Development and Validation of AMANDA, a New Algorithm for Selecting Highly Relevant Regions in Molecular Interaction Fields

Ángel Durán, Guillermo C. Martínez,[§] and Manuel Pastor*

Research Unit on Biomedical Informatics (GRIB), IMIM/Universitat Pompeu Fabra, Avinguda Dr. Aiguader 88, E-08003 Barcelona, Spain

Descriptors based on Molecular Interaction Fields (MIF) are highly suitable for drug discovery, but their size (thousands of variables) often limits their application in practice. Here we describe a simple and fast computational method that extracts from a MIF a handful of highly informative points (hot spots) which summarize the most relevant information. The method was specifically developed for drug discovery, is fast, and does not require human supervision, being suitable for its application on very large series of compounds. The quality of the results has been tested by running the method on the ligand structure of a large number of ligand−receptor complexes and then comparing the position of the selected hot spots with actual atoms of the receptor. As an additional test, the hot spots obtained with the novel method were used to obtain GRIND-like molecular descriptors which were compared with the original GRIND. In both cases the results show that the novel method is highly suitable for describing ligand−receptor interactions and compares favorably with other state-of-the-art methods.

INTRODUCTION

Molecular Interaction Fields (MIF) are useful tools for characterizing the ability of a molecule to interact with other molecules. In their most simple formulation (the Molecular Electrostatic Potentials), they represent the energy of interaction of a molecule with a positive charge located at any x, y, z coordinate of the molecule neighborhood. This basic concept can be expanded by replacing the positive charge by a more complex chemical probe representing any kind of functional group. The so obtained MIF can be represented graphically, highlighting areas of the space where molecules holding a probelike group can produce energetically favorable interactions (Figure 1).

In the pioneer work of Goodford[1] the MIF were recognized as a useful tool for identifying regions on the structure of biological receptors where the ligands can establish intense interactions. Since then, the MIF have been applied for characterizing both the structures of ligands and of receptors in different fields of computational chemistry and with different aims, including many applications in drug discovery which have been recently reviewed.[2]

MIF are intrinsically continuous functions which can be described using analytic expressions, but more often they are described by sampling them at regular intervals over the space surrounding the molecules at certain positions or "grid nodes". This method has the disadvantage of producing a large number of data points, in the order of thousands or hundreds of thousands for regular size molecules. However, not all regions around the molecules are equally interesting. Usually our interest is focused only in certain regions of this space holding the most negative (favorable) energy values,

which we will call "hot spots" here. If the MIF is used to describe the binding site of a biological receptor, these hot spots represent promising locations where a ligand can place a functional group similar to the probe. Conversely, when the MIF is used to characterize ligand molecules, these hot spots represent groups of the receptor binding site with which the molecule could establish favorable binding interactions. In either case, hot spots represent privileged regions holding highly relevant information for describing the ability of two molecules to interact.

Hot spots can be easily identified by simple visual inspection of a MIF graphical representation (like the one shown in Figure 1) as distinct areas holding the most negative values. In practice, methods involving human recognition are not convenient, and for most applications we need a computational method able to extract the hot spots automatically. The benefits of using hot spots in computational chemistry are obvious, since they can summarize in a handful of points the most relevant information contained in a MIF formed by hundreds of thousands of nodes. For example, the use of hot spots would allow the use of MIF-derived molecular descriptors in drug design methods involving large series of compounds (like virtual screening or library design), where they have been traditionally avoided due to computational performance reasons. The potential applications of hot spots are not limited to their use as molecular descriptors, and they could be applied in molecular superposition algorithms, docking simulations, etc. Unfortunately, the development of a computational algorithm able to extract these regions is not simple. Questions like the criteria used to recognize the MIF regions, the total number of points to extract, or the value of the energy cutoff that discriminates intense interactions from weak interactions have no unique answers and are probably required to be addressed differently depending on the field of application. In this sense, there

* Corresponding author e-mail: manuel.pastor@upf.edu.
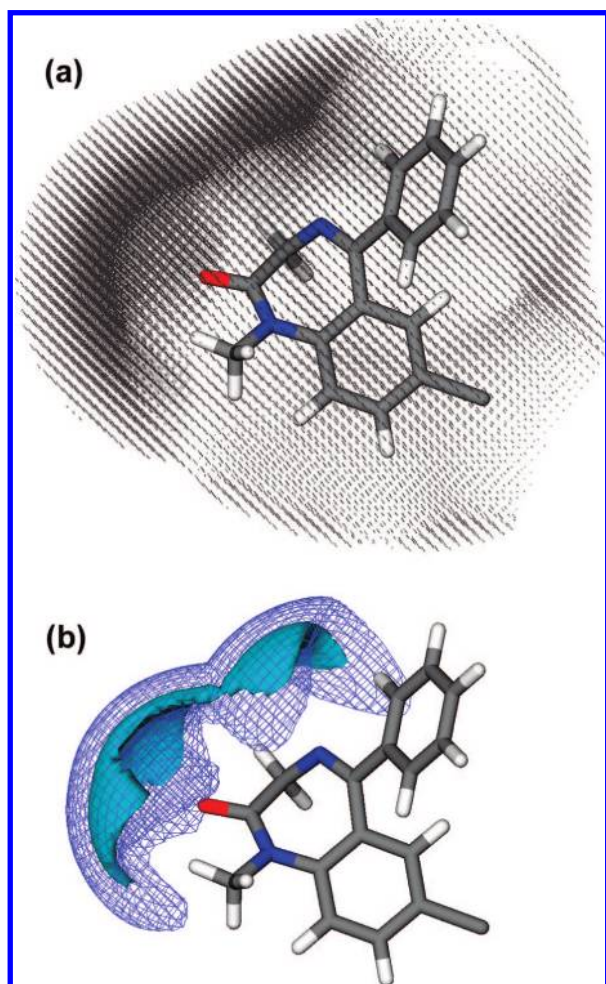§ Current address: Bioinformatics Unit, CNIO, C/Melchor Martínez Almagro 3, E-28029 Madrid, Spain.

