicted with the lowest reliability by homology modelling techniques [20]. The ability of our homology models to separate inhibitors from non-inhibitors was tested by computational docking of the Jak2 inhibitor AG-490 and 10 other tyrphostins known not to inhibit Jak2 [10].

To design a selective inhibitor for Tyk2, we need to identify interaction sites that can form the basis for selectivity. We have developed a method that we have called Protein Alpha Shape Similarity Analysis (PASSA), which identifies residues that are unique to one protein compared to several others. A number of methods exist to map protein binding sites. Some force field based methods, such as GRID [21] and Multiple Copy Simultaneous Search (MCSS) [22], use calculated interaction energies between probe molecules and the protein. In GRID the interactions are estimated by placing a probe atom at a number of fixed grid points in the protein. MCSS does not use a fixed grid. Instead, a geometry optimisation is performed on a large number of probe molecules placed randomly in the binding site. The probe molecules that bind strongly to the protein can then be taken as a basis for placement of functional groups in e.g. combinatorial library design. Other methods use the shape of the protein to find potential binding sites, without energy calculations. An example is alpha sphere-based methods [23]. Alpha spheres are geometrical representations of protein cavities. Alpha sphere centres are often found close to atoms of docked ligands [24].

In order to determine which sites can contribute to selectivity, a number of proteins must be mapped and the binding sites compared. GRID has been used to reveal structural differences between proteins [21]. The proteins are aligned prior to GRID calculations, and the data is analysed by Principal Component Analysis (PCA) [25]. MCSS and methods using alpha shapes are less suitable for direct comparison of proteins, because of the free movement of probe molecules and the absence of a fixed frame of reference, such as a grid. Force field methods can give spurious results due to errors in alignment or structures. The Lennard-Jones- and electrostatic terms of force fields are very steep close to atomic nuclei. Small changes in atomic positions can therefore lead to large changes in the calculated energy.

To avoid some of the problems mentioned above, we have combined Gaussian property distributions (similar to those used in Comparative Molecular Similarity Index Analysis (CoMSIA) [26]) and alpha shapes [23] in order to compare proteins. The value of a property field at each point of a grid on the aligned proteins is computed as a weighted sum of property Gaussians. The resulting similarity fields can be analysed by e.g. PCA or Discriminant Partial Least Squares (DPLS) regression [25]. In DPLS, the response matrix contains binary indicator variables indicating memberships in different classes. When the purpose is to differentiate between known classes of proteins, DPLS has the advantage that the most relevant differences are extracted and summarised as a single vector of regression coefficients. The regression coefficients can be displayed as isosurfaces on the 3D structures of the proteins. Such visualisations can be combined with the results from MCSS. This combines information about binding of functional groups and potential for selectivity.

STI-571 is a selective inhibitor of Abelson (Abl) kinase, platelet-derived growth factor (PDGF) receptor and c-kit [13,27]. We have tested whether the residues identified by PASSA to be unique to Abl kinase match the residues that interact with STI-571. The same test was carried out using the homology model of Jak2 in complex with the lowest docking energy conformation of AG-490. AG-490 has been reported to also inhibit Jak1 [28] and Jak3 [11,29], but not Tyk2 [30]. This test is valid, since neither Jak1 nor Jak3 were included in this analysis. We have utilised PASSA to identify unique properties of the Tyk2 tyrosine kinase domain. The results from this analysis, together with results from MCSS runs, have been used to suggest positioning of functional groups of a selective Tyk2 inhibitor.

Methods

Homology modelling

Homology models for the kinase domains of Tyk2 (F892-Q1177) and Jak2 (S833-N1129) were made using five templates simultaneously in SwissModel [31–34]. Suitable templates were found using the SwissModel Blast search [31,32]. In cases where more than one file were available in the RCSB Protein Data Bank (PDB) [14,15] for the same protein, the structure with the best resolution was used. All templates have sequence identities of 35%–45% to the target. SwissModel estimates the model reliability (Model B-factor) [32] for each atom in the model, based on the similarity between the target protein and the templates.

Hydrogens were added to the model structures in Molecular Operating Environment (MOE^{TM}) [35], and the structures were energy minimised to an RMS gradient of 0.01 using the AMBER94 force field [36] with a smooth non-bonded cut-off of 10–12 Å. The calculations were carried out in vacuum, using a distance-dependent dielectric to approximate the solvent screening effects. The energy minimisations were performed with fixed positions for backbone atoms of high reliability regions and heavy atoms of residues in inter-domain contact regions, because it is generally known that with extensive refinement, homology models tend to get worse [37]. In domain modelling, positions of atoms forming an interface to a missing domain should be fixed during energy minimisation. Free movement in these regions can lead to sidechain conformations that are preferable energetically, but not possible in the real protein structure because of interactions with the missing parts of the protein. The quality of the structures was verified by WHAT_CHECK [38,39]. The residues for which the WHAT_CHECK routines reported unusual conformations were relaxed, and the models were again optimised to an RMS gradient of 0.01.

