Molecular Field Extrema as Descriptors of Biological Activity: Definition and Validation

Tim Cheeseright, Mark Mackey, Sally Rose, and Andy Vinter*

Cresset Biomolecular Discovery Ltd., Spirella Building, Letchworth, Hertfordshire SG6 1ET, United Kingdom

Received August 31, 2005

The paper describes the generation of four types of three-dimensional molecular field descriptors or 'field points' as extrema of electrostatic, steric, and hydrophobic fields. These field points are used to define the properties necessary for a molecule to bind in a characteristic way into a specified active site. The hypothesis is that compounds showing a similar field point pattern are likely to bind at the same target site regardless of structure. The methodology to test this idea is illustrated using HIV NNRTI and thrombin ligands and validated across seven other targets. From the in silico comparisons of field point overlays, the experimentally observed binding poses of these ligands in their respective sites can be reproduced from pairwise comparisons.

INTRODUCTION

Overview. The basis of cheminformatics and virtual drug discovery in almost all their manifestations is structure comparison. Structure is here defined as the framework created by linking atoms with bonds in two or three dimensions. However, it is well-known that molecules of different structural classes can act at the same biological site (Figure 1). It is also accepted that molecules interact via their electronic properties: electrostatic and van der Waals forces. It therefore follows that if two molecules with diverse structures interact with an enzyme or receptor in a similar way, they will have similar surface properties in their bound conformations.

This paper describes an in silico method of defining molecular fields in a form that enables semiquantitative comparisons across molecules in three dimensions. Our aim is to show that these molecular fields can be used as nonstructural templates for defining behavioral similarity. The approach described in this paper has already been successfully used in the drug discovery arena.[1−3] Numerous previous efforts by others to define what is 'seen' by another molecule have failed for two major reasons: the inadequacy of the model used to define charge distribution and the unmanageable amount of data needed to describe surface molecular properties.

First, the definition of electrostatic potential over the molecular surface depends on the atomic charge distribution within the molecule. Adequately accurate spatial charge distributions can only be derived from quantum mechanics (QM) using full wave functions and will change with molecular conformation and environment. However, it is unreasonable at this time to expect QM to be fast enough to cope with even the accessible conformations of most natural hormones and therapeutics.[4] The problem of speed can be overcome using molecular mechanics (MM), but the most commonly used charge model places single point partial charges at atom centers (atom-centered charges or ACCs) and has proved to be too approximate to define surface electronic properties to a useful degree of accuracy. For example, the molecular electrostatic potential above the carbonyl group of acetone shows no sign of splitting to show two lone pairs when derived from atom-centered charges. However, molecular electrostatic potentials can be determined reliably using molecular mechanics with a more sophisticated charge model. The eXtended Electron Distribution (XED) force field[5−7] redefines the charge on electronegative and π atoms (atoms having p orbital valence involvement) in a molecule away from the conventional ACC monopole toward multipole electron distribution around the atom more in keeping with quantum orbital descriptors.[8,18] Over the past 10 years, the XED force field has been refined to include intramolecular interactions for carbon, oxygen, nitrogen, sulfur, phosphorus, and the halogens and has been validated against some of the commonly used MM force fields.[9,10]

Second, potential fields around a molecule are continuous functions and are hard to compare across molecules. Approximations based on grids or surfaces[11−14] produce too many data points for fast processing or are restricted in their accuracy by the grid resolution and lack of gauge invariance. Describing fields in terms of Gaussians[15] is elegant and quicker than grids but works best when describing fields which can be approximated by overlapping spheres (such as molecular volumes) and is less appropriate for probe-interaction energy fields.

What is needed is a fast way of generating molecular potential fields with sufficient spatial and quantitative accuracy to be suitable for quick estimates of similarity between them. In this paper, we revisit the approach that has utilized molecular field extrema[16−19] as a means of overcoming grid restrictions and large numbers of descriptors. We will describe the field generation and overlay procedures, how field patterns describe the nonbonding chemistry of ligands, and the results of simple overlay experiments.

Field Points. Proteins respond to the potential field around a molecule rather than to the 3D arrangement of its individual

* Corresponding author phone: +44-(0)1462-476320; e-mail: j.g.vinter@cresset-bmd.com.

10.1021/ci050357s CCC: $33.50
Published on Web 01/11/2006

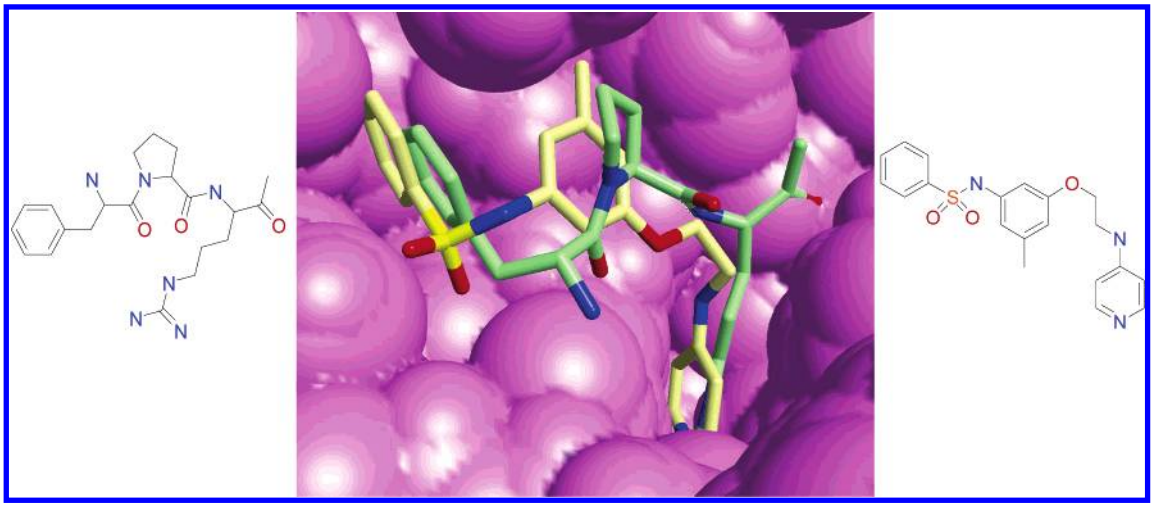

Figure 1. Thrombin inhibitors (X-ray data, 1ppb and 1uvt, from the Protein Data Bank) serve to exemplify two different chemotypes binding at the same site and giving the same biological effect.

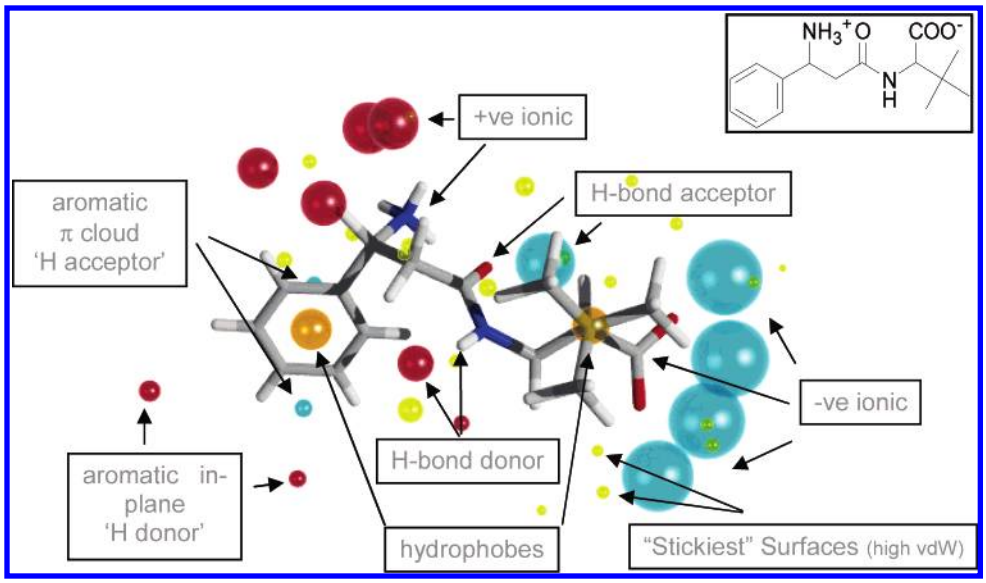

Figure 2. The molecule above (inset in 2D) serves to illustrate the relationship between some common chemical groups and the typical field point patterns associated with those groups. Large field points are generated by charged groups such as ammonium and carboxylate ions. The H-bond donor/receptor arrangements are represented by the amide linker, and the mixed hydrophobic and electrostatic character of a phenyl group is reflected in a combination of in-plane positive field points, π-cloud points above and below the plane of the ring, and the hydrophobic point at its center. Four field types are defined: electrophilic (red), nucleophilic (blue), van der Waals attractive (yellow), and hydrophobic (orange); see Methods section. The size of the 'balls' reflects the depths of each extremum energy well.

atoms. If we want to describe how a molecule appears to a protein, we need to define it as a set of properties near the molecular surface, not as a collection of atoms and bonds. A good description of these steric and electrostatic properties is vital if we are to understand biological activity via the interaction of two molecules.

