Journal of Molecular Graphics and Modelling 27 (2009) 860–870

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

A comparison of ligand based virtual screening methods and application to corticotropin releasing factor 1 receptor

Gary Tresadern [a,*], Daniele Bemporad [b], Trevor Howe [b]

[a] Johnson & Johnson, Pharmaceutical Research & Development, Janssen-Cilag S.A., Calle Jarama, 75, Poligono Industrial, 45007 Toledo, Spain
[b] Johnson & Johnson, Pharmaceutical Research & Development, Janssen Pharmaceutica N.V., Turnhoutseweg 30, 2340 Beerse, Belgium

ARTICLE INFO

Article history:
Received 13 December 2008
Received in revised form 12 January 2009
Accepted 14 January 2009
Available online 23 January 2009

Keywords:
Corticotropin releasing factor
CRF1
CRH1
Scaffold-hopping
Ligand based virtual screening

ABSTRACT

Ligand based virtual screening approaches were applied to the CRF1 receptor. We compared ECFP6 fingerprints, FTrees, Topomers, Cresset FieldScreen, ROCS OpenEye shape Tanimoto, OpenEye combo-score and OpenEye electrostatics. The 3D methods OpenEye Shape Tanimoto, combo-score and Topomers performed the best at separating actives from inactives in retrospective experiments. By virtue of their higher enrichment the same methods identified more active scaffolds. However, amongst a given number of active compounds the Cresset and OpenEye electrostatic methods contained more scaffolds and returned ranked compounds with greater diversity. A selection of the methods were employed to recommend compounds for screening in a prospective experiment. New CRF1 actives antagonists were found. The new actives contained different underlying chemical architecture to the query molecules, results indicative of successful scaffold-hopping.

1. Introduction

The corticotropin releasing factor 1 (CRF1) receptor is a target of interest in neuroscience [1]. The endogenous 41-amino acid peptide ligand, CRF (also known as corticotropin releasing hormone), regulates the body's response to stress through the release of adrenocorticotropic hormone (ACTH). CRF has been shown to mediate stress-induced changes in the autonomic system and to cause neuroendocrine and behavioural effects [2]. Clinical data has shown that patients with depression and post-traumatic stress disorder show significantly elevated concentrations of CRF in cerebrospinal fluid and may have down regulated CRF receptors [3–5]. Two receptor subtypes, CRF1 and CRF2, have been identified and shown to be widely distributed throughout the central nervous system (CNS) and periphery [6]. Compound R121919, Fig. 1, was the first CRF1 antagonist to be evaluated in a phase IIa clinical trial for depression and showed positive effects [7]. A selective CRF1 receptor antagonist would therefore represent a novel class of compounds for the treatment of anxiety, depression, and stress-related diseases.

The CRF1 receptor is a secretin-like class B GPCR [8]. A variety of NMR techniques have been used to solve the structure of the isolated extracellular N-terminal domain [9]. The small molecule antagonist binding domain, however, is believed to be located in the 7-transmembrane region [10]. Since the first reported non-peptidic small molecule CRF antagonists in 1991 the number of known compounds has increased significantly and have been reviewed elsewhere [11–13]. There are in excess of one hundred CRF composition of matter patents and although there are many chemically distinct series they often share the same pharmacophoric features. The reported CRF antagonists generally show high affinity and selectivity for the CRF1 receptor. Representative CRF1 compounds [14–19] are shown in Fig. 1. Given the expanding prior art it is increasingly difficult to find novel intellectual property space for CRF1 chemistry. For this reason we considered a virtual screening approach to identify new active compounds for this target.

High throughput screening (HTS) against a discrete biological target is a preferred technique for finding active starting points in early phase drug discovery. Confirmation and follow-up assays generally provide compounds of sufficient interest for a medicinal chemistry program to begin. Complementary to this, virtual screening can be used as a process for ranking molecules by their potential for activity against the target [20]. Therefore assuming the virtual screening method is of sufficient accuracy the potential benefits are substantial in terms of screening time, compound provision, library depletion and protein supply. As such, virtual screening of compound collections prior to assay or virtual libraries prior to chemical synthesis is widespread [21,22]. Available methods for virtual screening are often classified as structure or ligand based and the choice of which to apply is dependent on available information or desired outcome [23–25].

* Corresponding author. Tel.: +34 925 24 5782.
E-mail address: gtresade@its.jnj.com (G. Tresadern).

1093-3263/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2009.01.003

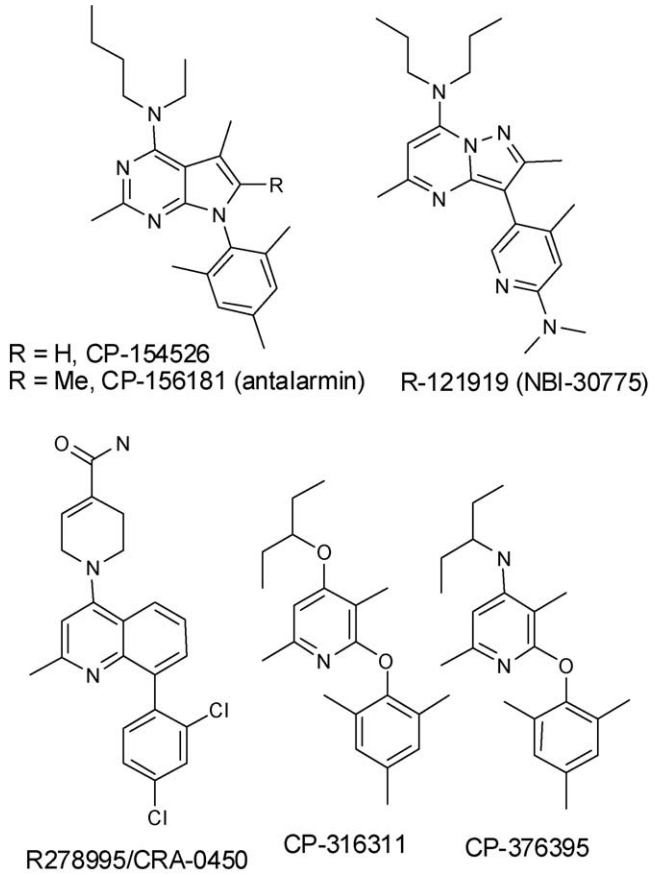


Fig. 1. Representative CRF1 compounds CP-154526 [14], antalarmin [15], R121919 [16], R278995 [17], CP-316311 [18] and CP-376395 [19].

Structure based virtual screening utilizes a 3D representation of the biological target whereas ligand based approaches have no such requirement. Ligand based virtual screening ranks compounds by their similarity towards known active ligands. In this work we compare a selection of ligand-based methods. There are thousands of 2D molecular property descriptors which could potentially be used for the calculation of molecular similarity [26,27]. Atom based 2D fingerprints describe the connectivity in molecules but tend to be limited to identifying analogues of close chemical structure, one example is extended connectivity fingerprints (ECFP) from Scitegic [28]. More abstract 2D molecular descriptors such as Feature Trees (FTrees) are based on graph representations of molecules and have less dependence on the underlying 2D structure [29]. Where molecules are represented and compared by their 3D properties the challenge of conformational flexibility and alignment arises. One such method which seeks to overcome the alignment problem is Topomers. Individual 3D fragments of the input molecules are aligned by their valence bonds [30]. The subsequent molecular comparison is performed using fragment steric fields with the inclusion of pharmacophoric features. Methods such as ROCS (Rapid Overlay of Chemical Structures) perform 3D shape based similarity on any number of user-supplied conformations which are often the output of a conformational search [31,32]. The whole molecules are aligned and the Tanimoto difference between their 3D steric fields is calculated. Throughout this work we refer to this method using the acronym OEST (OpenEye Shape Tanimoto). The ROCS colour score includes the option to measure 3D similarity with a feature-based definition of atom-centres and there is also the choice to combine both steric and colour-score, the so-called combo-score denoted as OECS (OpenEye Combo-Score). Comparison of the electrostatic potential of aligned molecules can be performed by means of the program EON which we refer to as OEET (OpenEye Electrostatic Tanimoto) [33]. The Cresset-Fieldscreen approach also performs complete 3D conformational analysis of compounds using a custom forcefield [34]. Subsequent molecular comparison uses four different 3D fields, positive and negative charge, steric shape and hydrophobicity. To render the comparison computationally tractable for virtual screening of large databases the fields are simplified to their maxima and minima [35,36]. The 3D shape and electrostatic field descriptors encode molecular characteristics likely to be important for biological recognition. However, given their more abstract nature they are less dependent on the underlying atom connectivity. As such they are well suited for scaffold-hopping which ideally should be ignorant of covalent structural frameworks. This is a topic receiving increasing attention [32,37–40].