For comparison, three homology models were made for both Tyk2 and Jak2 using only one template for each model in SwissModel. The hydrogens of the structures were optimised as described above. All non-hydrogens were held in fixed positions during this optimisation. A structure superpositioning of the α-carbons of these models was carried out in Swiss-PdbViewer [31,40], and the Cα Root Mean Square Distance (RMSD) was calculated.

Multiple Copy Simultaneous Search (MCSS)

We have carried out MCSS [22] runs in MOE to identify binding sites for acetamide, acetaldehyde, water, methane and benzene in the ATP binding pockets of Tyk2 and Jak2. MMFF94 [41] with implicit solvent electrostatic corrections [42–44] were used for the energy minimisations. 500 copies of each probe molecule were used.

Docking analysis

In addition to WHAT_CHECK verification, the model quality was tested by computational docking. All docking calculations were carried out using Tabu search [45] in MOE [35]. Tabu search is a stochastic searching algorithm that maintains a list of previously visited conformations, to guide the searching towards better conformations. The MMFF94 force field [41] was used, and the calculations were done in vacuum, with a distance-dependent dielectric and a smooth non-bonded cut-off of 10–12 Å. MMFF94 was chosen because it has more parameters for small molecules [46] than AMBER94. The AMBER force fields are more suited for calculations on proteins. Grid-based potential fields [35] were used for estimation of the interaction energies. Hence, the potential energy grids were calculated only at the beginning of the docking procedure.

The model of Tyk2 was aligned in MOE to the structure of human cyclin-dependent kinase 2 (CDK2) present in PDB [14,15] entry 1HCK. Sequence alignments were carried out using a modified version of the Needleman and Wunsch approach [47] with a structural correction and the Gonnet similarity matrix [48]. The 3D structures were superposed as described in [49]. ATP was docked 10 runs of 25 000 iterations using the experimentally determined ATP structure present in 1HCK as the starting position.

All fragments from the MCSS run on Jak2 having negative interaction energies with the receptor were combined into a ‘molecular cluster’. The structures of the tyrphostins were aligned onto this cluster using flexible alignment [50] in MOE. This generated a set of starting conformations for each tyrphostin. All starting conformations were docked in the homology model of Jak2, using the same docking box ($150\times150\times150$ grid points with 0.15 Å spacing). 1000 iterations were used for each starting conformation.

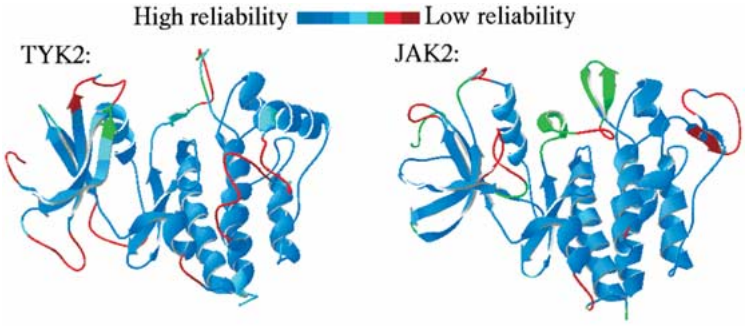

Figure 1. Predicted structures of the tyrosine kinase domains of Tyk2 and Jak2. The residues are coloured according to the Model B-factor [32] computed by SwissModel.

PASS Analysis

Property Gaussians were used to analyse the protein structures. Alpha-spheres [23] were used to identify potential hydrophobic and hydrophilic binding sites. 'Dummy' atoms placed at each alpha-sphere centre were assigned similarity field weights (ω) of +1 for either the hydrophobic or the hydrophilic field. In order to include steric effects, protein atoms were assigned a field weight of −1 for both fields (Equation (1)).

$$F(q, j) = \sum_{i=1}^{n} \frac{\omega_{ik}}{(\sigma_i \sqrt{2\pi})^3} \cdot e^{\frac{-r_{iq}^2}{2\sigma_i^2}} \qquad (1)$$

F is the value of the similarity field in grid point q of molecule j, ω_{ik} is the value of the physicochemical property k of atom i, r_{iq} is the distance between grid point q and atom i and σ_i corresponds to the atomic radii of atom i. PASSA works as follows:

1. Structural alignment of the proteins.
2. Placement of a grid surrounding the active sites of the proteins.
3. Determination of alpha-sphere positions.
4. Calculation of the value of the molecular similarity field in each grid point.
5. The molecular similarity field is taken as input to DPLS regression [25].

The regression coefficients from the DPLS regression can be mapped back onto the grid. This gives us the opportunity to visualise the regions having properties that are unique to a specific protein. All scripts were written in Scientific Vector Language (SVL) [35], and are available from the authors upon request.