A full surface description of a molecule over all of its accessible conformations is too complex to handle. We solve this problem by condensing the complex three-dimensional electrostatic/van der Waals fields down to their local extrema or 'field points'. Figure 2 exemplifies how some common functional groups relate to their associated field points although it is important to keep in mind that the generation of each field point takes into account effects from the whole molecule (vide infra).

Two molecules with different structures but similar biological activities present similar potential fields to their common binding site. As a result, they are expected to have similar sets of field points. This means that field patterns can be used to align molecules, to score active molecules, and to search through databases of compounds looking for potential hits. As the pattern is not directly related to the 2D connectivity of the molecule, but rather to its 3D properties, fields can be used to compare molecules from completely different structural classes.[20]

METHODS

(1) Field Point Generation. The fields that we will be describing are scalar fields which are derived in general from calculating the interaction energy of a 'probe' molecule with the target molecule. This has advantages over simply using the raw electrostatic field values: the electrostatic field is only sampled at points which are accessible to another molecule, and the field values are interaction energy scores that can be related directly to the energetics of molecular interactions.

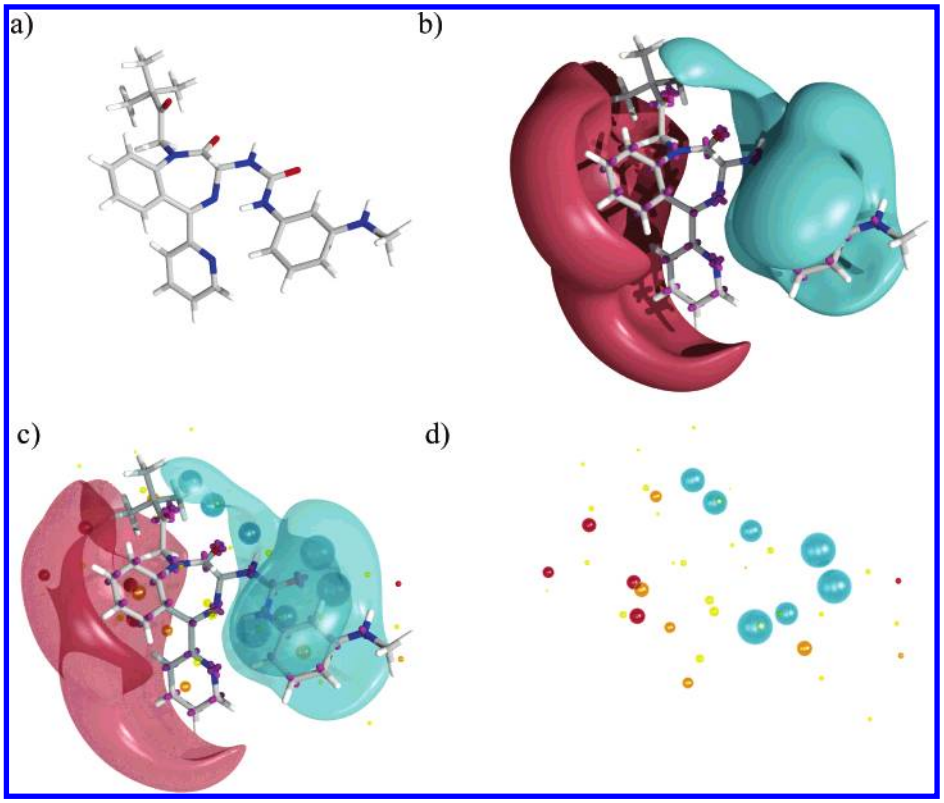

Figure 3. Steps in the creation of field points. (a) Structure of a candidate drug in its active conformation. (b) The molecular electrostatic potential map at the 0.5 kcal/mol contour showing the XED constructs on the electronegative and π atoms as mauve dummy atoms. (c) Distilled field point extrema for positive, negative, surface, and hydrophobic points superimposed onto (b). (d) The final field point pattern for the conformation of the drug shown in (a). All four field types are included: electrophilic (red), nucleophilic (blue), van der Waals attractive (yellow), and hydrophobic (orange); see Methods section. The size of the 'balls' reflects the depths of each extremum energy well.

Figure 3 introduces the methodology employed to create field points around a molecular conformer, showing the XED formalism necessary to create acceptable electrostatic fields, the distilled electrostatic field points superimposed on the potential contours, and the final field point pattern that is used as the basis of comparison with other molecules. The final pattern includes two extra field point types in addition to the positive and negative electrostatic points. These additions reflect the surface and hydrophobic character of the molecule.

One hundred twenty points are generated at regular intervals over a slightly diminished solvent accessible surface of each atom of the molecule. A probe atom is placed at each point and its interaction energy with the molecule optimized with a simplex that finds a true extremum and avoids grid techniques. The probe is given the van der Waals parameters of oxygen, and its charge is adjusted according to the potential to be used. The 120 points on each atom optimize down to common extrema. Extrema with very small values of the interaction energy are insignificant and are filtered out.

The first potential is a Morse description of the van der Waals interaction using a neutral probe (eq i).

$$V_{\mathrm{vdW}} = K_{\mathrm{v}} \bullet E_{\mathrm{vdW}}$$

$$E_{\mathrm{vdW}} = \sum_{1}^{j} \{-E_{\mathrm{pj}}(z^2 - 2z)\} \text{ kcal/mol}$$

$$z = \exp\{-b_{\mathrm{pj}}(1 - [r_{\mathrm{pj}}/r^0_{\mathrm{pj}}])\} \quad \text{(i)}$$

where V_{vdW} is the van der Waals energy of the neutral probe with the molecule containing j atoms. K_{v} is a constant, and E_{pj} and b_{pj} are parameters from the XED force field. r is the distance between the atom pair, and r^0 is the sum of the vdW radii

The second and third potentials calculate the Coulombic interaction for a negative probe and a positive probe according to eq ii

$$V_{\mathrm{c}} = \sum_{1}^{i} q_{\mathrm{p}} q_i / 4\pi e_0 \bullet 1/r_{\mathrm{ip}} \bullet 1/D \bullet 332.17 + E_{\mathrm{vdW}} \text{ kcal/mol} \quad \text{(ii)}$$

where V_{c} = the energy between the charged probe q_{p} ($\pm$1.0e) and a molecule with i charges (XED charges) at distances r_{ip} in a dielectric medium of $D = 4$. Note that the Coulombic potentials include the van der Waals potential.

The fourth potential (eq iii) calculates an attractive energy with a neutral probe. This potential reflects the hydrophobicity of a fragment or group. It is zero weighted on electronegative atoms relative to carbon to reflect low hydrophobicity and is weighted to 0.5 on hydrogens to reduce their importance without eliminating their effect

$$V_{\mathrm{h}} = K_{\mathrm{h}} \bullet - \sum_{1}^{j} \{9 \bullet E_{\mathrm{pj}}/(25 + r_{\mathrm{pj}}^3)\} \text{ kcal/mol} \quad \text{(iii)}$$

where V_{h} is the hydrophobic energy of a neutral probe p with the molecule containing j atoms and K_{h} is a constant.

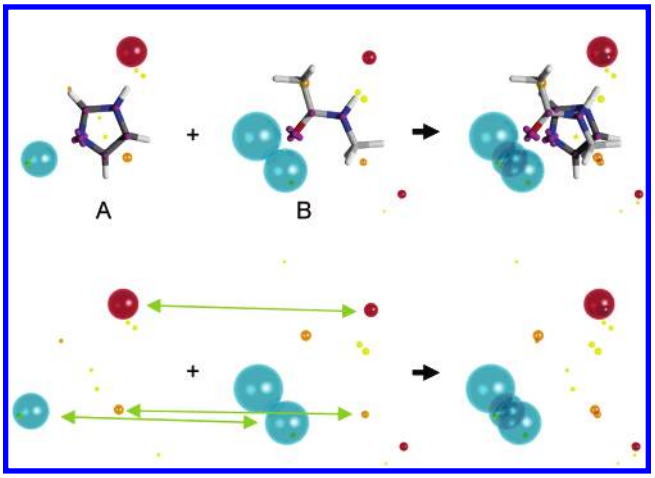

Figure 4. Simple field overlay scheme (see text). The upper picture shows the structural consequence of overlaying the field points of molecule A onto B (lower picture). Overlay is achieved using field information only and involves no structural information.

K_v and K_h are scaling factors set to 2.0 and 30.0, respectively.[21]

The two types of extremum derived from eq ii reflect centers of electrophilicity (displayed as 'red points') and electrophobicity (displayed as 'blue points'). Those calculated purely from eq i suggest where the 'stickiest' points occur on the molecular surface (displayed as 'yellow points'). The field extrema from eqs i and ii tend to occur on or near the molecular surface.