Given the lack of complete receptor 3D structure and abundance of small molecule ligands the CRF1 target is well suited to a ligand based virtual screening approach. We were particularly interested in finding new chemical series and as such employed 3D field based methods with scaffold-hopping potential. In this work we describe the comparison of a variety of commercially available techniques, FTrees, Topomers, ROCS (OEST and OECS), EON (OEET) and Cresset-Fieldscreen. The scitegic 2D fingerprint method, ECFP6, and simple descriptors such as MW, ALogP and element counts are used for comparison. We performed both retrospective and prospective virtual screening analyses. For the retrospective analysis we searched over compounds previously screened for CRF1 antagonism. We analysed the ability of the different methods to identify known active and reference compounds seeded into CRF1 inactive compounds from an in-house HTS. We also investigated the ability of the methods to successfully scaffold-hop. To achieve this we report the retrieval rate of different active scaffolds from amongst the active set. For the prospective work we used a selection of the same methods to recommend new compounds for screening versus CRF1. We demonstrate that 3D methods such as OEST and OECS had the best enrichment at identifying known active compounds. In addition the same methods retrieved more new active scaffolds amongst the proportion of ranked compounds studied. Methods such as OEET and Cresset appear to offer the most diversity when compared with fingerprints. The latter two methods delivered more new scaffold diversity amongst a fixed number of confirmed actives; however, it was at the cost of a lower active retrieval rate.

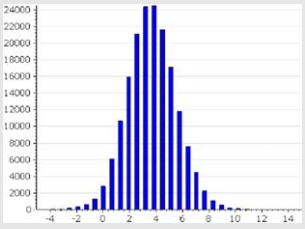
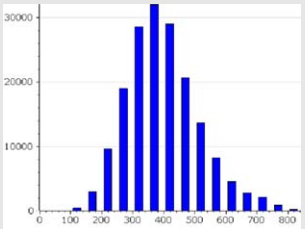
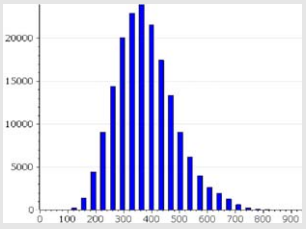
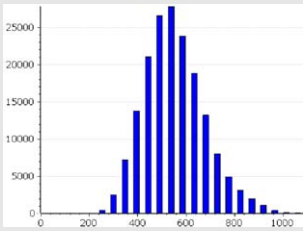
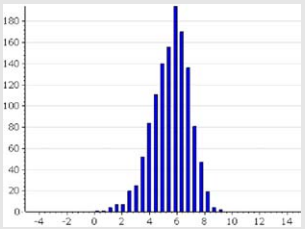
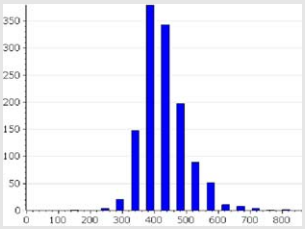
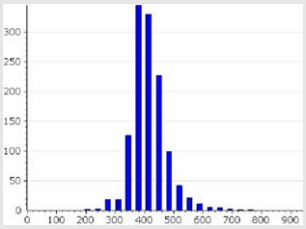
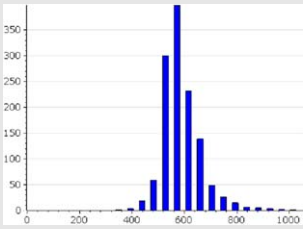
2. Computational and experimental details

2.1. Data set and query molecules

The dataset used for the retrospective analysis consisted of 1261 active and 175,196 inactive HTS molecules. The active set was a combination of 899 reported reference compounds extracted from CRF1 patents and 362 in-house CRF1 compounds with antagonistic pIC_{50} activity >6. The structures for the 899 reference actives are provided in the supplementary information accompanying this work. The confirmed inactive compounds were taken from the output of a previous in-house CRF1 antagonist HTS. For the prospective part of this work we searched over available compounds from the Johnson & Johnson corporate collection. The compounds in the datasets were considered in their neutral form. The datasets were converted into the required searchable database format for each software package. The case-by-case details are given in the following sections.

To shed further light on this proprietary dataset we compared histograms of simple descriptors ALogP, MW, surface area and volume, Table 1. The plots help to understand the potential elementary differences between actives and inactives. All plots

Table 1
Histograms of selected simple molecular descriptors comparing actives and inactives used in the retrospective dataset.

	AlogP	MW	Surf. area[a]	Vol[a]
Inactive				
Active				

[a] Descriptors calculated in pipeline pilot using a 2D method [28].

show unimodal behaviour. It can be seen that the distributions of the active set are narrower than the inactives. This is to be expected given that the actives are target focused compounds whereas the inactive set comprises over 175,000 compounds used for HTS against a range of biological targets. One may expect that the virtual screening approaches ought to easily discard compounds at the extremes of the inactive descriptor space. This may result in higher enrichment rates but does not affect the comparison between methods which is the focus of this work. For MW, surface area and volume the mean of the actives is well centred compared to the inactives. However, for ALogP it is clear that the actives have higher lipophilicity than the inactives. Therefore a method which places particular bias towards more lipophilic compounds is likely to perform well on this dataset.

We were particularly concerned with measuring the ability of the methods to scaffold-hop. Therefore we analysed the active subset of 1261 compounds and manually defined 89 different scaffold chemotypes which had at least two molecules per scaffold. Some of the most frequently occurring examples are shown in Fig. 2. We used these scaffolds to define scaffold enrichment plots. The scaffold retrieval rate was incremented each time a new scaffold from the defined list was identified in an active compound. Each scaffold was counted only once, the first time encountered. The data was calculated for each method with each query molecule and also averaged over all query molecules.

When selecting compounds to use as queries for virtual screening it is often useful to select those with a known binding conformation. As this was not available we used a CRF1 antagonist pharmacophore to generate the 3D search conformation where required. During previous project work we had used a subset of the active compounds in the dataset to build a pharmacophore with MOE [41]. A hypothesis generated by the pharmacophore elucidator tool was augmented via the inclusion of additional features based on project experience. The key features of the resultant pharmacophore, Fig. 3, are consistent with other reported examples [42,43]. The pharmacophore contains an aromatic scaffold which can be either monocyclic, bicyclic or tricyclic. The scaffold contains an acceptor atom which is usually an aromatic nitrogen. It is common for the scaffold to have three substituents. The first is a small hydrophobic group, often methyl, adjacent to the scaffold acceptor. The second is an often aliphatic lipophilic group frequently represented by open or cyclic carbon chains. We included an optional acceptor group in this distal part of the pharmacophore. The third group is an aromatic ring with an orthogonal orientation with respect to the scaffold. The conformational twist effect is produced by ortho-substituent(s). Also, to provide additional confidence in the 3D conformation of the query compounds we only considered actives with less than or equal to five rotatable bonds. Filters were also applied to molecular weight (MW) of 400 and number of H-bond acceptors <5. This excluded molecules containing too many non-essential features. From the remaining subset of actives four compounds were selected by hand which covered different chemical series, Fig. 3. Throughout this work the queries are referred to as query 01, 02, 03 and 04 respectively. Conformations of the four selected query molecules were generated using the stochastic conformational search with the MMFF94x force field available in MOE. The conformations selected were those with

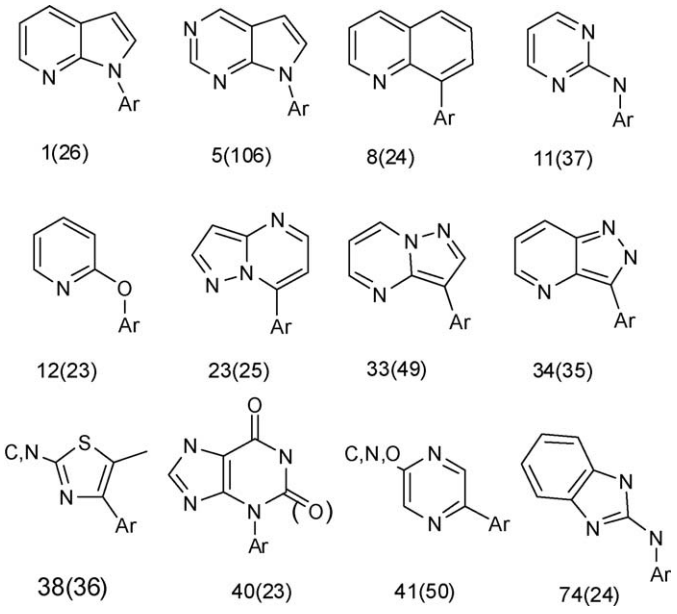


Fig. 2. Examples of frequently occurring scaffolds in the active compound subset. First number is scaffold id and number in parentheses is the frequency of occurrence.