The five- and one-template models of Tyk2 and Jak2, the structures used as templates in the homology modelling and PDB entries 1JST, 3LCK, 1VR2, 1IRK, 2SRC, 1AD5 and 1IEP were included in the PASS Analysis. Indicator variables for the following classes were used: Tyk2-models, Jak2-models, structures of non-Janus kinases and Abl kinase structures. The structures were superposed in MOE [35] using the same approach as mentioned in the previous section. A 3D grid was centred at the nucleotide-binding loop of Tyk2 (L903-V911). 50×50×40 grid points were used with a grid spacing of 0.75 Å. The DPLS regression was done in MATLAB™ [51], using the PLS Toolbox® [52]. All columns with standard deviation $< 10^{-4}$, or three or less nonzero entries were removed from the PASSA data. The data for each physicochemical property was set to equal variation by dividing each data point by the sum of the singular values of the data for this property. The resulting matrix was used as regressor data in the DPLS regression.

Table 1. Docking energies for the 11 tyrphostins used for verification of the model quality[a]

Tyrphostin	Docking energy (kJ/mol)
AG-490	−32.75
AG-30	−23.35
AG-18	−17.47
AG-126	11.19
AG-1295	19.73
AG-294	34.20
AG-370	40.29
AG-1112	47.23
AG-1007	56.52
AG-1478	69.41
AG-879	101.44

[a]AG-490 is an inhibitor of Jak2 [10], while all the other tyrphostins are non-active.
AG-1007 resembles AG-490 in structure (Figure 4).

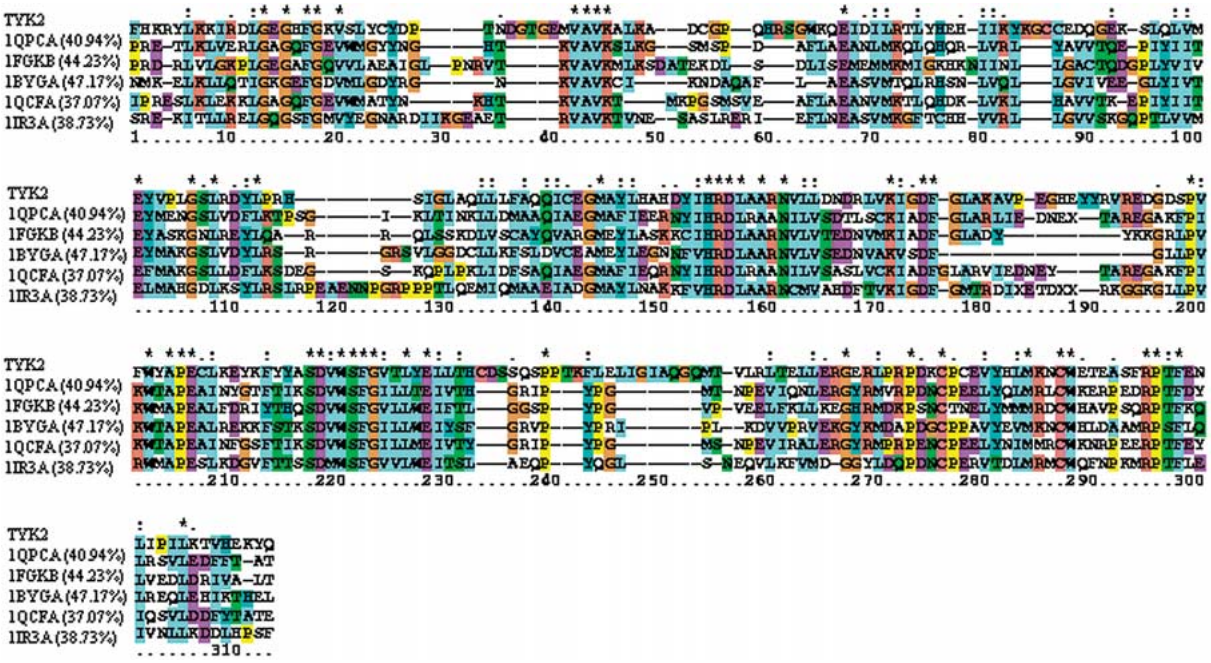

Figure 2. SwissModel multiple sequence alignment of Tyk2 (F892-Q1177) with the templates used for homology modelling. The sequence identities between the templates and the Tyk2 tyrosine kinase domain are shown together with the PDB [14,15] identifiers. The nucleotide-binding loop and the DFG motif are given in alignment positions 13–46 and 176–179, respectively.

