Journal of Molecular Graphics and Modelling 70 (2016) 275–283

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Topical Perspectives

Ensemble docking-based virtual screening yields novel spirocyclic JAK1 inhibitors

Dávid Bajusz, György G. Ferenczy, György M. Keserű *

Medicinal Chemistry Research Group, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Magyar tudósok krt. 2., Budapest 1117, Hungary

ARTICLE INFO

Article history:
Received 9 July 2016
Received in revised form 12 October 2016
Accepted 14 October 2016
Available online 17 October 2016

Keywords:
Janus kinase
JAK1
Kinase inhibitor
Virtual screening
Ensemble docking
Spirocyclic compounds

ABSTRACT

Small molecule inhibition of Janus kinases (JAKs) has been demonstrated as a viable strategy for the treatment of various inflammatory conditions and continues to emerge in cancer-related indications. In this study, a large supplier database was screened to identify novel chemistry starting points for JAK1. The docking-based screening was followed up by testing ten hit compounds experimentally, out of which five have displayed single-digit micromolar and submicromolar IC_{50} values on JAK1. Thus, the study was concluded with the discovery of five novel JAK inhibitors from a tiny screening deck with a remarkable hitrate of 50%. The results have highlighted spirocyclic pyrrolopyrimidines with submicromolar JAK1 IC_{50} values and a preference for JAK1 over JAK2 as potential starting points in developing a novel class of JAK1 inhibitors.

1. Introduction

Janus kinases are a family of tyrosine protein kinases associated with cytokine receptors, with crucial importance in signal transduction [1,2]. In particular, they are essential elements of the JAK/STAT signaling pathway, a key regulator of gene transcription [3,4]. Consequently, JAK inhibition has been proposed as a potential therapeutic intervention for various myeloproliferative and inflammatory diseases, including myeloproliferative neoplasms (MPNs) [5,6], rheumatoid arthritis (RA) [7], psoriasis [8] and inflammatory bowel disease (IBD) [9,10]. Recently, Sanz-Moreno and colleagues have identified JAK1 to contribute to the generation of actomyosin contractility, which is required to provide contractile force for tumor cell movement [11]. Thus, they have established the role of JAK1 in cancer metastasis and have concluded their work with proposing small-molecule JAK1 inhibitors as metastasis-blocking agents.

For the past two decades, significant research efforts have been dedicated to the discovery of small molecule JAK inhibitors [12]. So far, two drugs have been approved by the US FDA that act via the inhibition of Janus kinases: the JAK1/JAK2 inhibitor ruxolitinib [13] and the JAK1/JAK3 inhibitor tofacitinib [14,15]. While they have been demonstrated to be safe and effective in the treatment of myelofibrosis, polycythemia vera (ruxolitinib) and rheumatoid arthritis (tofacitinib), a novel generation of JAK inhibitors would be desired to diminish the side effects that occur during the use of these currently marketed drugs. A proposed and generally accepted strategy to achieve this goal would be to improve upon the subtype selectivity of these drugs within the JAK family (JAK1, JAK2, JAK3 and TYK2). In these efforts, medicinal chemists are facing a great challenge, as the ATP-binding sites of these and more generally, of all kinases share a great deal of homology and structural similarity. Alternative strategies might target other binding sites [16], however, the design of ATP-competitive agents is definitely the most accessible strategy, particularly for structure-based approaches.

Structure-based virtual screening has been shown numerous times to be an invaluable tool for providing novel starting points for kinase-directed drug discovery [17]. The current study can be considered a continuation of our recently published work, where we reported a successful virtual screen that yielded JAK2 inhibitors

Abbreviations: AUC, area under the (ROC) curve; EF, enrichment factor; FDA, Food and Drug Administration; FPR, false positive rate; HTVS, high throughput virtual screening; JAK, Janus kinase; LLE, lipophilic ligand efficiency; LELP, lipophilicity-corrected ligand efficiency; PME, particle mesh Ewald [method]; ROC, receiver operating characteristic [curve]; SP, standard precision; STAT, signal transducers and activators of transcription; TPR, true positive rate.

* Corresponding author.
E-mail address: keseru.gyorgy@ttk.mta.hu (G.M. Keserű).

http://dx.doi.org/10.1016/j.jmgm.2016.10.014
1093-3263/

with promising subtype selectivities [18]. In the current work, we have applied a similar ensemble docking protocol to provide novel inhibitors of JAK1 with even better, submicromolar potencies. In our previous study we have introduced an interaction fingerprint (IFP) based strategy to identify JAK2 selective inhibitors: in that work, an IFP scoring scheme that could discriminate selective JAK2 inhibitors has been developed. In this study – due to the smaller amount of available JAK1-related experimental data – we could not have implemented the prediction of subtype selectivity as a part of the virtual screening process, which in retrospect could have been a key reason for the better potencies and more favorable hit rate (50%) we report here. The hit compounds presented here are structurally novel, and three of them even define a broader compound class: 4-substituted spirocyclic pyrrolopyrimidines. Showing preference for JAK1 over JAK2, they are worth following up in future optimization studies aiming for selective JAK1 inhibitors.