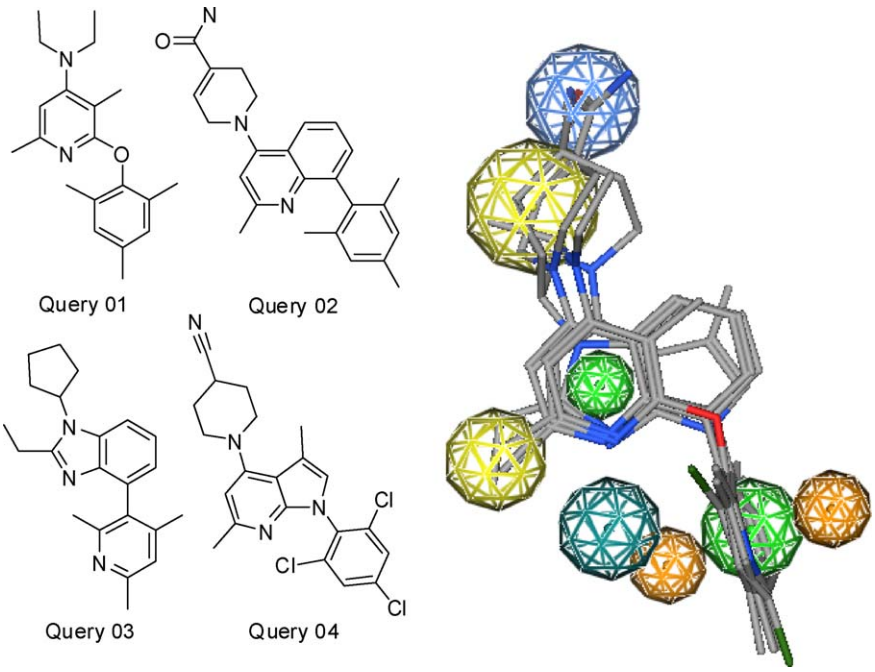


Fig. 3. Query molecules 01–04 and their overlay in a CRF1 antagonist pharmacophore. Pharmacophore features are identified by their colour, aromatic green, projected aromatic orange, acceptor blue/cyan and hydrophobe yellow.

the lowest RMSD to the pharmacophore. These were retained and used as input for searches requiring 3D conformations.

2.2. Computational methods

Extended connectivity fingerprints (ECFP6) fingerprints were generated with SciTegic's Pipeline Pilot software [28]. Each feature represents the presence of a structural unit. A molecule may have a fingerprint consisting of tens or hundreds of individual features. The features are constructed from the atom connections, element type, charge and mass. Structural units were only calculated up to a maximum diameter of six bonds with respect to each central atom. The set of all features is the fingerprint. The Tanimoto coefficient was chosen to quantify the similarity between molecules in the 2D virtual screening process

$$\text{Tanimoto} = \frac{\text{FC}}{\text{FA} + \text{FB} + \text{FC}}$$

This is the number of features present in both query and target molecule (FC) divided by the sum of itself, the number of features present only in the query (FA) and the target (FB). The coefficient ranges between 0, for molecules with no features in common, to 1, in the case of identical molecules.

Feature Trees (FTrees) descriptors although a 2D method are somewhat more abstract than fingerprints in that they use a graph representation of features [29]. The descriptor represents hydrophobic fragments and functional groups of the molecule as nodes and encodes the way they are linked together, therefore defining a tree-like topology rather than a more common bit string or vector. Each node is labelled with a chemical and steric feature computed from the molecular fragment the node represents. The steric feature is the van der Waals volume and the number of ring closures. The chemical feature incorporates the molecular interactions the fragment can form, such as aromatic, hydrophobic, donor or acceptor. The similarity between two FTrees is defined as the score for the best possible alignment or superposition of the two trees, which is performed by comparing their associated subtrees. The technique is well suited to scaffold-hopping as the underlying chemical scaffold is disregarded and the descriptor reproduces the connectivity of the node features and volume. The datasets for FTrees were prepared according to the guidelines by the program authors. Formal molecular charges were assigned using the MOE wash function and compounds exported with defined atom types.

Topomers (dbtop) is a method for the 3D comparison of molecules [44]. Query and dataset molecules are fragmented across all acyclic bonds. The open valence of fragments is filled with a cap atom and one single 3D conformation of each fragment is generated with the CONCORD program [45]. The fragment is then oriented in space according to alignment with a vector defined along the valence bond. The remainder of the 3D fragment is oriented using a rule-based scheme defining the direction of each proceeding atom based on its atomic mass. Spatially aligned fragments can then be compared based on their 3D steric field. The steric fields are generated in a similar way to CoMFA using an sp^3 carbon atom probe on a 2 Å lattice surrounding the fragment. The steric field differs from the CoMFA equivalent as the atomic contribution diminishes with distance from the valence bond, being scaled according to the formula 0.85^n where n is essentially the number of intervening double bonds [46]. The absence of pharmacophoric features in 3D space yields a penalty to the similarity between two fragments. The features include aromatic, positive, negative, H-bond donor and acceptor. The total topomer similarity (also called topomer distance) between two molecules is calculated in a geometric fashion as $(A^2 + B^2)^{1/2}$ where A and B are the topomer similarity between two constituent fragments. Dataset and queries were prepared in an identical way using the dbtranslate utility with the +standardize and +concord option to generate the single conformation required for each molecule.

Openeye ROCS Shape Tanimoto (OEST), ComboScore (OECS) and EON (OEET) are all 3D comparison methods. The ROCS program uses multiple conformations per molecule as input and outputs 3D aligned molecules with the calculated OEST and OECS similarity metrics. The EON program uses the aligned molecule output from ROCS to calculate the electrostatic field similarity. OMEGA was used to generate conformers for ROCS. OMEGA is a rule-based

method. It disassembles each molecule into fragments, whose conformations are either retrieved from a pre-generated library or constructed on-the-fly using distance constraints. These are used as starting points for a torsional search around rotatable bonds and finally the fragments are reassembled together to yield the conformations for the entire molecules. In ROCS shape based comparison atoms are approximated to Gaussian distributions to make the problem computationally tractable. ROCS aligns each conformation of the target molecule onto each conformation of the query, and then keeps the alignment with the highest shape overlap. The shape Tanimoto (OEST) similarity is given by the following:

$$\text{Shape Tanimoto} = \frac{\text{VC}}{\text{VA} + \text{VB} - \text{VC}}$$

where VC is the volume in common between the 2 molecules, VA is the total volume of the query molecule, VB is the total volume of the target molecule. The coefficient ranges from 0, no shape overlap, to 1, identical shape.

Additional similarity metrics are calculated by ROCS. The colour force field is used to measure chemical complementarity and the 3D overlap is refined based on the matching of simple pharmacophoric features. In the default implementation of the force field used here the features are H-bond acceptor and donor, hydrophobes, anions, cations and rings. We have not directly considered the colour force field similarity but instead have used an additional metric from ROCS, combo-score. The OpenEye ComboScore (OECS) coefficient is the sum of the shape Tanimoto and colour-score.

In EON, the Poisson–Boltzmann equation is solved to calculate the electrostatic potential on the nodes of a 3D grid constructed around a molecule. The molecule has MMFF94 nuclei-centred charges assigned. The OpenEye Electrostatic Tanimoto (OEET) similarity is given by:

$$\text{Electrostatic Tanimoto} = \frac{A{*}B}{A{*}A + B{*}B - A{*}B}$$

where A is the electrostatic potential around the query molecule and B is the electrostatic potential around the target molecule. The coefficient ranges between $-1/3$ when the two molecules have equal but opposite potentials, and 1, when the two molecules have identical potentials. EON does not generate new alignments to maximise the electrostatic Tanimoto, it only scores the alignments generated from the ROCS shape overlap.

For the dataset preparation conformers were generated with OMEGA 2.1.0 using default parameters. This included a maximum of 400 conformers per compound. The maximum of the energy window for allowed conformers was reduced from 25 to 10 kcal mol^{-1}. The force field used was the 94 s variant of the MMFF without coulomb interactions and the attractive part of the van der Waals interactions. The virtual screening was performed with ROCS 2.2 and EON 1.1, using default parameters. EON was run on compounds having Shape Tanimoto coefficient greater than 0.6 from ROCS.

The Cresset-Fieldscreen approach considers multiple 3D conformations of each compound. The conformers are generated with a custom force field based on non-atom centred extended electron distribution charges (XED) [34]. Four different 3D field properties are then considered. The molecular electrostatic potential is calculated based on the XED charges. From this the negative and positive electrostatic charge fields are derived using the oppositely charged probe atoms. A steric shape field is calculated via a neutral probe and the van der Waals radii. A hydrophobic field uses a similar neutral probe and parameters weighted towards hydrophobic atoms. All fields are then simplified to maxima and minima which are stored for each conformer and are used in the similarity calculations.