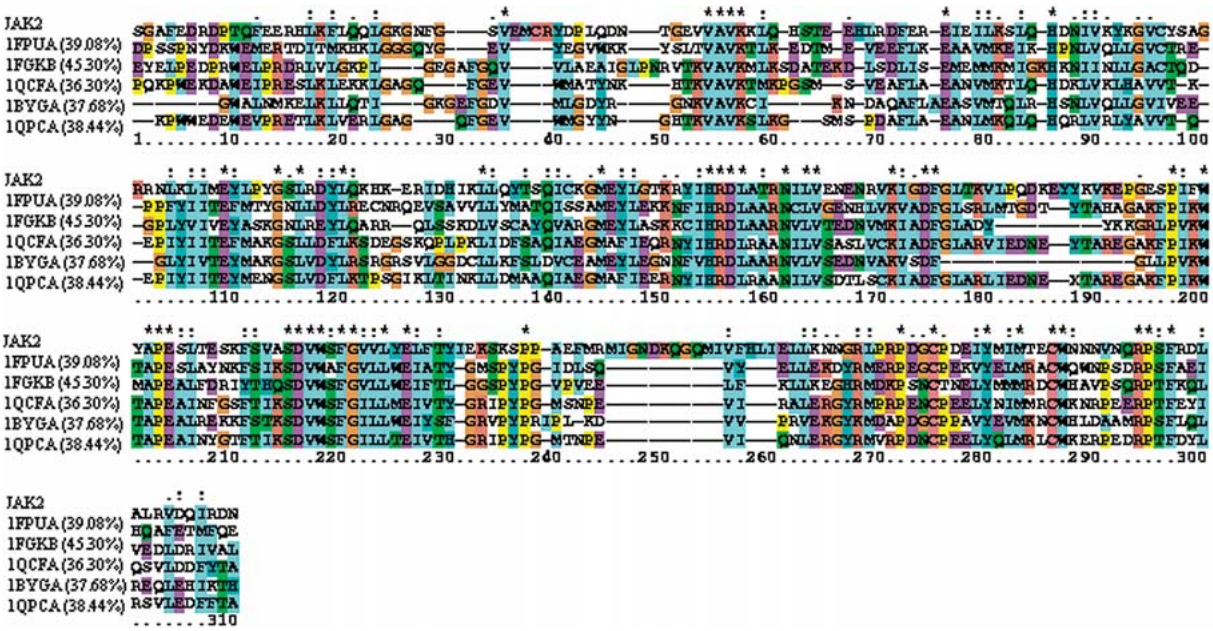

Figure 3. SwissModel multiple sequence alignment of Jak2 (S833-N1129) with the templates used for homology modelling. The sequence identities between the templates and the Jak2 tyrosine kinase domain are shown together with the PDB [14,15] identifiers. The nucleotide-binding loop and the DFG motif are given in alignment positions 24–58 and 176–178, respectively.

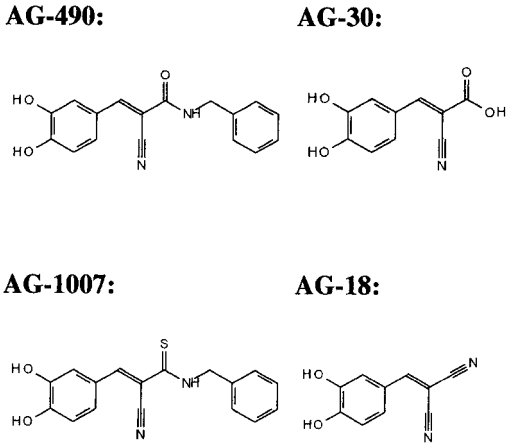

Figure 4. The structures of AG-490 and the other two tyrphostins with negative docking energies, together with the structure of AG-1007.

Results and discussion

Homology modelling

Figure 1 shows that the estimated reliability of our models of Tyk2 and Jak2 is high in large areas of the structures. Most of the nucleotide-binding loop and the activation loop of Tyk2 are predicted with relatively high reliability, but we have some problem areas with large gaps in the sequence alignment (Figure 2). The reliability of the structure prediction of the nucleotide-binding loop and the activation loop of Tyk2 seems to be somewhat higher than for Jak2. The sequence identity between the modelled domains of Tyk2 and Jak2 is 47.9%.

The influence of the choice of template was verified by comparing three different one-template models for both Tyk2 and Jak2. 1QPCA, 1QCFA and 1IR3A were used as templates for Tyk2, while 1BYGA, 1QPCA and 1IR3A were used as templates for Jak2. The other individual templates did not contain sufficient information to construct reasonable homology models when used alone. Structure superpositioning gave a Cα RMSD of 1.31 Å between the three models of Tyk2, and a Cα RMSD of 1.18 Å between the models of Jak2. The Cα RMSD between the five-template models of Tyk2 and Jak2 was 0.75 Å, and increased to 0.77 Å after geometry optimisation. The relatively high RMSD values between the homology models made using one template illustrates the importance of the choice of templates.

Figures 2 and 3 show the multiple alignments of Tyk2 and Jak2, respectively, with their templates. SwissModel makes a structural correction of the alignments. Conformational differences between the templates therefore lead to gaps in the DFG motifs and in the nucleotide-binding loop of Jak2. Figure 2 shows that 1QPCA and 1IR3A, which are both crystal structures of activated kinase domains, give the best alignments with Tyk2 in the activation loop. Hence, Tyk2 is modelled in its active conformation. We see from Figure 3 that 1FPUA and 1QCFA give the best sequence alignments with Jak2 in these regions. Since both 1FPUA and 1QCFA are structures of inactive kinases, Jak2 is modelled in its inactive conformation.