In contrast, the so-called 'hydrophobic' extrema (displayed as 'orange points') from eq iii penetrate the molecular surface and furnish us with a general measure of structural bulk in nonelectrostatic regions of the molecule. For example, an adamantyl group would have one 'hydrophobic' point at its center. A cyclohexyl would have a smaller one comparable with that for phenyl. However, only the phenyl group generates electrostatic points in addition to its 'hydrophobic' points in accordance with the general chemical intuition that benzene is hydrophobic and displays electrostatic properties,[22,23] while the saturated hydrocarbons have little or no electrical influence.

(2) Field Point Comparisons. Having devised a way to define the essential properties of a molecule in terms of a tractable number of field points (approximately equal to the number of heavy atoms), the aim of the project is to use these points to compare structurally diverse molecules that are known to behave in biologically, and possibly chemically, similar ways. This implies that structural features are no longer of consequence in the comparison stage, that structure is merely the underlying generator of a 'molecular field', and that only the molecular 'fields' are important in molecular recognition.

Let us assume that we wish to compare molecule A (imidazole) and molecule B (*N*-methylacetamide) in Figure 4 using their field points alone. A simple metric would involve assigning a pseudo-Coulombic potential between pairs of field points

$$S_{\mathrm{fpBfpA}} = \frac{\mathrm{size}(\mathrm{fp_A}) \times \mathrm{size}(\mathrm{fp_B})}{\mathrm{dist} + \mathrm{offset}} \qquad \text{(iv)}$$

where negative field points on molecule A attract negative field points on molecule B but repel positive ones, while surface field points and hydrophobic field points simply attract points of like type and ignore others. Summation of the scores over all field point pairs leads to an overall score for any given alignment of two field patterns. Maximizing this score by moving one field pattern relative to the other (using a simplex optimizer, for example) then gives a field-based molecular alignment method. Normalizing the score provides a field-based similarity metric.

This metric has some problems. Although a field point has a quantitative size associated with it (determined from the depth of the potential energy well it represents), the above procedure takes no account of the shape and width of the energy well associated with a field point. Neither does it investigate the environment surrounding the energy well where the various potentials have finite values but have no associated extrema. Figure 5 uses pyridine field patterns to illustrate the problem.

To overcome these hurdles, an improved field overlay metric was developed. Rather than simply base the score on the relative positions of the field points, the field points on each molecule are used as sampling points into the actual field potentials of the other molecule

$$E_{\mathrm{A}\rightarrow\mathrm{B}} = \sum_{\mathrm{fp_A}} \sqrt{\mathrm{size}(\mathrm{fp_A}) \times F_{\mathrm{B}}(\mathrm{position}(\mathrm{fp_A}))} \qquad \text{(v)}$$

where $F_B(x)$ is the value of the appropriate field on B at position x, and the sum is over all field points on A.

This score is asymmetric, so we repeat for the field points of B sampling into the actual field of A and average the two to give a symmetric score. The score can then be normalized to give a Dice field similarity metric

$$S_{\mathrm{AB}} = \frac{2E_{\mathrm{AB}}}{E_{\mathrm{AA}} + E_{\mathrm{BB}}} \qquad \text{(vi)}$$

and maximizing this metric between two conformations of A and B gives both the best conformational overlay (in terms of field similarity) and a single field similarity value for the two conformations.

Because the energies are analytically recalculated, the entire 'true' field is used in the calculation, and the potential well widths are implicitly included. However, only a few field values need to be calculated in any given orientation so the technique is fast enough to be applied to large structures and many conformations in reasonable computing time. The fields of each molecule are sampled at only a few places, but the use of the field extrema of the other molecule as the sampling points ensures that the fields are sampled at biologically relevant points. It is also worth noting that this calculation is gauge-invariant[24] and hence avoids many of the issues involved in grid-based similarity metrics.

A further problem arises when a portion of the potential surface is close to flat such that small changes in geometry can lead to either one extremum or two being presented (Figure 6). This can be somewhat alleviated by scaling the field point sizes before field overlay. A scaling function is applied which reduces each field point's size if any other field points of the same type are near by: in the limit of two field points overlapping, each is halved in size. Now two closely spaced field points will give the same overlay

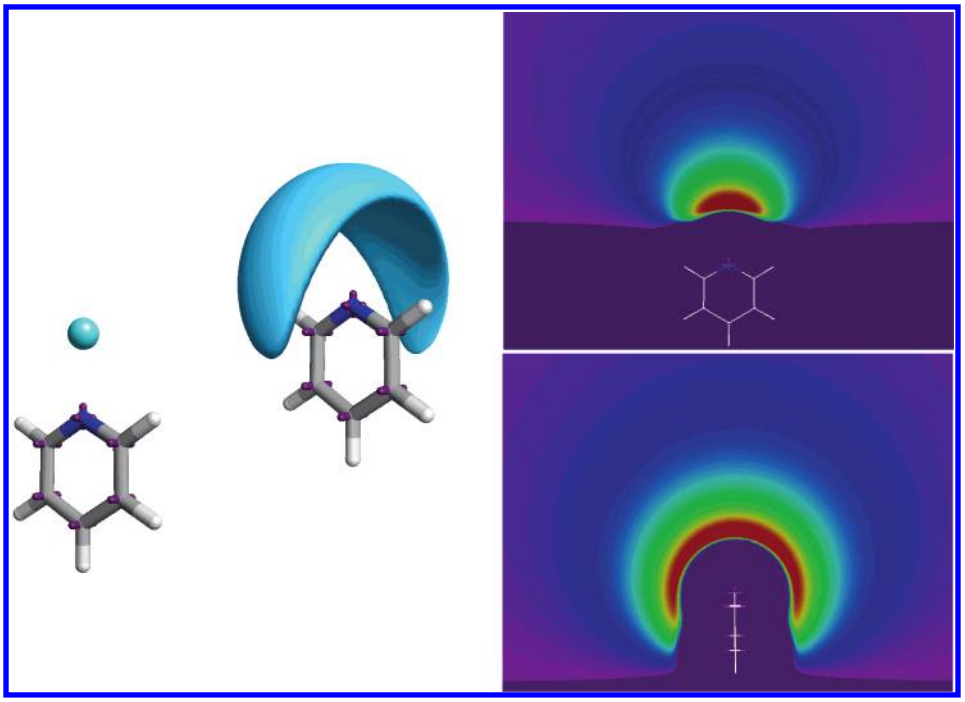

Figure 5. On the far left, the negative field point extremum of pyridine (blue) loses information about its shape and extent of influence as reflected by a plot of the −0.5 kcal/mol contour to its right. The full electrostatic potential plots are shown on the far right in the two orthogonal planes of the ring (the inner red contour corresponding to the blue areas on the left picture). This problem is overcome by recalculating the potentials 'on the fly' (see text).

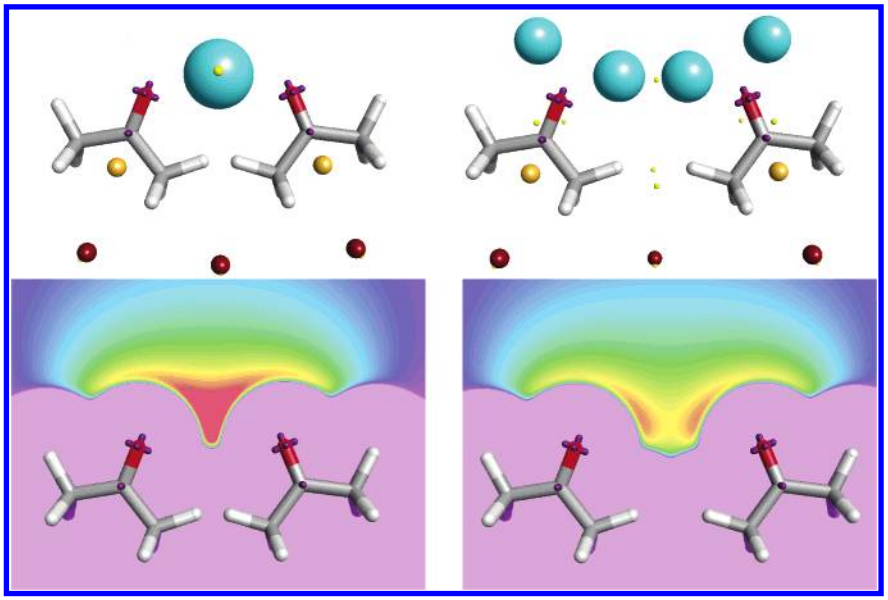

Figure 6. The effect of the proximity of two atoms in space on the field point pattern is illustrated using acetone fragments as convenient examples. The upper picture shows the pattern of field point extrema after full optimization when two carbonyls are close (left) and separated (right). The lower pictures record the full electrostatic potential on a plane through the carbonyl groups (the inner red contour corresponding to the blue field points on the upper picture). When comparing the field points of two molecules whose field patterns are perturbed by close proximity, scaling field point sizes before overlaying the fields can alleviate the inconsistency (see text).

energy as one point, thus removing this ambiguity when the final overlay score is calculated.