The XED force field was used within the XedeX program to generate a multi conformer version of the J&J database. A Monte Carlo process randomizes rotatable bonds and the resultant conformers are minimised and closely related and high energy conformers removed. A maximum of 50 conformers per compound were retained. The Cresset-Fieldscreen method combines a step wise approach to database searching. In the first step the most dissimilar compounds are discarded based on a fast comparison of the 1D encoded field prints. The field prints are calculated from the distance matrix of field points for each conformer. The remaining top ranking 50% of the database was subjected to similarity assessment via the 3D field points, so-called clique refinement. In the final step the subsequent top ranking 20% of the output from the clique refinement were re-ranked via an optimised alignment. A simplex minimiser is used to optimise with respect to both field similarity and volume of steric overlap.

3. Results and discussion

3.1. Retrospective enrichment analysis

The enrichment plots for the searches with each of the four query compounds using all methods are shown in Fig. 4. The plots show the percent of actives captured versus the top ranked 1.2% of the dataset. This percentage corresponded to ~2100 compounds. A sub-selection for screening is often made from within such a number. Frequently compounds from virtual screening need special compound handling and formatting requirements which prevents considering a bigger number. In each plot random and perfect behaviour are indicated. It should be noted that the line of random behaviour runs close to the abscissa across the bottom of the plots.

Immediately striking was the good performance of all methods with all query molecules. The virtual screening methods performed significantly better than random at identifying the 1261 active compounds from the entire set of HTS inactives. In general queries 02, 03 and 04 showed similarities in performance of the methods whereas query 01 was somewhat different.

Firstly considering queries 02, 03 and 04 the best performing methods were OEST and OECS. The latter reached 60% of actives captured amongst the top ranked 1.2% of the dataset. This corresponded to 757 active compounds amongst the top 2100 for the OECS results with query 03. The best methods appear to plateau at around 60% retrieval of actives. This is likely a manifestation of the false negative problem which blights ligand based virtual screening. Active compounds with significantly different shapes cannot be identified and therefore some series in the active set were missed. Also for the same three queries Topomers performed well capturing between 40 and 50% of actives within the top ranked 1.2% of the dataset. With the exception of the OEET search with query 04, FTrees and all 3D approaches outperformed fingerprints (ECFP6) for queries 02, 03 and 04. As the active set of compounds was rich in scaffold diversity this implies that these methods are more successful at scaffold-hopping. Queries 02, 03 and 04 show that performance of the Cresset and EON methods was worse than OEST, OECS and Topomers. This was somewhat surprising as the former two methods offer more detailed 3D field descriptions of the molecules whereas the latter are predominantly 3D steric descriptors with the inclusion of pharmacophoric features. However the result suggests that CRF1 actives are best separated from the inactives by methods which place a larger emphasis on steric considerations. In addition the similar performance of OEST and OECS suggests that steric factors are sufficient to achieve the best enrichment and "features" offer little benefit. The 3D shape of a large proportion of the CRF1 actives is not well represented amongst the inactives and is an important

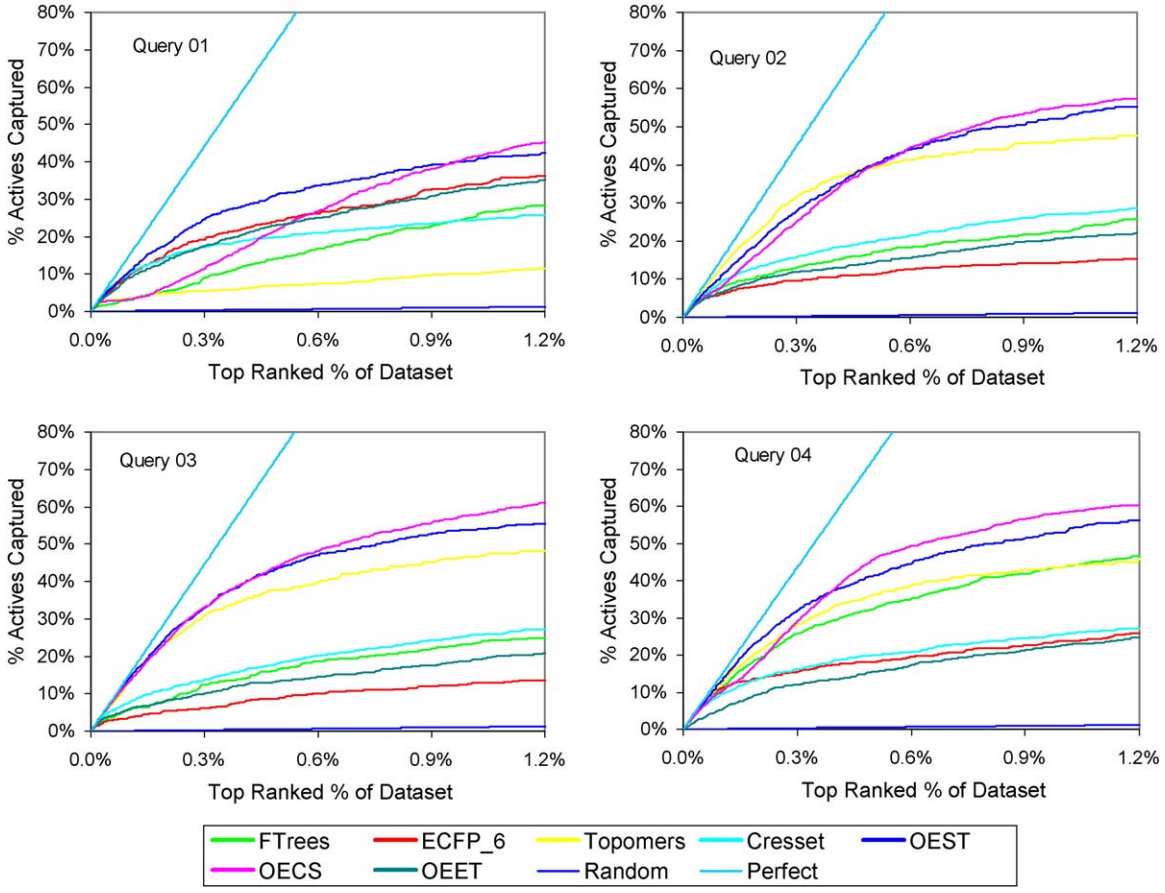


Fig. 4. Enrichment plots showing the retrieval of actives versus % of database searched for each method with the four query molecules. The top ranked 1.2% of samples corresponded to ~2100 compounds.

discriminator for CRF1 activity. However it cannot be said that electrostatics or features are not important for CRF1 activity just that the descriptors do not offer a noticeable benefit to enrichment for this dataset.

Turning to the enrichment curves for query 01, despite a poor active retrieval rate by OECS method over the first ~0.2% of the dataset it went on to perform the best capturing greater than 40% of actives within the top 1.2%. It is noticeable that the OEST method selected many more actives amongst the initial ~0.2% of ranked compounds than OECS. Given that OECS is a combination of OEST plus colour-score the difference between the OECS and OEST curves reflects the impact of features from the colour-score. Close inspection of the top ranked 0.2% of samples from OECS identified a series of low MW inactive compounds containing the same scaffold and aromatic features as the query. They were ranked significantly higher with OECS than by OEST. Two examples of the series are shown in Fig. 5. The inactivity of the compounds may be rationalized due to the lack of certain pharmacophoric elements such as the aliphatic hydrophobic substituent on the scaffold. They were lower ranked by the 3D shape only method, OEST. However, when combining with feature matching via OECS the 3D shape had less weighting in the overall similarity and new emphasis was given to the matching features. Likewise the Topomers method identified the same series of inactives. This illustrates a drawback of the Topomer approach that it tends to over bias towards compounds that contain a very close matching or near identical fragment shared between the two molecules being compared. This is likely due to the geometric approach for calculating the Topomer distance between molecules. The Cresset method performed similarly for query 01 as for queries 02, 03 and 04. In relative terms this is an improvement compared to the other methods. Also of note was the high performance of ECFP6 fingerprints for this query.

The poorer performance of the 3D methods for query 01 compared to query 02, 03 and 04 can be partly explained by its different MW and volume compared to the active set. The MW and van der Waals volume of query 01 were 312 and 401 Å^3 compared to the average values for the active set of 430 and 495 Å^3 respectively. The other three query molecules were closer to the mean, with MW's of 386, 333 and 434 and volumes of 483, 423 and 474 Å^3 for queries 02, 03 and 04 respectively. Bender and Glen have shown that simple descriptors such as MW and especially element counts can deliver enrichment better than random [47]. Therefore

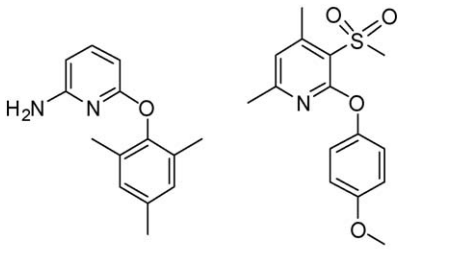


Fig. 5. Two inactive pyridyl hits from the searches with query 01. OECS and Topomers ranked these compounds much higher than the 3D shape only method, OEST.