As described by Schindler et al. [13], the selective inhibitor STI-571 binds to the inactive conformation of Abl kinase. We assume that the greater diversity in inactive kinases makes them better targets for selective drug therapy. Attempts to make a model of the inactive conformation of Tyk2 have not yet given reliable results. The model of the active conformation of the Tyk2 tyrosine kinase domain was therefore used.

Docking analysis

As the purpose of the structure modelling of Tyk2 is inhibitor design, the ability of the models to distinguish inhibitors from non-inhibitors is essential. Table 1 shows the results from the computational docking of the 11 tyrphostins in the Jak2-model. Table 1 shows that AG-490 is identified as the most potent inhibitor of Jak2. All the other tyrphostins are predicted to be less potent than AG-490. The separation of AG-1007 from AG-490 is especially interesting, since their structures are very similar. In AG-1007 the carbonyl oxygen of AG-490 is replaced with a sulphur atom. The structures of AG-490 and the other two tyrphostins with negative docking energies are shown in Figure 4, together with the structure of AG-1007. The Model B-factors of the Tyk2 and Jak2 models from SwissModel (see Figure 1) and the WHAT_CHECK report indicate that the model of Tyk2 is at least as good as the model of Jak2. Hence, we assume that the Tyk2-model may be used in inhibitor design.

PASS Analysis

The score plot from the DPLS regression (Figure 5) indicates that the PASSA data is able to predict the correct memberships for the Tyk2- and Jak2 models, and the other kinases. Since 1IR3 and 1QPC are crystal structures of activated kinases, JAK2-1IR3 and JAK2-1QPC lie closer to the Tyk2-models than the other models of Jak2. The models of Tyk2 and Jak2

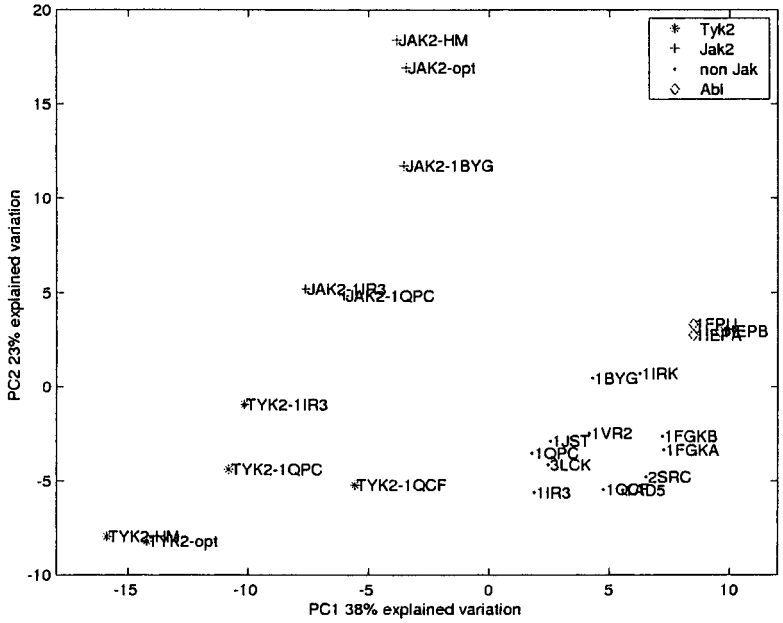

Figure 5. Score plot of the two first principal components (PCs) from the DPLS regression. JAK2-1BYG = Model of Jak2 made using only 1BYGA as template (likewise for the other one-template models). Postfix 'HM' means five-template homology model before geometry optimisation. Postfix 'opt' means the final optimised model made using five templates.

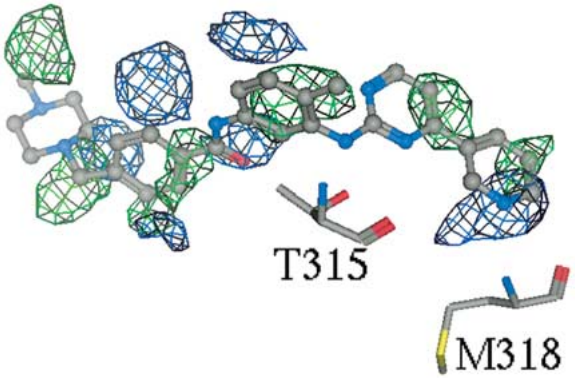

Figure 6. Plot of the regression coefficients for the hydrophilicity (blue) and the hydrophobicity (green) from the DPLS analysis mapped back onto the grid surrounding the ATP binding pocket of Abl kinase in complex with STI-571 (PDB entry 1IEP).

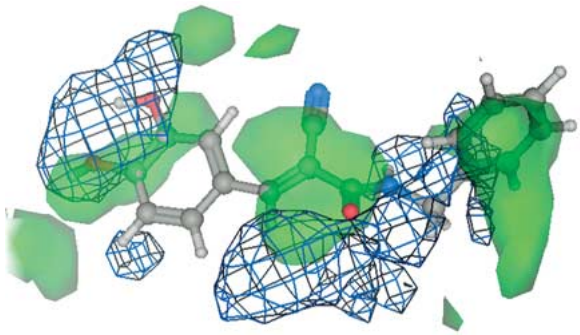

Figure 7. Plot of the regression coefficients for the hydrophilicity (blue) and the hydrophobicity (green) from the DPLS analysis mapped back onto the grid surrounding the model of the ATP binding pocket of Jak2 containing the lowest docking energy conformation of AG-490.

made using five different templates are the most reliable models of these two enzymes. These models are most widely separated from each other and from the other kinases in the score plot. This indicates that these models have properties that are more characteristic to Tyk2 and Jak2, respectively, than the models made using only one template.