(3) The Field Point Overlay Technique. As mentioned previously, given two rigid conformations A and B, maximizing the field similarity S_{AB} over all possible relative orientations provides a field-based 'best alignment" of A and B and concomitant field similarity value. The problem is to determine this best orientation. The surface of S_{AB} with respect to translation and rotation of B is generally quite complex, with numerous local maxima, so this is a global optimization problem. Given an appropriate starting alignment of A and B, a simplex optimizer can be used to maximize S_{AB} so the problem becomes one of finding suitable starting alignments such that the global maximum (and high-scoring local maxima) can be located with a high degree of confidence.

To prepare reasonable starting orientations for a simplex optimization of each field point pattern, we use a variant on a colored clique matching algorithm. The field patterns are seen as graphs with nodes colored by field type and edges labeled with the distances between pairs of field points. A search is performed for colored cliques across the two graphs,

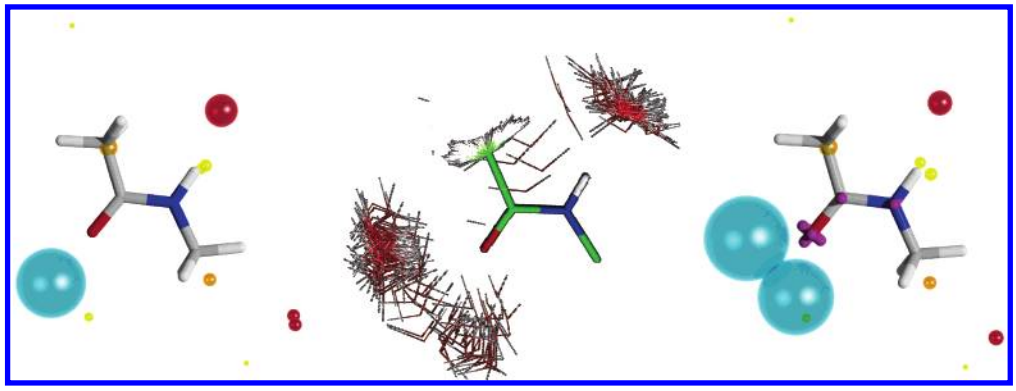

Figure 7. The field points of *N*-methylacetamide showing the loss of lone pair definition when atom-centered charges are used (far left). The Isostar plot of waters binding to *N*-methylacetamide moieties (middle) clearly indicates the presence of the lone pairs. The XED force field copes with this problem (far right). See ref 18 for more examples.

with each clique being scored by the number and size of the field points matched, reduced by a penalty for having mismatched distances in the edges. The search tree is pruned by discarding cliques whose distance mismatches are too large. The best-scoring collection of cliques found (generally 50−100 cliques) is used to generate a set of initial maps of the field points of A and B: the molecules are aligned according to least-squares fitting of the mapped field points and then submitted to the simplex optimizer. In testing this procedure against a Monte Carlo method which simply started the simplex at several thousand random orientations, this clique-matching technique found the best alignment in almost all cases and was significantly faster.

RESULTS AND DISCUSSION

(1) Field Point Validation. The use of a distributed multipole model as implemented in the XED force field is of paramount importance if field calculations are to yield useful results. Field point patterns generated from quantum mechanics[18] correlate well with our molecular mechanics patterns using the XED force field. Isostar[25] provides good experimental validation data for checking if field point patterns are qualitatively correct. For example, Figure 7 shows a typical example of the difference in field pattern when extended electron distribution is included compared with when atom-centered charges are used. The lone pairs on the carbonyl of the example molecule are lost in the latter case when experimental evidence, in the form of an Isostar plot of water molecules interacting with *N*-methylacetamide moieties, plainly indicates their presence.

(2) Overlay Process and Validation Results. To illustrate the process and validity of the field point overlay procedure named 'FieldCompare', two examples are reported in detail, both of which are rooted in therapeutic case histories. In section A, using three HIV NNRTI inhibitors in their active conformations extracted from X-ray data,[26] we show that molecular field overlays can generate the experimentally observed orientation of each inhibitor in the protein. This amounts to a feasibility study using the single, active conformation of neutral ligands. In section B, the procedures described in section A are applied to a larger data set and compared with literature results to test their general applicability. Finally, although section B contains some formally charged ligands, we investigate in depth whether formally charged ligands are handled properly. It goes without saying that formal charge gives rise to gross changes in field pattern, and the unpredictable effects of solvation and pH on the extent of formal charge must be resolvable. In addition, if we are to use the FieldCompare tool without the knowledge of the active conformation, can we still detect the experimental binding arrangement of a set of ligands from a comparison of the field points of all their conformations? Finally, will irrelevant overlay patterns mask the detection of the relevant ones? A small set of positively charged thrombin inhibitors has been chosen to test these points in section C.

(A) Neutral HIV NNRTI Inhibitors in their Active Conformations. The active conformations of three HIV NNRTI inhibitors were extracted from X-ray data. Each structure was atom typed according to the XED force field rules. Hydrogens and XEDs were added and minimized, and a field pattern was calculated for each ligand (Figure 8).

The question to be asked is as follows: do the field point overlays of each pair of compounds correspond to the structural overlays from the X-ray data?

Table 1 catalogues the best 10 field overlays for each pair of the three chosen HIV ligands. The correspondence of each overlaid pair to the experimental X-ray data is shown in the last column (X-RMS) as an rms deviation of heavy atoms. For all pairs (duos), correlation of the experimental X-ray overlay with both the field similarity ($S_{AB,}$ eq vi) and the raw field overlay energy score for the first 5 overlays is good. The raw score is less useful for other targets (see thrombin later), reflecting the need for a normalized field similarity metric.

In the ordering of the HIV duo (pairwise) overlays for the HIV ligands (Table 1), it is encouraging that the field similarity (S_{AB}) correlates well with the goodness of fit to the X-ray data (X-RMS). The first five duos of the EFZ+NVP overlay are correct in overall orientation but vary slightly in the rotational aspects of each ligand. Beyond duo number 5, one of the ligands has turned through 180 degrees and resulted in a marked deterioration of the rms fit to experiment. For the EFZ+AAP overlay, the first 6 duos correspond to the X-ray data with small rotational variations. Overlay 7 and 9 again reflect a reversal of one of the ligands. For the NVP+AAP duo, the top 3 overlays correlate well with X-ray data.

We estimate errors to all overlay raw energy score values after simplex optimization to be ±0.5. It should be empha-

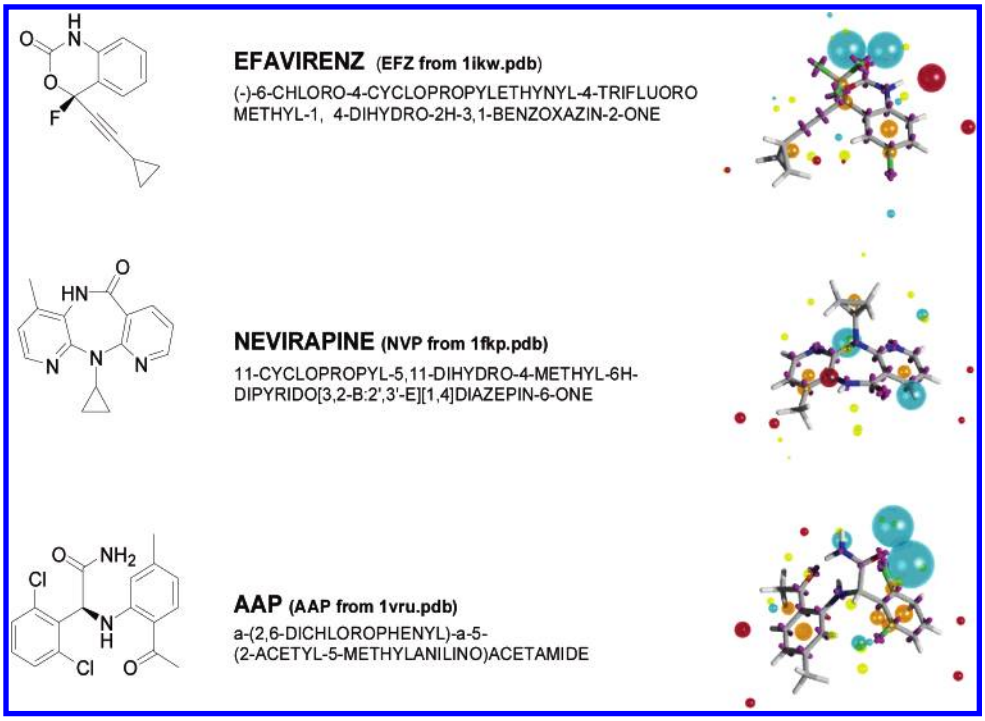

Figure 8. Details of the three HIV NNRTI inhibitors used to investigate the field overlay principle. The abbreviations for the ligands along with the Protein Data Bank reference codes from which they were derived are in brackets (1ikw resolution 3.0 Å, 1fkp resolution 2.9 Å, 1vru resolution 2.2 Å). The relevant XED constructs and field extrema patterns are shown on the right of the figure for each active conformer.