Table 2
Enrichment factors at 1% of database retrieval compared to random.

	Query molecule			
	01	02	03	04
ECFP6	33.9	14.4	12.7	23.7
FTrees	25.1	22.5	23.2	43.6
Topomers	10.2	46.3	46.6	43.5
OEST	40.1	51.9	53.8	53.0
OECS	41.2	55.2	57.7	58.3
OEET	32.8	20.7	18.8	22.5
Cresset	24.3	27.1	25.5	25.5
ALogP	2.6	1.9	1.9	3.6
MW	0.2	1.7	0.3	2.1
Surf area[a]	1.0	2.5	0.5	2.3
PSA[a]	1.5	1.2	2.3	1.3
Vol[a]	0.3	1.9	1.0	2.1
Element counts	1.3	1.0	3.1	6.8

[a] Descriptors calculated in pipeline pilot using a 2D method [28].

it is understandable that query 01, which has the largest difference in MW and volume between queries and dataset actives, would perform worse.

It was of interest to consider if the high enrichments shown in Fig. 4 were due to elementary differences between actives and inactives which could be identified with simpler descriptors. In Table 2 we report the enrichment factors at 1% of the retrieved dataset for methods from Fig. 4 alongside ALogP, MW, surface area, polar surface area (PSA), volume (Vol) and element counts. For the property descriptors similarity was assessed by rank order of the absolute difference between query and dataset compounds. Element count similarity was calculated as the sum of the absolute differences between elements in the query and dataset molecules. Compounds were then ranked accordingly. The latter is analogous to the atom counts used by Bender and Glen where they reported examples of greater than tenfold enrichment with this "dumb descriptor" [47]. Here we see in Table 2 that the simple descriptors performed best on query 04 suggesting it was the easiest for separating actives from inactives. The element counts descriptor displayed an enrichment of 6.8 compared to random at 1% of the ranked dataset. For query 01 the MW and Vol descriptors picked more inactives than expected compared to random as shown by enrichment factors of 0.2 and 0.3 respectively. This was expected given the difference in MW and volume of query01 compared to the actives as discussed above. Interestingly for this dataset ALogP and PSA consistently displayed enrichment for all four query molecules. The enrichment for ALogP was generally higher than PSA. The two descriptors manifest the same effects due to their inverse correlation. For ALogP the enrichment factors varied between 1.9- and 3.6-fold. From the comparison of the actives and inactives discussed in Section 2.1 it was shown that the actives have higher lipophilicity than the inactives. This explains the origin of the ALogP enrichment. For three of the query molecules the simple descriptors were able to show at best close to 3-fold enrichment. For query 04 this was even higher at 6.8-fold enrichment. Therefore these values add context to the relatively high reported enrichments of over 40- or 50-fold displayed by the more complex methods.

The results observed from query 01 demonstrate how performance of virtual screening changes depending on query selection. They also highlight the difficulties in using enrichment curves to compare virtual screening methods. From these data some compounds presented a particular challenge to the OECS and Topomers methods. Recently McGaughey et al. compared virtual screening methods with searches over both public and Merck proprietary datasets [48]. They too noticed a difference in performance for ligand based virtual screening depending on the datasets. This seems entirely reasonable given our observations for query 01. Retrospective enrichment analyses are very dependent on the presence of certain molecules in the dataset that are able to challenge a given computational method.

We have calculated the average active retrieval rates for each method over the four query molecules. The average performance is dominated by the comparable results for queries 02, 03 and 04 discussed above. The OECS and OEST methods performed similarly and were superior. The ranking then followed Topomers > FTrees > Cresset > OEET > ECFP6. The analogous performance of OECS and OEST again suggests that the feature contribution provided by the colour score does not greatly improve the ability to retrieve active CRF1 compounds. The Cresset and OEET methods displayed similar ability to identify actives. From the average plot ECFP6 fingerprints performed the worse, which is consistent with the need for 3D shape to best identify the CRF1 actives. All plots are given in supplementary information, Fig. 1S.

3.2. Retrospective scaffold-hopping analysis

The retrieval rates for active scaffolds are shown in Fig. 6. The plots cover the top ranked 2000 dataset compounds by each method for each of the four queries. The scaffold retrieval rate was incremented the first time one of the predefined active scaffolds was identified in an active compound. As the active set was quite rich in scaffolds the plots mirror to some extent the retrieval rates of active compounds discussed previously. Similar performance was seen for the searches with queries 02, 03 and 04 whereas query 01 exhibited different behaviour. Nevertheless the OEST method performed best overall at identifying different scaffolds with OECS a close second. The feature definitions which are part of the colour-score included in the OECS method appear to be detrimental for the scaffold-hopping ability as the shape only method OEST performs marginally better for three out of four queries. Nevertheless, both approaches retrieved greater than 80% of the defined scaffolds for queries 02, 03 and 04. Results for Topomers, FTrees and ECFP6 fingerprints were consistent with the retrieval rates of active compounds.

Fig. 7 shows the average over the four queries for the retrieval rate of active scaffolds amongst the top 2000 ranked compounds in the dataset. The OEST method performed best and the OECS method reached a similar proportion of scaffolds after considering the top ranked 2000 compounds. Interestingly the OEET and Cresset methods, which use the most detailed 3D field descriptors, on average identify more active scaffolds than methods such as Topomers and FTrees. This is despite identifying fewer active compounds than the latter two approaches. OEET and Cresset also out-perform OECS over the first four hundred ranked compounds from the dataset whilst again identifying fewer active compounds. These methods were superior to Topomers, FTrees and ECFP6 fingerprints at scaffold-hopping in this study. However, it should be noted that the average performance of Topomers in Fig. 7 is influenced significantly by the poor performance for query 01 whereas for queries 02, 03 and 04 it performed better. As discussed in Section 3.1, Topomers performed badly for query 01 due to its approach of fragmentation and subsequent similarity metric. Of all the methods it performed worse with the query molecule that had the biggest difference MW and volume compared to the actives.

Fig. 8 shows a plot of % scaffolds retrieved versus the top ranked two hundred active compounds. For the OEET method this corresponded to an average of ~50% of the active scaffolds. The Cresset method was the second best in this regard having identified close to 45% of active scaffolds amongst the same number of actives. Therefore within a given number of known active compounds the OEET and Cresset methods contain more scaffolds than OEST and OECS methods. However, this benefit in