The regression coefficients from the DPLS analysis indicate areas where each class of proteins has properties that separate them from the other proteins. Interactions between the protein and inhibitor groups placed in areas of high regression coefficients can contribute to selectivity. We have verified the performance of PASSA by testing whether residues known to interact with selective inhibitors are among the residues identified to be unique. STI-571 is a selective inhibitor of Abl kinase [13,27]. The regression coefficients for Abl were mapped back onto the grid on the X-ray structure of this protein present in PDB entry 1IEP. We see from Figure 6 that the areas of high regression coefficients for the hydrophobicity correspond well to

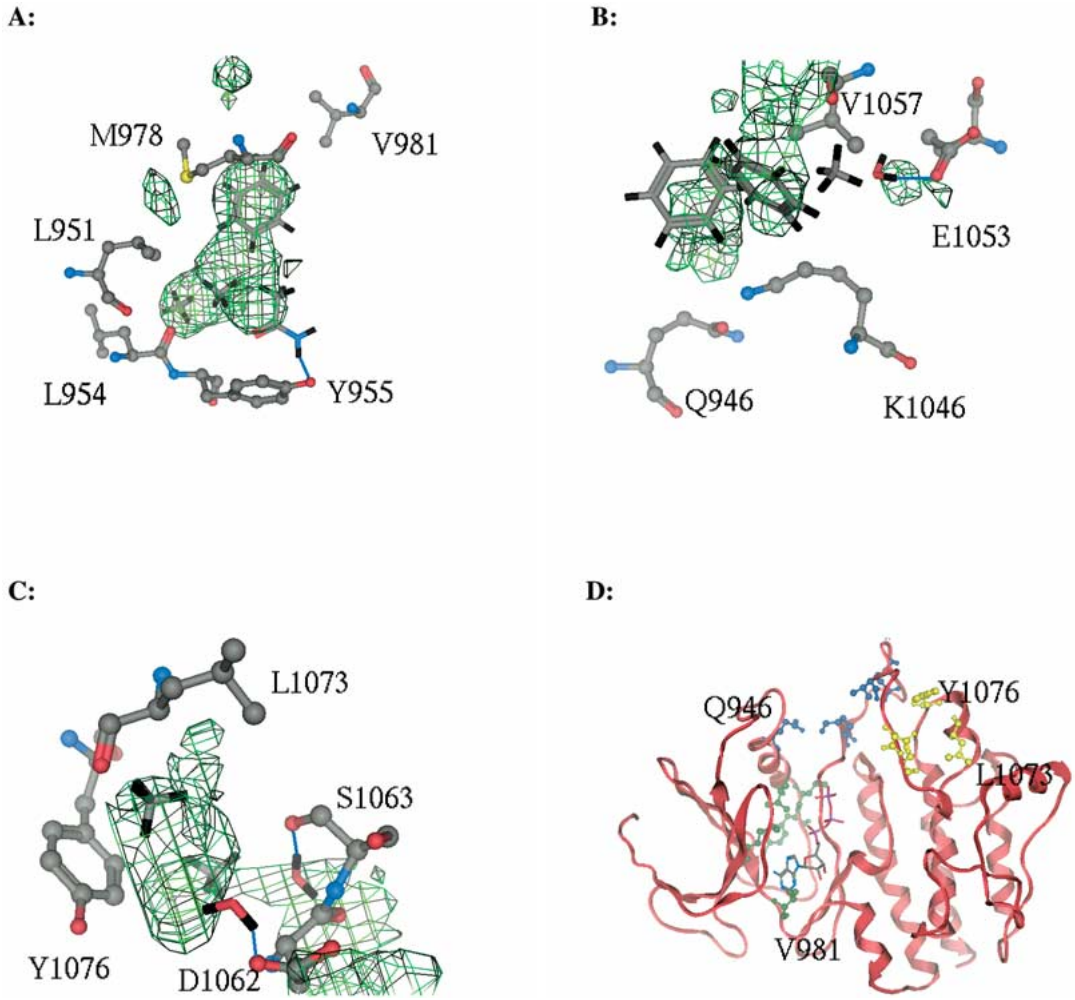

Figure 8. A–C: Plots of the regression coefficients for the hydrophobicity (green) from the DPLS analysis mapped back onto the grid surrounding the model of the ATP binding pocket of Tyk2. Residues identified by PASSA to be unique to Tyk2 are shown, together with selected fragments from the MCSS. Possible hydrogen bonds between MCSS fragments and hydrogen acceptors/donors of Tyk2 identified to be unique are shown as blue lines. D: The residues from Figure 8A (green), 8B (blue) and 8C (yellow) shown together with the result from the computational docking of ATP in the Tyk2 model.