sized that the accuracy of the experimental overlay pattern is limited by the X-ray crystal preparation conditions, resolution, and interpretation of the X-ray electron density and the possible variations in the way the protein/ligand complexes are overlaid to retrieve the experimental ligand overlays.[26] The rmsd values are obtained using a least squares heavy atom fit of each overlaid pair with each pair of X-ray structures. In other words, the rmsd is an average across all atoms of each pair of molecules. Experience has strongly suggested to us that any field overlay between 0.0 and 1.25 rmsd may be 'correct'. Indeed, if there is movement of a ligand within its protein active site, all results between 0.0 and 1.25 rmsd may be right.[27] This 'acceptability cutoff' is in agreement with that of Poso et al. (see below).[14]

So far, we have shown that in silico field pattern overlays are successful in deducing the experimental binding orientation of three ligands acting at the same protein binding site given the active conformations of the ligands as starting structures. It is now necessary to validate these illustrative observations on a larger data set.

(B) Validating the Field Point Overlay Technique with Larger Data Sets Derived from X-ray Conformations. The formal charge state of any ligand drastically changes field patterns, and it is important to show that field generation and overlay techniques are able to handle charged molecules. For a ligand at or around its target binding site, the charges expected on ionizable groups are not easy to ascertain. Experience with field point patterns for many chargeable species has led us to apply a general rule that field comparison is most reliable when a formal charge is set at between one-eighth and one-quarter of the full formal charge; for example, a charged primary amine at a formal charge of +1 will be seen by the field generation software as a formal charge of +0.125 when the divisor is set to 8. Charged arginine spreads the 1/8th charge over three nitrogens, the two oxygen atoms on the carboxylate anion each carry a formal charge of −0.0625, and so on. If more emphasis on charge has been required, no more than 0.25 of a full formal charge has been used (divisor = 4) if field distortions were to be avoided. The divisor is acting as a formal charge dielectric that varies between 4 and 8, a value possibly reflecting what would be expected at a binding site in 'protein' phase.[28] Only in the gas phase would the dielectric be unity, extending exponentially to 80 only in the full aqueous phase. The HIV NNRTI inhibitors in section A are representative of uncharged ligands. The following experiments extend the principles described in section A to a larger data set and use a mixture of charged and uncharged ligands. In section C we deal exclusively with charged thrombin ligands.

We decided to compare our results with those published by Poso et al.[14] where possible. This group has developed a system called BRUTUS along the same conceptual lines as FieldCompare but using a different approach. They construct a grid of electrostatic (based on atom-centered charges) and van der Waals points for rigid-body molecular superposition and similarity searching.

We ran their data set through our FieldCompare protocol. In each case the bound conformation of a single ligand was compared against a range of other ligands for the same target in their bound conformations. The data consisted of 35 HIV protease inhibitors (HIVp, target ligand from 1qbr), 13 rhinovirus coating inhibitors (HRV, target from 1ruc), 11 elastase inhibitors (elastase, target 1ele), 7 thermolysin inhibitors (therm, target 5tmn), and 5 matrix metalloproteinase 8 inhibitors (MMP8, target 1mmb). To these, we added

Table 1. First 10 Field Overlays of Each Pair of the Three HIV NNRTI Inhibitors Defined in Figure 1 in Their Active Conformations (from X-ray Data)[a]

EFZ+NVP overlay number	raw score	S_{AB}	X-RMS (Å)
1	−37.5	0.539	0.830
2	−37.2	0.535	0.721
3	−36.8	0.529	0.759
4	−36.7	0.527	0.367
5	−35.6	0.512	0.726
6	−34.6	0.498	3.575
7	−33.6	0.483	2.699
8	−33.2	0.478	3.478
9	−32.0	0.460	2.610
10	−27.0	0.388	3.619
EFZ+AAP overlay number	**raw score**	S_{AB}	**X-RMS (Å)**
1	−44.4	0.567	0.924
2	−42.5	0.544	1.118
3	−40.7	0.521	0.893
4	−40.3	0.515	1.004
5	−39.6	0.506	0.890
6	−39.5	0.505	0.516
7	−37.2	0.476	3.262
8	−34.1	0.436	1.013
9	−34.1	0.436	3.191
10	−33.4	0.428	1.208
NVP+AAP overlay number	**raw score**	S_{AB}	**X-RMS (Å)**
1	−44.5	0.581	0.535
2	−44.4	0.580	0.667
3	−42.9	0.561	1.206
4	−41.8	0.546	2.846
5	−41.3	0.540	0.661
6	−40.8	0.533	2.877
7	−39.2	0.513	0.926
8	−39.2	0.513	3.058
9	−38.0	0.496	1.265
10	−36.8	0.481	3.022

[a] 'Raw score' is the field overlay score, 'S_{AB}' is the field similarity of overlay (eq vi), and X-RMS is the rms deviation of heavy atoms of the overlaid pair from the corresponding pair derived from the experimental X-ray data. The results are ordered by S_{AB}.

11 thrombin inhibitors (thrombin, target 1ets) and 8 HIV-nonnucleotide reverse transcriptase inhibitors (HIVnnrti, target 1ikw) not processed by Poso et al. As mentioned above, our 'acceptability cutoff' is 1.25 Å rmsd. Their cutoff is twice this value (2.5 Å) because their fitting proceeds only over one molecule in each pair after fully superimposing the other. The two criteria are thus entirely equivalent.

Table 2 summarizes the comparison of results from the two approaches using our cutoff criterion of 1.25 rmsd. In each example, Poso et al. chose the best result from the top five overlays output by BRUTUS. We have done likewise with the output from FieldCompare, and the results in the 'B5' and 'FC5' columns of Table 2 record the comparable results from the two methods. We have added a column marked 'FC1' in Table 2 that records how many of our top answers correspond to the correct overlay. Finally, the results in Table 2 have been subdivided not only to cover those overlays that are within the 1.25 Å rmsd cutoff but also to reflect resolution at the 0.5 Å and 0.25 Å rmsd levels.

As can be seen, the comparison of results from each group is most encouraging and reinforces the concepts of molecular field overlays as good descriptors of binding poses. Failure

Table 2. Summary of Results from Applying Molecular Field Pairwise Overlay Protocols to 7 Sets of Inhibitors[a]

		rmsd ≤ 0.25			rmsd ≤ 0.5			rmsd ≤ 1.25		
target	*n*	B5	FC5	FC1	B5	FC5	FC1	B5	FC5	FC1
HIVp (1qbr)	35	3	13	11	17	23	19	32	32	22
HRV (1ruc)	13	2	3	2	7	7	6	11	12	10
elastase (1ele)	11	0	3	3	4	4	4	4	7	4
Therm (5tmn)	7	2	1	1	5	4	1	7	7	6
MMP8 (1mmb)	5	0	0	0	1	0	0	5	5	3
thrombin (1ets)	12	-	1	0	-	9	6	-	12	11
HIVnnrti (1ikw)	8	-	1	0	-	3	1	-	8	6

[a] Each column records the number of inhibitors falling within the specified rmsd range of the experimental X-ray data. Two protocols are reported for the first 5 targets: BRUTUS[14] and FieldCompare (vide infra). *n* = number of inhibitors, B5 = BRUTUS top 5 overlay results, FC5 = FieldCompare top 5 overlay results, FC1 = FieldCompare top overlay result only.

to find the correct overlay in most cases can be explained by inhibitors addressing different active sites (Figure 9a) or where there are significant field differences between the two ligands (Figure 9b). If the technology is to be extended to ligands for which no X-ray data are available, we want to be confident that the top answer in a list of overlays is the true pose. This is a tall order if we insist on working in ligand space only, and there is no reason to suppose that a 'top' answer will comply with the constraints imposed by the target protein. However, our added column (FC1) suggests that this may be possible with further clever filters.

(C) Thrombin Inhibitors as Sets of Conformations. An important question in this introduction to our field overlay technology is whether duos corresponding to experimental X-ray data can be found from a series of conformations derived from a conformational search engine. More specifically, if three ligands are introduced to the procedure with no knowledge of their active conformations, can common overlay patterns be extracted that contain the active conformers of each ligand and reflect the observed binding patterns of each ligand?