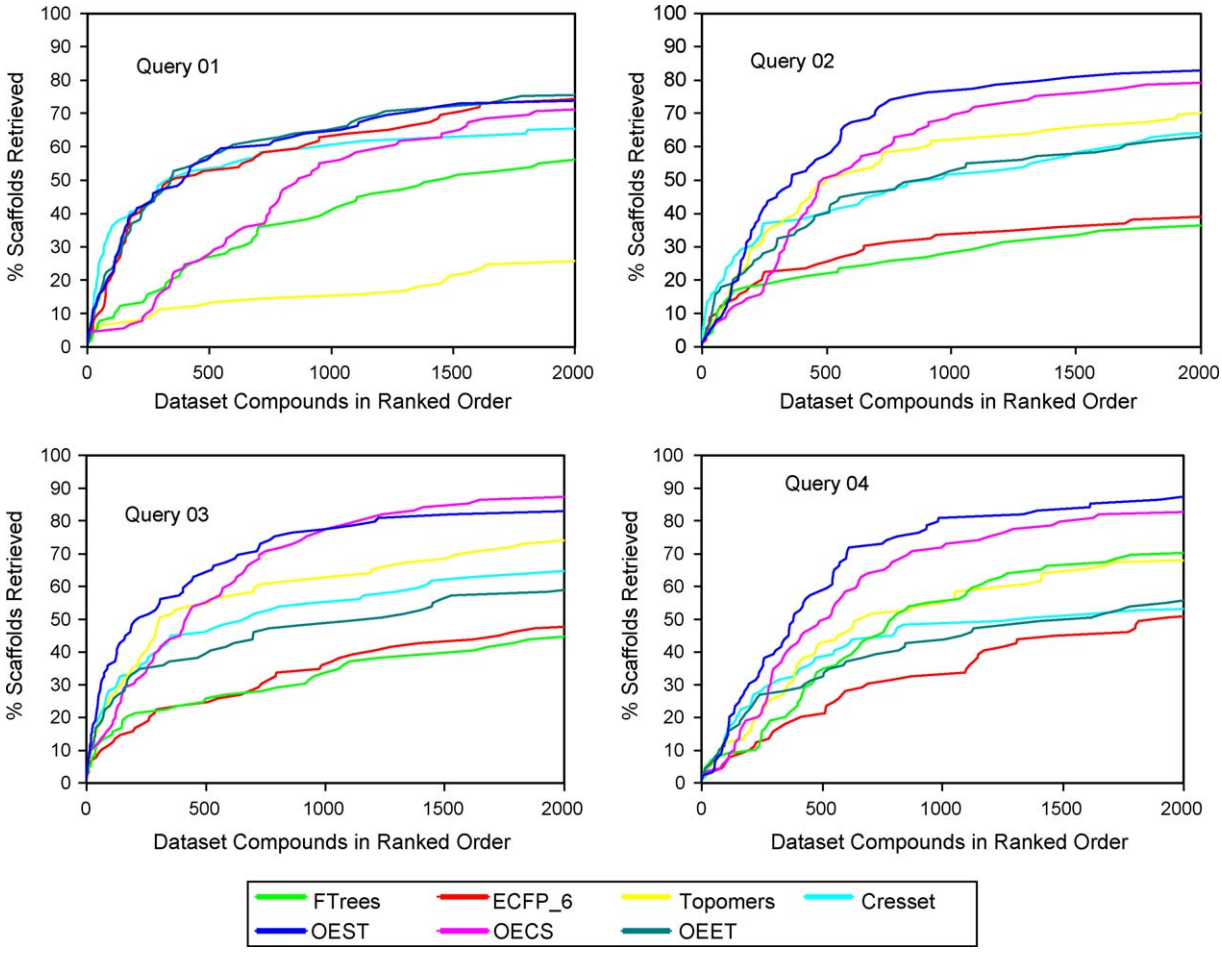


Fig. 6. Plots of retrieval rates of active scaffolds amongst the top ranking 2000 hits for each query molecule with the different virtual screening methods.

diversity must be traded with the worse performance at identifying actives in this dataset.

We invested further efforts to understand the diversity of the compounds selected by the different methods. Figures are given in the supplementary information that accompanies this work. One of the several approaches we considered was to calculate the distribution of Tanimoto ECFP6 fingerprint nearest non-identical neighbour similarity for top ranking compounds from each search, Fig. 2S. The plots clearly show that Cresset and OEET methods select compounds with more diverse nearest neighbours. This is valuable in itself, suggesting that these two methods offer something beyond the other approaches and more distinct to a fingerprint. However, it must be remembered that the enrichment

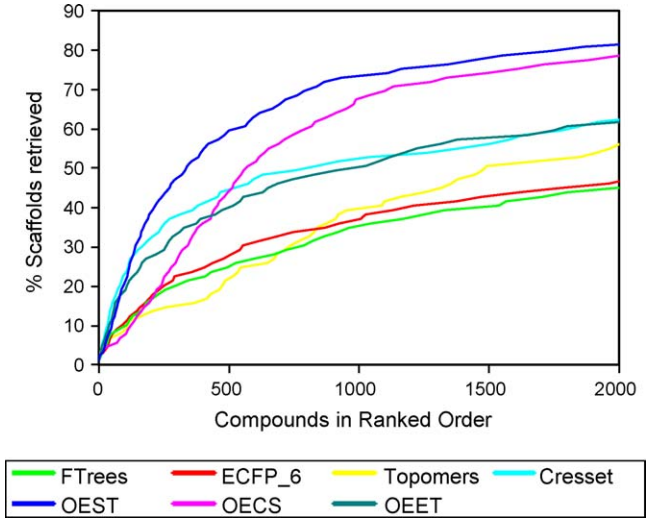


Fig. 7. Plots of average retrieval rates of active scaffolds amongst the top ranking 2000 hits for each query molecule with the different virtual screening methods.

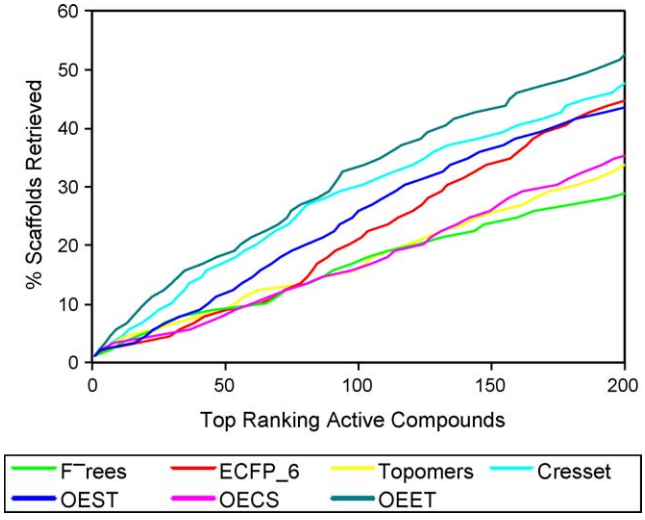


Fig. 8. Plots of average retrieval rates of active scaffolds amongst only the top ranking 200 confirmed active hits for each query molecule with the different virtual screening methods.

of active compounds was lower. We also analysed the cumulative number of ECFP6 fingerprints contained within the active molecules amongst the top ranking hits from each search, Fig. 3S. These plots show that for a given number of active compounds the Cresset and OEET methods contained more fingerprint diversity. Ultimately the OEST and OECS methods surpassed the total number of unique fingerprints identified, as they hit so many more active compounds, but they required more compounds to do this.

3.3. Prospective selection of compounds for screening

The prospective aspect of this study involved assaying compounds identified with the virtual screening methods in the CRF1 antagonist assay. New searches were performed on the Johnson & Johnson in-stock compound collection. This required building new databases in the correct format for all methods applying the same parameters as explained in Section 2.1. We used a selection of methods from the retrospective study, Topomers, Cresset, OEST and OECS approaches. The same four query molecules and conformations were used as previously described. The search results for each method were then combined. The molecules were uniqued and the ranking to each of the four query molecules was retained. Compounds with molecular weight >500, matching known in-house CRF1 scaffolds or previously tested in CRF1 assays were removed from the hits. For these experiments there was no additional filtering of the compounds prior to screening. This is not typical practice but was done to avoid introducing bias into the selections based on previous project experience. From the remaining compounds a total of 655 were chosen and requested for screening. The final selection was based on the best ranked ~160 compounds from each of the four methods.

From the CRF1 antagonist assay ten of the screened compounds showed >20% effect at 10 μM concentration, Table 3. Although this was a low hit rate of ~1.5% it was favourable compared to a CRF1 antagonist HTS performed after these experiments. The second HTS had a hit rate of 1.1% following assay at single point concentration. However, this included many compounds with MW > 500 and importantly still included known CRF1 scaffolds. Considering only those compounds of sufficient interest for the confirmation assay yielded a hit rate of 0.3%. Such a low hit rate is not unknown for CRF1 which is often perceived as a "difficult" GPCR target. The unusually high enrichments seen in the retrospective virtual screening could be seen as a forewarning for the low hit rate. We observed an astonishingly successful separation of CRF1 actives from inactives. All methods performed well compared with random. It was still of use to study the relative performance between approaches but such overall good enrichments suggest that the CRF1 dataset was a relatively easy case for the virtual screening methods. In other words actives and inactives were easily separated and one may infer that the majority of inactives were quite dissimilar to the queries. In addition the hit-rate for this prospective work may have been negatively impacted by the lack of intervention prior to recommending the compounds for screening. It has been shown elsewhere that the use of additional similarity metrics via processes such as data fusion can improve hit-rates [20].

With only a small number of active compounds from this work some caution must be used in statistically analysing in detail the relative performance of the different methods. Nevertheless, for interest we provide the ranking of the active compounds by each method in Table 3. A ranking was assigned for each compound in the order it was retrieved from the entire database for searches with each query molecule and all methods. The reported value in Table 3 is the lowest ranking to any of the four queries. The rankings were allocated prior to applying the MW, scaffold and previously screened filters described above. This had an impact on the rankings and makes them more difficult to interpret. In general the active compounds were lower ranked by the OEST and OECS methods than by Topomers.

The structures of three of the actives identified are shown in Fig. 9. Of particular interest was the difference in 2D structure between these molecules and any of the four queries. Some of these molecules show departure from the conventional pharmacophore presented in Fig. 3. They lack the combination of certain features otherwise thought to be essential for activity. For instance, 4 does not contain the small hydrophobic group alpha to a scaffold acceptor. Molecule 5 was highly ranked by the Cresset-Fieldscreen method and we show the predicted overlay with the closest query molecule, query 03, in Fig. 10. It can be seen that the new active replaces the scaffold acceptor by the carbonyl of the ester. The ester and amide groups pendant from the central thiophene have a restricted conformation due to interaction between amide NH and alkoxy oxygen of the ester. These groups and their intramolecular interaction replace the scaffold centre present in the pharmacophore. Although the pendant aryl ring does not contain the usual ortho-substitution twisted conformations were nevertheless identified. The two molecules have similar size and good overall steric overlap. The distal thiophene group provides small hydrophobic and positive charge regions which match well with similar field points from the cyclopentane of the query. Overall molecule 5 represents an interesting modification to known CRF1 compounds. However, the actives only show weak activity and it may be that they would revert to the conventional pharmacophore upon

Table 3
Actives identified from the prospective screening work.