the hydrophobic parts of STI-571. The phenyl-moiety of STI-571 known to interact with T315 in Abl [13] is placed in an area where Abl is particularly hydrophobic compared to the other proteins. The interaction between T315 and STI-571 is known to be important for selectivity [13,53]. According to our results, Abl kinase is particularly hydrophilic in the areas around the carbonyl group of STI-571, and around the nitrogen interacting with M318. A similar analysis was carried out using the homology model of Jak2 and the lowest docking energy conformation of AG-490. AG-490 inhibits Jak2, but none of the other proteins included in this analysis [10,11,28–30]. In Figure 7, the regression coefficients from the DPLS regression are mapped back onto the grid on the model of the ATP binding pocket of Jak2. The results from the PASSA indicate that Jak2 is particularly hydrophilic compared to the other proteins in the areas close to the OH-groups, the NH-group and the carbonyl group of AG-490. We see from Figure 7 that the areas where Jak2 is particularly hydrophobic correspond well to the hydrophobic parts of AG-490. The fact that the interactions between Abl kinase and STI-571, and between Jak2 and AG-490 are identified by PASSA, indicates that this approach may be utilised in the design of selective drugs.

Figure 8 shows the results from the PASS analysis of Tyk2. According to our results, Tyk2 has three

unique hydrophobic pockets that can be utilised by an inhibitor (shown in Figure 8A, B and C, respectively). Similar analysis for the hydrophilicity identified useful hydrogen acceptors and donors close to these pockets. According to our results, interactions with hydrogen acceptors/donors on the following residues can be utilised to achieve selectivity towards Tyk2: Y955, E1053, D1062 and S1063. Fragments from the MCSS placed in regions of high DPLS regression coefficients indicate possible functional groups for a selective Tyk2 inhibitor. These results can be used as a starting point for combinatorial library design, database searching and de novo ligand design. Figure 8A shows that the hydrophobic pocket created by V981, M978, L951, L954 and Y955 can be utilised by an inhibitor having hydrophobic groups pointing towards M978 and L954. In addition, a hydrophobic group with a hydrogen donor or acceptor interacting with the OH-group of Y955 may be advantageous. The hydrophobic pocket shown in Figure 8B can be occupied by a relatively large, aromatic structure, containing a hydrogen donor group in hydrogen-bonding position to E1053. According to our results, a selective inhibitor should also contain a hydrogen donor or acceptor that could interact with the OH-group of S1063, and a hydrogen donor close to D1062. These groups can be connected by e.g. a hydrocarbon chain occupying the space between Y1076 and L1073 (Figure 8C). Figure 8D shows the residues that according to our analysis may be utilised to achieve selectivity towards Tyk2. One can, of course, not guarantee that other protein structures not included in this analysis do not have the same properties as Tyk2 in some of these areas.

In conclusion, we have developed a useful method that identifies binding sites for functional groups that can lead to selectivity. The method has been tested using both X-ray structures and homology models, and appears to be robust against small structural errors caused by the homology modelling process and the computational docking.

Acknowledgements

We want to thank Anders Sundan and Magne Børset at The Department of Cancer Research and Molecular Biology at The Norwegian University of Science and Technology for helpful discussions. We also want to thank The Norwegian Research Council and Amersham Health for financial support.