Field patterns are sensitive to conformation, substitution pattern, charge state, solvation, and other influences. Consequently, if two diverse structures are known to act at the same site, only their 'active' conformers are expected to yield matching field point patterns. This idea leads to a possible way to derive the active conformers of two or more ligands without any knowledge of the target site. The first step would be to derive a reasonable conformation collection for each ligand and second, to add a field pattern to each conformer in the collection. The third step would involve the pairwise overlay comparison of the field of every conformer of the first molecular field with every one on the second. The analysis of these data would be expected to show that the best matches, in terms of the similarity metric S_{AB} (eq vi), should resemble the active conformers and reflect at least some of the important features that are required by the target active site.

To test whether active conformers can be inferred from conformational sets, three charged thrombin inhibitors were chosen for study (Figure 10).

The intention of this paper is to show how active conformations can be described in terms of structureless field patterns which can be overlaid to simulate experimental

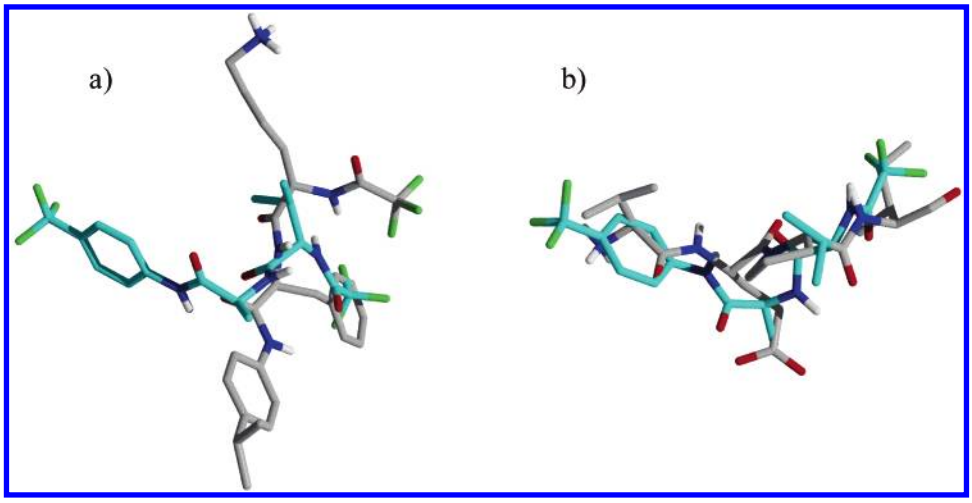

Figure 9. The "correct" overlay of the elastase ligands from (a) 1ELE and 1ELC and (b) 1ELE and 1H9L. In (a) the volume of overlap is low (volume Tanimoto similarity 0.45), with the two ligands addressing significantly different parts of the elastase active site. In (b) the ligand from 1H9L has significant field points from the charged carboxylate that are unmatched in 1ELE. The carboxylate is largely solvated and therefore should be downgraded as an important binding moiety. However, our field alignment procedure treats all field points as equally important for binding and hence creates 'wrong' overlays that use these carboxylate field points.

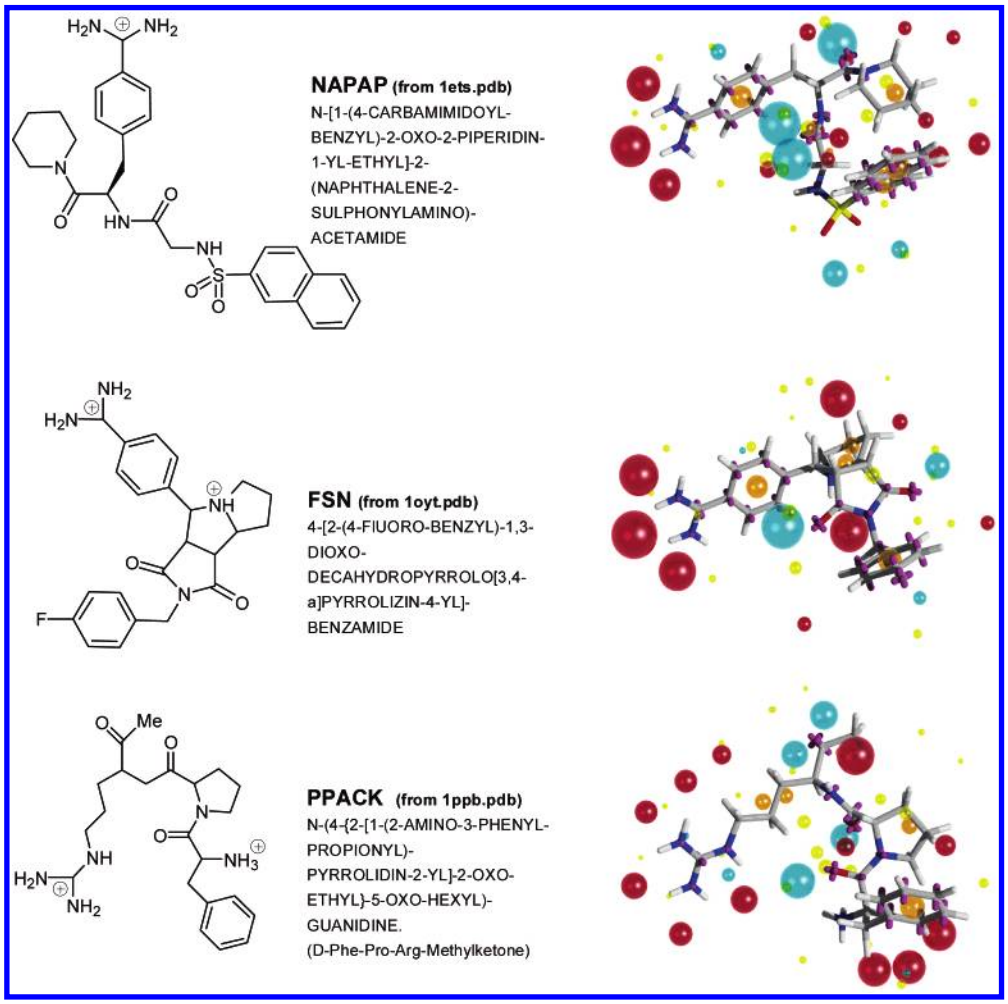

Figure 10. Details of the three thrombin inhibitors used to investigate the field overlay principle for charged molecules. The abbreviations for the ligands along with the Protein Data Bank reference codes from which they were derived are in brackets (1ets resolution 2.3 Å, 1oyt resolution 1.67 Å, 1ppb resolution 1.92 Å). Each formally charged center is set to 1/8th unit charge (see Discussion). The relevant XED constructs and field extrema patterns are shown on the right of the figure for each active conformer.

binding orientations. The question of how to find active conformations, starting with 2D input structures (sdf files), is irrelevant in this context and will be dealt with elsewhere. It will therefore be assumed that a conformation search engine has been able to find a close relative of the active conformer of each of the ligands of interest somewhere in its output conformer list. The active conformers for the HIV NNRTI overlays in section A were taken directly from the X-ray data. In this section, conformation searches were performed for all three thrombin ligands starting from a 2D sdf file, and the resulting conformer list was checked to confirm that it contained the active conformer or a close relative (within 0.5 rmsd). Note that, because of the large number of conformations expected from PPACK, its conformation search was controlled with a single constraint that forced the arginine group to remain in an extended state. This reduced the conformer count to 29 from several hundred.

The field point patterns for all conformers from each charged ligand were overlaid to create duos. Table 3 summarizes the top 10 results from all pairwise overlays of 88 NAPAP conformers, 2 FSN conformers, and 29 PPACK conformers.

The 'X-RMS' column has been included in Table 3 for easy reference later but should be ignored while we ask the following question: what indications have we for inferring which conformers might be the active ones? For the two duos involving NAPAP, only conformation 56 is common. For the PPACK duos, conformers 10, 3, and 26 are common. Conformer 10 turns up in two of the high-scoring duos. FSN is limited to 2 conformers only. We are concerned only with showing that the active conformers of a set of known ligands binding at the same site show high field similarity with each other. However, it may be noted that the use of conformationally restricted ligands is of considerable value in determining bound conformers under this regime and should be taken advantage of where possible.

Further examination of Table 2 reveals that conformer 56 of NAPAP is associated with conformer 2 of FSN, conformer 2 of FSN is associated with conformation 10 of PPACK, and conformer 10 of PPACK is associated with conformer 56 of NAPAP creating the only complete cyclic path across the three sets of 10 overlays.

We can conclude that the field overlay procedure has probably found the active conformation for each ligand; NAPAP conformer 56, FSN conformer 2, and PPACK conformer 10.