2. Computational and experimental methods

2.1. Dataset preparation

We have utilized three sources to collect ligand structures and activity data for retrospective purposes: the ChEMBL Kinase SARfari [19], the Thomson Reuters Integrity database [20] and primary literature [21–31]. From the collected compounds, we have assembled two ligand sets: a training and a test set, containing 62 and 59 known JAK1 inhibitors, respectively. In the training set, we have included those ligands with good reported inhibitory activities towards JAK1 ($IC_{50} \leq 1\ \mu M$) and preference towards JAK1 instead of JAK2 (lower IC_{50} for JAK1 than JAK2). (While we have not considered subtype selectivity explicitly at this stage, our long-term aim is the design of JAK1-selective inhibitors.) These 53 compounds were supplemented with those ligands that have been cocrystallized with JAK1 (with a reported $IC_{50} \leq 1\ \mu M$), providing a total of 62 compounds for the training set. Due to the lack of further compounds conforming to the above mentioned criteria, the test set contains 59 ligands with excellent inhibitory activities reported for JAK1 ($IC_{50} \leq 70$ nM), but with a slight preference towards JAK2.

Both sets were mixed to a CDK2 decoyset (2074 compounds) downloaded from the Directory of Useful Decoys (DUD) [32]. (At the time of our research, CDK2 was the closest kinase to JAK1 – in terms of sequence homology – with a published decoyset in DUD.) The resulting two datasets with active:decoy ratios around 1:30 were used for training and testing the virtual screening protocol presented in this work.

For prospective purposes, the Mcule Purchasable Compounds Database (containing at the time approx. 5.1M compounds) was utilized [33].

All dataset operations were conducted with KNIME [34] and Instant Jchem from ChemAxon [35].

2.2. Ligand docking

We have applied a similar ensemble docking procedure as published in our recent work [18]. Schrödinger's Virtual Screening Workflow was applied for ensemble docking into five protein structures. First, LigPrep is used to prepare the ligands, including the removal of duplicates and high-energy tautomer states, generating stereoisomers and ring conformations, and generating protomers at a target pH of 7.4 with Epik [36,37]. Then, docking is carried out in two steps: first Glide HTVS is used for a faster docking step, and then the top 20% of best scoring poses are submitted to Glide SP docking [38,39]. In this step, establishing a hydrogen bond with at least one of the anchor groups of the hinge region of the kinase (Glu957, Leu959) is required. The hydroxyl group of Ser963 near the binding site is allowed to rotate. The final output is the best scoring pose for each ligand.

2.3. MD simulations

Desmond was applied to conduct an all-atom MD simulation [40–42]. The PDB entry 4IVC [31] was retrieved and solvated in TIP3P water [43] in a cubic box, keeping the original ligand and water molecules as parts of the system. Four sodium ions were added to neutralize the system and 29 more (along with 29 chloride ions) to set the salt concentration to 0.15M. Equilibration was carried out with the default, six-step equilibration protocol of Desmond, consisting of two minimization and four short simulation steps, as summarized in our recent work [18].

A 20-ns-long production run was conducted in the NPT ensemble at 300 K using a Nosé-Hoover thermostat [44,45] and a Martyna-Tobias-Klein barostat [46]. The Smooth PME method [47] was applied for the treatment of long-range ($d > 9.0$ Å) interactions. The RESPA integrator was applied for timestepping, employing a 2/2/6 fs multistepping scheme.

2.4. Virtual screening evaluation

The proposed virtual screening protocol was evaluated retrospectively with early enrichment factors and receiver operating characteristic curves. Following the suggestion of Jain and Nicholls, we have defined enrichment factors as the *TPR/FPR* ratios at specific points (*i.e.* specific FPR values) on the ROC curve (*ROC enrichments*), to provide a measure of enrichment that is independent of the size of the ligand set [48]. We also report "conventional" enrichment factors in the Supplementary information, defined as:

$$EF_{x\%} = (N_{act,x\%}/N_{x\%})/(N_{act}/N) \quad (1)$$

Here, $N_{act,x\%}$ and $N_{x\%}$ are the number of actives and the total number of compounds, respectively, in the top *x%* of the compound list (ranked by Glide docking score), while N_{act} and N are the number of actives and the total number of compounds in the whole dataset, respectively.

In the respective tables, we also report 95% confidence intervals of the mentioned performance parameters, to provide an error estimate for them [49].

2.5. JAK inhibition measurements

JAK1 inhibition of the hit compounds was tested in a Z'-LYTE kinase inhibition assay of Life Technologies. Z'-LYTE employs a fluorescence-based, coupled-enzyme format and is based on the differential sensitivity of phosphorylated and non-phosphorylated peptides to proteolytic cleavage.

The details of the measurement protocol are identical to that presented in our recent work [18]. A more detailed description of the process is available on the website of Life Technologies [50].

3. Results and discussion

3.1. Retrospective screening

In a computational sense, retrospective screening can be thought of as the training procedure of the virtual screening protocol that we present and apply here. While the protocol itself is similar to the one we presented in our recent work involving JAK2 inhibitors [18], the individual steps had to be optimized nonetheless (X-ray structures had to be selected, datasets assembled, *etc.*). The procedure is reported in this subsection (and in the Experimental section in more detail) and summarized in Fig. 1, followed by its prospective application in the next subsection.