References

1. Johnson, L.N., Noble, M.E.M. and Owen, D.J., Cell, 85 (1996) 149.
2. Ihle, J.N., Witthuhn, B.A., Quelle, F.W., Yamamoto, K. and Silvennoinen, O., Annu. Rev. Immunol., 13 (1995) 369.
3. Pellegrini, S. and Dusanter-Fourt, I., Eur. J. Biochem., 48 (1997) 615.
4. Van der Geer, P., Hunter, T. and Lindberg, R.A., Annu. Rev. Cell Biol., 10 (1994) 251.
5. Richter, M.F., Dumenil, G., Uze, G., Fellous, M. and Pellegrini, S., J. Biol. Chem., 273 (1998) 24723.
6. Hubbard, S.R. and Till, J.H., Annu. Rev. Biochem., 69 (2000) 373.
7. Pautsch, A., Zoephel, A., Ahorn, H., Spevak, W., Hauptmann, R. and Nar, H., Structure, *9* (2001) 955.
8. Heldin, C.H., Cell, 80 (1995) 213.
9. Anderson, K., Semin. Oncol., 26 (1999) 10.
10. Meydan, N., Grunberger, T., Dadi, H., Shahar, M., Arpaia, E., Lapidot, Z., Leeder, J.S., Freedman, M., Cohen, A., Gazit, A., Levitzki, A. and Roifman, C.M., Nature, 379 (1996) 645.
11. Wang, L.H., Kirken, R.A., Erwin, R.A., Yu, C.R. and Farrar, W.L., J. Immunol., 162 (1999) 3897.
12. Lindauer, K., Loerting, T., Liedl, K.R. and Kroemer, R.T., Protein Eng., 14 (2001) 27.
13. Schindler, T., Bornmann, W., Pellicena, P., Miller, W.T., Clarkson, B. and Kuriyan, J., Science, 289 (2000) 1938.
14. The RCSB Protein Data Bank, http://www.rcsb.org/pdb/.
15. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E., Nucleic Acids Res., 28 (2000) 235.
16. Bajorath, J., Stenkamp, R. and Aruffo, A., Protein Sci., 2 (1993) 1798.
17. Chothia, C. and Lesk, A.M., EMBO J., 5 (1986) 823.
18. Lesk, A.M. and Chothia, C., J. Mol. Biol., 136 (1980) 225.
19. Chothia, C. and Lesk, A.M., J. Mol. Biol., 160 (1982) 309.
20. Read, R.J., Brayer, G.D., Jurásek, L. and James, M.N.G., Biochemistry, 23 (1984) 6570.
21. Pastor, M. and Cruciani, G., J. Med. Chem., 38 (1995) 4637.
22. Miranker, A. and Karplus, M., Proteins Struct. Func. Genet., 11 (1991) 29.
23. Edelsbrunner, H., Facello, M., Fu, R. and Liang, J., Proceedings of the 28th Hawaii International Conference on Systems Science¸ Maui, January 1995, pp. 256-264.
24. Liang, J., Edelsbrunner, H. and Woodward, C., Protein Sci., 7 (1998) 1884.
25. Martens, H. and Martens, M. Multivariate Analysis of Quality, An Introduction, John Wiley & Sons, Ltd., Chichester, 2000.
26. Klebe, G., Abraham, U. and Mietzner, T., J. Med. Chem., 37 (1994) 4130.
27. Zimmermann, J., Buchdunger, E., Mett, H., Meyer, T. and Lydon, N.B., Bioorg. Med. Chem. Lett., 7 (1997) 187.
28. Xuan, Y.T., Guo, Y.R., Han, H., Zhu, Y.Q. and Bolli, R., Proc. Natl. Acad. Sci. USA, 98 (2001) 9050.
29. Kirken, R.A., Erwin, R.A., Taub, D., Murphy, W.J., Behbod, F., Wang, L.H., Pericle, F. and Farrar, W.L., J. Leukocyte Biol., 65 (1999) 891.
30. Bright, J.J., Du, C.G. and Sriram, S., J. Immunol., 162 (1999) 6255.
31. Guex, N. and Peitsch, M.C., Electrophoresis, 18 (1997) 2714.
32. Peitsch, M. C., Biochem. Soc.Trans., 24 (1996) 274.
33. Guex, N., Diemand, A. and Peitsch, M.C., TiBS, 24 (1999) 364.
34. Peitsch, M.C., Bio/Technology, 13 (1995) 658.

35. Molecular Operating Environment™, Version 2001.01, Chemical Computing Group, Inc., 2001.
36. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S. and Weiner, P., J. Am. Chem. Soc., 106 (1984) 765.
37. Charifson, P.S. Practical Application of Computer-Aided Drug Design, Marcel Dekker, Inc., New York, 1997.
38. Vriend, G., J. Mol. Graph., 8 (1990) 52.
39. Hooft, R.W.W., Vriend, G., Sander, C. and Abola, E.E., Nature, 381 (1996) 272.
40. Swiss-PdbViewer, Version 3.7b2, Glaxo Wellcome Experimental Research, 2001.
41. Halgren, T.A., J. Comp. Chem., 17 (1996) 490.
42. Still, W.C., Tempczyk, A., Hawley, R.C. and Hendrickson, T., J. Am. Chem. Soc., 112 (1990) 6127.
43. Qiu, D., Shenkin, S., Hollinger, F.P. and Still, W.C., J. Phys. Chem., 101 (1997) 3005.
44. Schaefer, M. and Karplus, M. A., J. Phys. Chem., 100 (1996) 1578.
45. Baxter, C.A., Murray, C.W., Clark, D.E., Westhead, D.R. and Eldridge, M.D., Proteins Struct. Funct. Genet., 33 (1998) 367.
46. Halgren T.A., J. Comp. Chem., 20 (1999) 730.
47. Needleman, S. B. and Wunsch, C. D., J. Mol. Biol., 48 (1970) 443.
48. Gonnet, G.H., Cohen, M.A. and Benner, S.A., Science, 256 (1992) 1433.
49. Shapiro, A., Botha, J.D., Pastore, A. and Lesk, A.M., Acta Cryst., A48 (1992) 11.
50. Labute, P. and Williams, C., J. Med. Chem., 44 (2001) 1483.
51. MATLAB™, Version 5.3, MathWorks, Inc., 1999.
52. PLS Toolbox®, Version 2.1, Eigenvector Research, Inc., 2001.
53. Gorre, M.E., Mohammed, M., Ellwood, K., Hsu, N., Paquette, R., Rao, P.N. and Sawyers, C.L., Science, 293 (2001) 876.