We can now check our conclusion by referring to the relevant X-ray data:

1. Do the proposed active conformers correspond to the experimental X-ray data?

The usual heavy atom least-squares fitter was used to check that the surmised active conformers corresponded to the active conformers derived from X-ray data: NAPAP conformer 56; rmsd 0.22, FSN conformer 2; rmsd 0.14, and PPACK conformer 10; rmsd 0.59.

2. Do the derived duos correspond to the experimental X-ray overlay poses?

Column 6 of Table 3 records the correspondence of each duo with the X-ray overlays of each pair. Conformer 56 of NAPAP with conformer 2 of FSN is the top field overlay answer and corresponds to the best X-ray overlay. Conformer 56 of NAPAP with conformer 10 of PPACK is the fourth best field overlay answer and corresponds to the best X-ray overlay. Conformer 10 of PPACK with conformer 2 of FSN is the third best field overlay answer and corresponds to the best X-ray overlay.

Table 3. Best 10 Field Overlay Duos (Ordered by S_{AB} (Eq vi)) of the Three Thrombin Inhibitors Defined in Figure 9[a]

NAPAP+FSN overlay number	NAPAP conf no.	FSN conf no.	raw score	S_{AB}	X-RMS
1	56	2	−72.3	0.606	0.366
2	23	1	−68.4	0.558	3.147
3	48	1	−64.8	0.528	3.500
4	67	2	−61.8	0.518	4.108
5	49	1	−60.6	0.493	2.276
6	67	1	−60.5	0.493	3.569
7	48	2	−58.3	0.489	2.878
8	20	2	−57.9	0.486	4.339
9	18	2	−56.0	0.469	4.067
10	24	2	−54.6	0.458	3.980

NAPAP+PPACK overlay number	NAPAP conf no.	PPACK conf no.	raw score	S_{AB}	X-RMS
1	42	1	−63.0	0.569	4.545
2	41	25	−76.6	0.567	2.988
3	10	15	−70.5	0.543	5.063
4	56	10	−72.7	0.534	0.542
5	31	12	−64.6	0.532	2.957
6	22	21	−60.7	0.532	2.887
7	50	13	−67.9	0.521	4.551
8	10	6	−62.0	0.482	4.996
9	52	3	−67.7	0.469	3.298
10	9	26	−61.5	0.451	5.174

FSN+PPACK overlay number	FSN conf no.	PPACK conf no.	raw score	S_{AB}	X-RMS
1	1	7	−56.8	0.483	3.917
2	1	19	−65.4	0.479	3.808
3	2	10	−59.4	0.436	0.752
4	1	24	−59.7	0.425	3.540
5	1	26	−56.4	0.414	4.003
6	1	3	−59.3	0.410	3.495
7	2	29	−55.5	0.405	1.874
8	1	10	−52.6	0.386	1.913
9	1	17	−54.2	0.385	4.037
10	1	16	−57.8	0.383	3.249

[a] From a conformation search on each molecule over a 6 kcal/mol energy window and a filter set to eliminate duplicates at ≤0.5 rms, 88 conformers were found for NAPAP, 2 conformers were found for FSN, and 29 conformers were found for PPACK (constrained search − see Discussion). 'Conf no.' is the number of the conformers involved in the overlay, 'raw score' is the field overlay energy, 'S_{AB}' is the field similarity of overlay (eq vi), and X-RMS is the rms deviation of heavy atoms of the overlaid pair from the corresponding pair derived from the experimental X-ray data.

We can now conclude that the field overlay procedure has found the active conformation for each ligand in this particular case.

Finally, can a 'trio' be derived from the three proposed active duos that corresponds to the experimental X-ray data?

If the three sets of 'active' duos in Table 3 are superimposed to form a 'trio' of actives (i.e. conformer 56 of NAPAP, conformer 2 of FSN, and conformer 10 of PPACK with duplicates removed), the correspondence to the X-ray trio overlay deviates by rms 0.8 (Figure 11) and therefore corresponds to the X-ray overlay by our 'acceptable fit' criterion.

The processing of the three chosen thrombin ligands in the above experiments has given rise to just one definitive

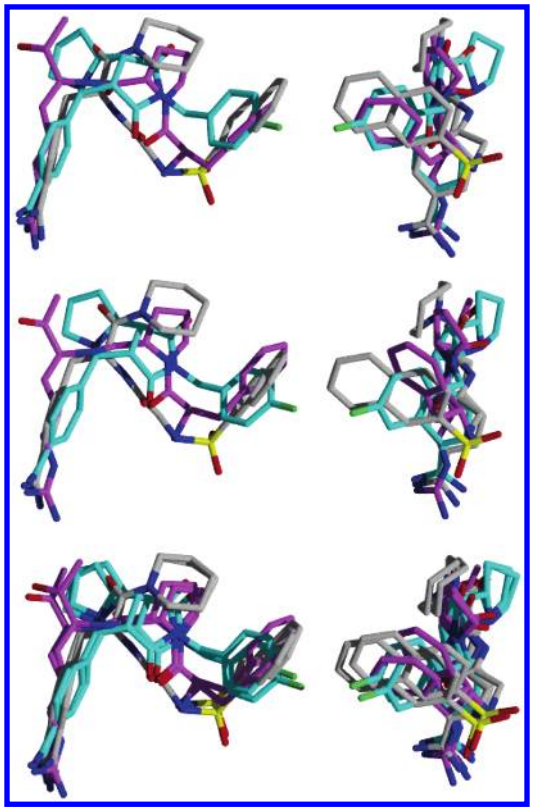

Figure 11. The overlays of the thrombin inhibitors NAPAP, FSN, and PPACK. Top: The experimentally observed overlay derived from the X-ray data of the ligands in their protein (Figure 10). Middle: The 'trio' overlay (with duplicates removed). Bottom: The heavy atom least-squares fit of the X-ray overlay (top) and trio overlay (middle) representing a rmsd of 0.77.

answer for the set of active conformations (Figure 11). We have deliberately used one inhibitor (FSN) with a small set of conformers to simplify the explanation of the procedures. However, most searches for duos that we have carried out, using multiple conformation sets for different ligands and different targets, usually result in several feasible duos. This again begs the question of which duo is the 'true' representation of the binding pattern, and the arguments put forward in section B on this question still apply. In the next paper, we will be reporting more fully on the criteria useful for distilling the most likely candidates and discussing 'field templating'—the creation of 'trios' and higher order active ligand clusters—in more detail using a wide variety of data sets.

CONCLUSION

In this paper, we report on the techniques necessary to create a minimal set of meaningful field points on ligands known to act at the same biological site and a way in which they can be overlaid to reproduce experimental observations without using any protein structural information. The principles of our approach depend on a definition of the electrostatic and hydrophobic fields around a molecular surface and deliberately move away from the restrictions associated with purely structural comparisons. The groups of Poso[14] and Maggiora[29] have covered many of the problems and pitfalls associated with this approach, most of which vanish with the use of the XED force field (giving better electrostatic potentials), the generation of field extrema (avoiding grid constructs and reducing the amount of data), and the technique of recalculating potentials 'on the fly' (taking account of the complete potential characteristics and imparting gauge invariance).

One of the chief advantages of using molecular fields to investigate biological activity is their avoidance of structural prejudice. We chose three HIV NNRTI inhibitors because of their relative conformational simplicity and lack of charge. In contrast, our final example used three charged thrombin inhibitors, two of which have considerable conformational flexibility. Furthermore, peptide ligands are often found to be good therapeutic agents but are not acceptable as drug candidates. We have shown that the well-known peptidic thrombin inhibitor, PPACK, can be successfully overlaid by field comparison with two nonpeptidic derivatives thus opening up the opportunity to jump from peptides to nonpeptides while retaining the desired biological activity. This process is currently undertaken in the pharmaceutical industry as an empirical exercise costing much time and resource.

We have successfully shown that molecular field patterns can be used to deduce the molecular binding requirements for several protein targets using diverse chemotypes known to bind in the active site of each target. For each target active site, the structurally different ligands give rise to similar field patterns that reflect what the binding site expects. Using this information, it is possible to derive the active conformer of each ligand without any protein structural information.

ACKNOWLEDGMENT

The authors would like to thank Caroline Low, Sir James Black, and the staff at the James Black Foundation for their scientific input and support over the years.

Supporting Information Available: An expansion of Table 2 giving detailed results of the overlays of the various sets of inhibitor and a pdb formatted file of the structures shown in Figure 11. This information is available free of charge via the Internet at http://pubs.acs.org/.

REFERENCES AND NOTES

(1) Kalindjian, S. B.; Buck, I. M.; Davies, J. M.; Dunstone, D. J.; Hudson, M. L.; Low, C. M.; McDonald, I. M.; Pether, M. J.; Steel, K. I.; Tozer, M. J.; Vinter, J. G. Non-peptide cholecystokinin-B/gastrin receptor antagonists based on bicyclic, heteroaromatic skeletons. *J. Med. Chem.* **1996**, *39*, 1806−1815.