Molecule	Ranking with each method				% Effect at 10 μM	pIC_{50}
	Topomers	OEST	OECS	Cresset		
1	466	1873	289	52	76.3	
2	236	332	648	78	70.4	5.7
3	305	1124	168	66799	63.3	5.2
4	608	3013	2588	330	62.7	5.5
5	287	518	2248	171	55.3	5.3
6	308	367	562	317	50.8	5.2
7	214	2216	342	135	29.3	
8	146	458	893	1872	26.2	
9	571	1791	1859	34243	23.9	
10	42	397	55	961	20.7	

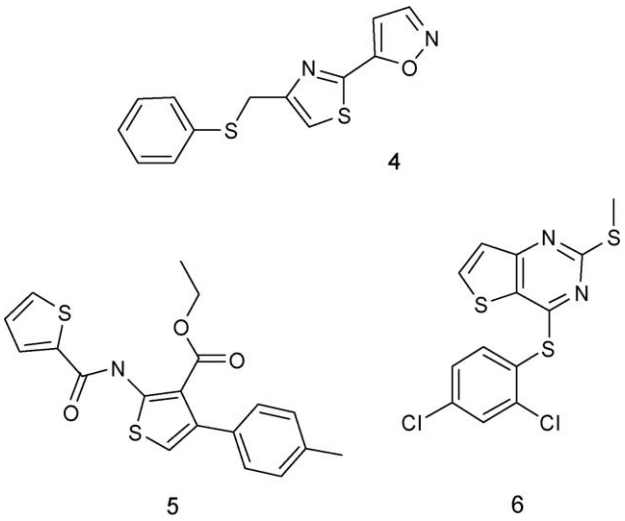


Fig. 9. A selection of the actives identified amongst the compounds recommended for screening.

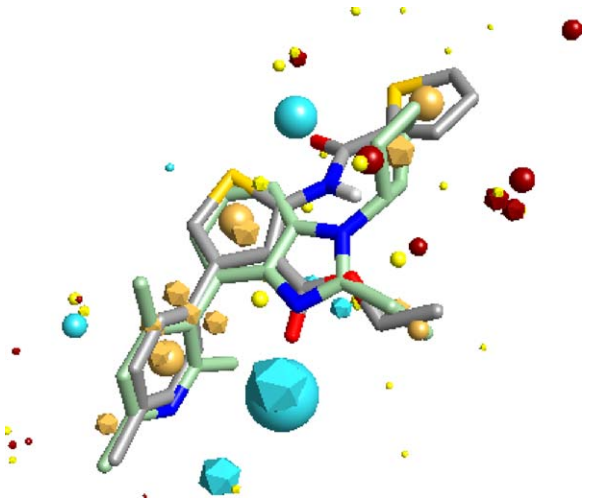


Fig. 10. Cresset-Fieldscreen overlay between query molecule 03 (green carbons) and active molecule 5 (grey carbons). Field points for query and active are shown as icosohedra and spheres respectively. Negative field is blue, positive red, steric yellow and hydrophobic orange.

further medicinal chemistry optimisation. Also, it should be pointed out that thienopyrimidine 6 has scaffold similarity to some of the examples in Fig. 2 and may be a less optimal hit for an already known series [49].

4. Conclusion

We have performed a retrospective analysis of active retrieval rates for a CRF1 antagonist dataset. The high enrichment rates compared to random indicate that all methods performed well at distinguishing the CRF1 actives from the HTS inactives. Nevertheless, OEST and OECS 3D methods performed better than 2D techniques such as ECFP6 fingerprints and FTrees. The similar performance between OEST and OECS indicated that the inclusion of features on top of 3D steric similarity offered little benefit for active–inactive differentiation. Of the simple descriptors ALogP consistently showed enrichment for all four queries and element counts produced the highest single enrichment factor. These data help to place in context the high enrichment factors from the more complex methods. In addition we also showed how the performance with query 01 was quite different to the other three queries on the same dataset. This was due to the presence of compounds in the dataset which presented more of a challenge for some methods than others. Such results highlight the need for caution in the interpretation of retrospective enrichment analyses and the value in using multiple query molecules. It was of interest that the more complex methods OEET and Cresset did not out perform OEST. These methods performed well in the diversity of the compounds they retrieved and contained more scaffolds amongst a given number of active compounds. In this regard they may be seen as better suited for scaffold-hopping but it is at the cost of active retrieval rate. Overall we have demonstrated a target and dataset example where virtual screening with 3D shape out performs other methods such as 2D fingerprints.

The prospective aspect of this study allowed a selection of the methods to recommend new compounds for screening. Ten active compounds were identified. This was a favourable hit rate compared to a subsequent HTS but well short of the high enrichment rates seen in the retrospective part of this study. Some of the identified actives were shown to be successful scaffold-hops given their different underlying 2D structures compared to the query molecules.

Acknowledgements

We are grateful to Dr. Hilde Lavreysen and Ilse van der Linden for CRF1 antagonist functional assay support. We thank Berndt Wendt from Tripos, Mark Mackey at Cresset and Jeremy Yang at Openeye.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2009.01.003.

References

[1] T. Steckler, F.M. Dautzenberg, Corticotropin-releasing factor receptor antagonists in affective disorders and drug dependence—an update, CNS Neurol. Dis. Drug Targets 5 (2006) 147–165.
[2] M.J. Owens, C.B. Nemeroff, Physiology and pharmacology of corticotropin-releasing factor, Pharmacol. Rev. 43 (1991) 425–473.
[3] C.B. Nemeroff, E. Widerlov, G. Bissette, H. Walleus, I. Karlsson, K. Eklund, C.D. Kilts, P.T. Loosen, W. Vale, Elevated concentrations of CSF corticotropin-releasing factor-like immunoreactivity in depressed patients, Science 226 (1984) 1342–1344.
[4] F. Holsboer, U. Von Bardeleben, A. Gerken, G.K. Stalla, O.A. Muller, Blunted corticotropin and normal cortisol response to human corticotropin-releasing factor in depression, N. Engl. J. Med. 311 (1984) 1127.
[5] P.C. Wynn, J.P. Harwood, K.J. Catt, G. Aguilera, Corticotropin-releasing factor (CRF) induces desensitization of the rat pituitary CRF receptor-adenylate cyclase complex, Endocrinology 122 (1988) 351–358.
[6] D.T. Chalmers, T.W. Lovenberg, D.E. Grigoriadis, D.P. Behan, E.B. De Souza, Corticotrophin-releasing factor receptors: from molecular biology to drug design, Trends Pharmacol. Sci. 17 (1996) 166–172.
[7] A.W. Zobel, T. Nickel, H.E. Kunzel, N. Ackl, A. Sonntag, M. Ising, F. Holsboer, Effects of the high-affinity corticotropin-releasing hormone receptor 1 antagonist R121919 in major depression: the first 20 patients treated, J. Psychiat. Res. 34 (2000) 171–181.
[8] T.L. Bale, W.W. Vale, CRF and CRF receptors: role in stress responsivity and other behaviors, Annu. Rev. Pharmacol. Toxicol. 44 (2004) 525–557.
[9] M.F. Mesleh, W.A. Shirley, C.E. Heise, N. Ling, R.A. Maki, R.P. Laura, NMR structural characterization of a minimal peptide antagonist bound to the extracellular domain of the corticotropin-releasing factor 1 receptor, J. Biol. Chem. 282 (2007) 6338–6346.
[10] C.W. Liaw, D.E. Grigoriadis, T.W. Lovenberg, E.B. De Souza, R.A. Maki, Localization of ligand-binding domains of human corticotropin-releasing factor receptor: a chimeric receptor approach, Mol. Endocrinol. 11 (1997) 980–985.
[11] M.E. Abreu, W. Rzeszotarski, D.J. Kyle, R.L. Elliot, Corticotropin releasing factor antagonist compounds. Patent 1991, Nova Pharmaceutical, US5063245.
[12] M. Lanier, J.P. Williams, Small molecule corticotropin-releasing factor antagonists. Expert Opin. Ther. Patents 2002, 12, 1619–1630.
[13] P.A. Keller, L. Elfick, J. Garner, J. Morgan, A. McCluskey, Corticotropin releasing hormone: therapeutic implications and medicinal chemistry developments, Bioorg. Med. Chem. 8 (2000) 1213–1223.
[14] Y.L. Chen, R.S. Mansbach, S.M. Winter, E. Brooks, J. Collins, M.L. Corman, A.R. Dunaiskis, W.S. Faraci, R.J. Gallaschun, A. Schmidt, D.W. Schulz, Synthesis and oral efficacy of a 4-(butylethylamino)pyrrolo[2,3-d]pyrimidine: a centrally active corticotropin-releasing factor 1 receptor antagonist, J. Med. Chem. 40 (1997) 1749–1754.
[15] T. Deak, K.T. Nguyen, A.L. Ehrlich, L.R. Watkins, R.L. Spencer, S.F. Maier, J. Licinio, M.L. Wong, G.P. Chrousos, E. Webster, P.W. Gold, The impact of the nonpeptide corticotropin-releasing hormone antagonist antalarmin on behavioral and endocrine responses to stress, Endocrinology 140 (1999) 79–86.
[16] C. Chen, K.M. Wilcoxen, C.Q. Huang, Y.F. Xie, J.R. McCarthy, T.R. Webb, Y.F. Zhu, J. Saunders, X.J. Liu, T.K. Chen, H. Bozigian, D.E. Grigoriadis, Design of 2,5-dimethyl-3-(6-dimethyl-4-methylpyridin-3-yl)-7-dipropylaminopyrazolo[1,5-a]pyrimidine (NBI 30775/R121919) and structure-activity relationships of a series of potent and orally active corticotropin-releasing factor receptor antagonists, J. Med. Chem. 47 (2004) 4787–4798.
[17] S. Chaki, A. Nakazato, L. Kennis, M. Nakamura, C. Mackie, M. Sugiura, P. Vinken, D. Ashton, X. Langlois, T. Steckler, Anxiolytic- and antidepressant-like profile of a new CRF1 receptor antagonist, R278995/CRA0450, Eur. J. Pharma. 485 (2004) 145–158.
[18] Y.L. Chen, J. Braselton, J. Forman, R.J. Gallaschun, R. Mansbach, A.W. Schmidt, T.F. Seeger, J.S. Sprouse, F. David, E. Winston, D.W. Schulz, Synthesis and SAR of 2-aryloxy-4-alkoxy-pyridines as potent orally active corticotropin-releasing factor 1 receptor antagonists, J. Med. Chem. 51 (2008) 1377–1384.
[19] Y.L. Chen, R.S. Obach, J. Braselton, M.L. Corman, J. Forman, J. Freeman, R.J. Gallaschun, R. Mansbach, A.W. Schmidt, J.S. Sprouse, I.I.I. Tingley, E. Winston, D.W. Schulz, 2-Aryloxy-4-alkylaminopyridines: discovery of novel corticotropin-releasing factor 1 antagonists, J. Med. Chem. 51 (2008) 1385–1392.
[20] J. Hert, P. Willett, D.J. Wilton, P. Acklin, K. Azzaoui, E. Jacoby, A. Schuffenhauer, New methods for ligand-based virtual screening: use of data fusion and machine