(2) McDonald, I. M.; Dunstone, D. J.; Kalindjian, S. B.; Linney, I. D.; Low, C. M.; Pether, M. J.; Steel, K. I.; Tozer, M. J.; Vinter, J. G. 2,7-Dioxo-2,3,4,5,6,7-hexahydro-1H-benzo[*h*][1,4]diazonine as a new template for the design of CCK(2) receptor antagonists. *J. Med. Chem.* **2000**, *43*, 3518−3529.

(3) Low, C. M. R.; Buck, I. M.; Cooke, T.; Cushnir, J. R.; Kalindjian, S. B.; Kotecha, A.; Pether, M. J.; Shankley, N. P.; Vinter, J. G.; Wright, L. Scaffold Hopping with Molecular Field Points: identification of a CCK2 receptor pharmacophore and its use in the design of a prototypical series of pyrrole- and imidazole-based CCK2 antagonists. *J. Med. Chem.* **2005**, *48*, 6790−6802.

(4) Klähn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. On possible pitfalls in ab initio quantum mechanics/molecular mechanics minimisation approaches for studies of enzymatic reactions. *J. Phys. Chem.* B **2005**, *109*, 15645−15650.

(5) Hunter, C. A.; Sanders, J. K. M. The nature of $\pi-\pi$ interactions. *J. Am. Chem. Soc.* **1990**, *112*, 5525−5534.

(6) Vinter, J. G. Extended electron distributions applied to the molecular mechanics of intermolecular interactions. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 653−668.

(7) Vinter, J. G. Extended electron distributions applied to the molecular mechanics of intermolecular interactions. II−Organic Complexes. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 417−426.

(8) Stone, A. J.; Price, S. L. Some new ideas in the theory of intermolecular forces: anisotropic atom−atom potentials. *J. Phys. Chem*. **1988**, *92*, 3325−3335.
(9) Lozman, O. R.; Bushby, R. J.; Vinter, J. G. Complementary polytopic interactions (CPI) as revealed by molecular modelling using the XED force field. *J. Chem. Soc., Perkin* **2001**, *2*, 1446−1453.
(10) Chessari, G.; Hunter, C. A.; Low, C. M. R.; Packer, M. J.; Vinter, J. G.; Zonta, C. An Evaluation of Force Field Treatments of Aromatic Interactions. *Chem. Eur. J*. **2002**, *8*, No. 13, 2860−2867.
(11) Goodford, P. J. A Computational Procedure for Determining Energetically Favourable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem*. **1985**, *28*, 849−857.
(12) Cramer, R., D., III.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA) 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc*. **1988**, *110*, 5959−5967.
(13) Bender, A.; Mussa, H. Y.; Gill, G. S.; Glen, R. C. Molecular Surface Point Environments for Virtual Screening and the Elucidation of Binding Patterns (MOLPRINT 3D). *J. Med. Chem.* **2004**, *47*, 6569−6583.
(14) Tervo, A. J.; Ronkko, T.; Nyronen, T. H.; Poso, A. BRUTUS: Optimization of a Grid-Based Similarity Function for Rigid-Body Molecular Superposition. 1. Alignment and Virtual Screening Applications. *J. Med. Chem*. **2005**, *48*, 4076−4086.
(15) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *J. Med. Chem*. **1994**, *37*, 4130−4146.
(16) Davis, A.; Warrington, B.; Vinter, J. G. 'Strategies in Drug Design II− Modelling studies on phosphodiesterase substrates and inhibitors'. *J. Comput.-Aided Mol. Des.* **1987**, *1*, 97−120.
(17) Vinter, J. G.; Saunders, M. R. Molecular Modelling approaches to host−guest complexes. In *Host−Guest Molecular Interactions: from Chemistry to Biology*; Wiley: Ciba Foundation Symposium 1991; No 158, pp 249−261.
(18) Apaya, R. P.; Lucchese, B.; Price, S. L.; Vinter, J. G. The matching of electrostatic extrema: A useful method in drug design? A study of phosphodiesterase III inhibitors. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 33−43.
(19) Vinter, J. G.; Trollope, K. I. Multi-conformational Composite Molecular Fields in the Analysis of Drug Design. (1) Methodology and First Evaluation using 5HT and Histamine Action as examples. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 297−307.
(20) Cheeseright, T.; Mackey, M.; Vinter, J. G. Peptides to non-peptides: leads from structureless virtual screening. *DDT BIOSILICO* **2004**, *2*, 57−60.
(21) The constant K_v scales the van der Waals potential points around the molecular surface. It is estimated by assuming that the sum of all the yellow field points derived from eq i gives the overall energy of transferring a molecule from polar phase to a nonpolar binding site (vdW_{sum}). The equivalent energy (vdW_{surf}) can also be crudely derived from multiplying the nonpolar area of a molecule by 0.06 kcal·mol^{-1} Å^{-1} (See Sharman, G. J.; Searle, M. S.; Benhamu, B.; Groves, P.; Williams, D. H. Burial of Hydrocarbon Causes Cooperative Enhancement of Electrostatic Binding. *Angew. Chem., Int. Ed. Eng.* **1995**, *34*, 1483−1485.). Taking 1894 molecules ranging from 3 to 86 heavy atoms, K_v was adjusted to give the closest correspondence between vdW_{sum} and vdW_{surf}. K_h can be estimated using the same set of molecules by assuming that the sum of positive plus negative electrostatic field points tends on average to equal the sum of vdW plus hydrophobic points.
(22) Suzuki, S.; Green, P. G.; Bumgarner, R. E.; Dasgupta, S.; Goddard, W. A.; Blake, G. A. Benzene Forms Hydrogen Bonds with Water. *Science* **1992**, *257*, 942−945.
(23) Dougherty, D. A. Cation-pi Interactions in Chemistry and Biology: A New View of Benzene, Phe, Tyr and Trp. *Science* **1996**, *271*, 163−168.
(24) If fields are calculated on a grid constructed around a molecule, their values will change on rotation of the molecule subject to (a) the resolution of the grid and (b) the degree of rotation. This gauge variance must be avoided if molecules are to be optimally overlaid by field comparison. Using a simplex optimization regime for both the generation of field points and the overlaying of field point patterns ensures grid avoidance and gauge invariance.
(25) Bruno, I. J.; Cole, J. C.; Lommerse, J. P. M.; Rowland, R. S.; Taylor, R.; Verdonk, M. L. Isostar: A library of information about nonbonded interactions. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 525−537.
(26) When extracting experimental ligand overlays from protein X-ray data, our practice is to overlay all relevant proteins (with their ligands) by least-squares fit on all alpha carbon atoms along all homologous sequences. The procedure is performed graphically, enabling visual judgment to be made of the goodness of fit both of the proteins and ligands. The rare occasions when overlays larger than 0.8 rms were encountered, they were easily recognized as unacceptable and discarded. Having overlaid all relevant proteins, their ligands were isolated by deleting all protein structures and saving the ligands without further coordinate change. The ligand clusters so isolated were subsequently used to calculate the root-mean-square deviation (X-RMS) of our field overlay patterns.
(27) Das, K.; Clark, A. D., Jr.; Lewi, P. J.; Heeres, J.; de Jonge, M. R.; Koymans, L. M. H.; Vinkers, H. M.; Daeyaert, F.; Ludovici, D. W.; Kukla, M. J.; De Corte, B.; Kavash, R. W.; Ho, C. Y.; Ye, H.; Lichtenstein, M. A.; Andries, K.; Pauwels, R.; de Be'thune, M.-P.; Boyer, P. L.; Clark, P.; Hughes, S. H.; Janssen, P. A. J.; Arnold, E. Roles of Conformational and Positional Adaptability in Structure-Based Design of TMC125−R165335 (Etravirine) and Related Non-nucleoside Reverse Transcriptase Inhibitors That Are Highly Potent and Effective against Wild-Type and Drug-Resistant HIV-1 Variants. *J. Med. Chem.* **2004**, *47*, 2550−2560.
(28) (a) Lockhart D. J.; Kim, P. S. Internal stark effect measurement of the electric field at the amino terminus of an alpha helix. *Science* **1992**, *257*, 947−951. (b) Lockhart D. J.; Kim, P. S. Electrostatic screening of charge and dipole interactions with the helix backbone. *Science* **1993**, *260*, 198−202.
(29) (a) Mestres, J.; Rohrer, D. C.; Maggiora, G. M. A molecular-field-based similarity study of nonnucleoside HIV-1 reverse transcriptase inhibitors. *J. Comput.-Aided Mol. Des*. **1999**, *13*, 79−93. (b) Mestres, J.; Rohrer, D. C.; Maggiora, G. M. A molecular-field-based similarity study of nonnucleoside HIV-1 reverse transcriptase inhibitors. 2. The relationship between alignment solutions obtained from conformationally rigid and flexible matching. *J. Comput.-Aided Mol. Des*. **2000**, *14*, 39−51.

CI050357S