learning to enhance the effectiveness of similarity searching, J. Chem. Inf. Model. 46 (2006) 462–470.
[21] G. Klebe, Virtual ligand screening: strategies, perspectives and limitations, Drug Discov. Today 11 (2006) 580–594.
[22] R.D. Cramer, M.A. Poss, M.A. Hermsmeier, T.J. Caulfield, M.C. Kowala, M.T. Valentine, Prospective identification of biologically active structures by topomer shape similarity searching, J. Med. Chem. 42 (1999) 3919–3933.
[23] T.J. Hou, X.J. Xu, Recent development and application of virtual screening in drug discovery: an overview, Curr. Pharma. Des. 10 (2004) 1011–1033.
[24] F.L. Stahura, M. Bajorath, New methodologies for ligand-based virtual screening, Curr. Pharma. Des. 11 (2005) 1189–1202.
[25] A. Pozzan, Molecular descriptors and methods for ligand based virtual high throughput screening in drug discovery, Curr. Pharma. Des. 12 (2006) 2099–2110.
[26] D.J. Livingstone, The characterization of chemical structures using molecular properties. A survey, J. Chem. Inf. Comp. Sci. 40 (2000) 195–209.
[27] R.P. Sheridan, S.K. Kearsley, Why do we need so many chemical similarity search methods? Drug Discov. Today 7 (2002) 903–911.
[28] SciTegic, Accelrys Inc. 10188 Telesis Court, Suite 100, San Diego, CA 92121.
[29] M. Rarey, J.S. Dixon, Feature trees: a new molecular similarity measure based on tree matching, J. Comput.-Aided Mol. Des. 12 (1998) 471–490.
[30] R.D. Cramer, R.J. Jilek, K.M. Andrews, dbtop: toporner similarity searching of conventional structure databases, J. Mol. Graphics Model. 20 (2002) 447–462.
[31] J.A. Grant, M.A. Gallardo, B.T. Pickup, A fast method of molecular shape comparison: a simple application of a Gaussian description of molecular shape, J. Comp. Chem. 17 (1996) 1653–1666.
[32] T.S. Rush, J.A. Grant, L. Mosyak, A. Nicholls, A shape-based 3-D scaffold hopping method and its application to a bacterial protein–protein interaction, J. Med. Chem. 48 (2005) 1489–1495.
[33] A. Nicholls, N.E. MacCuish, J.D. MacCuish, Variable selection and model validation of 2D and 3D molecular descriptors, J. Comput.-Aided Mol. Des. 18 (2004) 451–474.
[34] J.G. Vinter, Extended electron distributions applied to the molecular mechanics of some intermolecular interactions, J. Comp.-Aided Mol. Des. 8 (1994) 653–668.
[35] T. Cheeseright, M. Mackey, S. Rose, A. Vinter, Molecular field extrema as descriptors of biological activity: definition and validation, J. Chem. Inf. Model. 46 (2006) 665–676.
[36] J.G. Vinter, K.I. Trollope, Multiconformational composite molecular-potential fields in the analysis of drug-action. 1. Methodology and first evaluation using 5-HT and histamine action as examples, J. Comput.-Aided Mol. Des. 9 (1995) 297–307.
[37] G. Schneider, W. Neidhart, T. Giller, G. Schmid, "Scaffold-hopping" by topological pharmacophore search: a contribution to virtual screening, Angew. Chem.-Int. Ed. 38 (1999) 2894–2896.
[38] R.D. Cramer, R.J. Jilek, S. Guessregen, S.J. Clark, B. Wendt, R.D. Clark, "Lead hopping". Validation of topomer similarity as a superior predictor of similar biological activities, J. Med. Chem. 47 (2004) 6777–6791.
[39] A. Nicholls, J.A. Grant, Molecular shape and electrostatics in the encoding of relevant chemical information, J. Comput.-Aided Mol. Des. 19 (2005) 661–686.
[40] E.J. Barker, D. Buttar, D.A. Cosgrove, E.J. Gardiner, P. Kitts, P. Willett, V.J. Gillet, Scaffold hopping using clique detection applied to reduced graphs, J. Chem. Inf. Model. 46 (2006) 503–511.
[41] Molecular Operating Environment, version 2006.06, 2006, Chemical Computing Group Inc., Montreal, Quebec, Canada.
[42] P.A. Keller, M. Bowman, K.H. Dang, J. Garner, S.P. Leach, R. Smith, A. McCluskey, Pharmacophore development for corticotropin-releasing hormone: new insights into inhibitor activity, J. Med. Chem. 42 (1999) 2351–2357.
[43] A. McCluskey, P.A. Keller, J. Morgan, J. Garner, Synthesis, molecular modeling and biological activity of methyl and thiomethyl substituted pyrimidines as corticotropin releasing hormone type 1 antagonists, Org. Biomol. Chem. 1 (2003) 3353–3361.
[44] R.J. Jilek, R.D. Cramer, Topomers: a validated protocol for their self-consistent generation, J. Chem. Inf. Comp. Sci. 44 (2004) 1221–1227.
[45] CONCORD was developed by R.S. Pearlman, A. Rusinko, J.M. Skell, R. Baducci, at the University of Texas, Austin, and distributed by Tripos, Inc., 1699 S. Hanley Rd., St. Louis, MO 63144.
[46] R.D. Cramer, R.D. Clark, D.E. Patterson, A.M. Ferguson, Bioisosterism as a molecular diversity descriptor: steric fields of single "topomeric" conformers, J. Med. Chem. 39 (1996) 3060–3069.
[47] A. Bender, R.C. Glen, Discussion of measures of enrichment in virtual screening: comparing the information content of descriptors with increasing levels of sophistication, J. Chem. Inf. Model. 45 (2005) 1369–1375.
[48] G.B. McGaughey, R.P. Sheridan, C.I. Bayly, J.C. Culberson, C. Kreatsoulas, S. Lindsley, V. Maiorov, J.F. Truchon, W.D. Cornell, Comparison of topological, shape, and docking methods in virtual screening, J. Chem. Inf. Model. 47 (2007) 1504–1519.
[49] A. Nakazato, T. Okubo, D. Nozawa, T. Tamita, L.E.J. Kennis, Thienopyrimidine and thienopyridine derivatives substituted with cyclic amino group. Patent 2005, Taisho Pharmaceutical Co. Ltd., WO2005066182.