Figure 1. MIF computed using a water probe on a molecule of diazepam. (a) Energies under 0.5 Kcal/mol, represented as crosses of a size proportional to the absolute value of the computed energy. (b) Isovolumes enclosing regions for which the energy of interaction is lower than −2.5 Kcal/mol (contour) and −5.0 Kcal/mol (solid).

are many ways to extract hot spots, each one adequate for a different practical purpose. The present work will be focused on the application of hot spots for small molecules design where these hot spots must be, above all, representative of the ligand ability to establish noncovalent bonding interactions with a biological receptor.

The problem of extracting hot spots from MIF has been addressed by diverse authors who have proposed solutions for MIF obtained with proteins[3,4] as well as on MIF obtained with small molecules.[5,6] A method for selecting hot spots starting from GRID[1] MIF was also proposed by one of us, as part of the algorithm for obtaining the alignment-independent descriptors GRIND.[7] In our opinion, none of the methods proposed so far is fully satisfactory, and all of them have different limitations and drawbacks that hamper their general application in the field of drug design. In the present work we will describe AMANDA, a simple but efficient algorithm which can be used to extract from any MIF a set of hot spots highly suitable for being applied in the field of small molecule design. The quality of the algorithm has been validated by comparing the agreement between the hot spots obtained and the actual receptor atom positions for a set of nearly 800 ligand−receptor complexes extracted from the PDB crystallographic database. A detailed

Table 1. Prefilter Threshold Values Used for Diverse GRID Probes

	AMANDA	GRIND
DRY	−0.5	−0.35
N1	−2.6	−0.61
O	−4.2	−0.72

comparative analysis with other hot spots methods will also be presented. To conclude, the suitability of the novel method for drug design application was additionally tested by using the resulting hot spots for computing GRIND-like molecular descriptors and building 3D QSAR models, the quality of which compares favorably with those obtained with the original GRIND.

METHODS

The AMANDA Algorithm. The method requires that the starting MIF nodes were tagged by the atom contributing most to the field energy. In the program GRID[8] this feature is already present, and such MIF can be easily obtained using the ALMD directive. A node prefiltering is first carried out simply by applying a suitable energy cutoff, in order to discriminate between relevant nodes and nodes representing weak or nonspecific interactions, which are removed. The cutoff values used in AMANDA (Table 1) were selected by carrying out energy values distribution analysis for large collections of MIF obtained on druglike compounds (data not shown). Then a list of the remaining m_i MIF nodes assigned to every atom (i) in the molecule was built. Provided that this list is not empty, a maximum of n_i nodes will be selected for atom i, where n_i is computed using eq 1.

$$n_i = e^{[a.\ln(m_i)]} \tag{1}$$

The purpose of this equation is to obtain a nonlinear fraction of the total number of nodes, containing a few nodes for small regions and some more nodes for large regions, but to avoid extracting too many nodes for some atoms and too few nodes (or none at all) for some others. Figure 2 shows the value of n obtained for different values of m and of the weighting parameter a. According to our tests, the
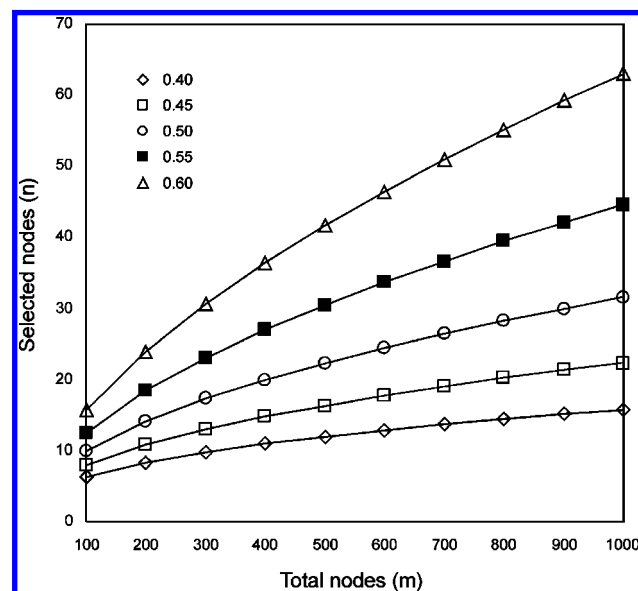

Figure 2. Values of n obtained for diverse values of m and of the weighting parameter a, using eq 1.

DEVELOPMENT AND VALIDATION OF AMANDA

J. Chem. Inf. Model., Vol. 48, No. 9, 2008 **1815**

optimum value of this weighting parameter a for drug design was 0.55, but other values could be used for other applications. In order to select these nodes, the $2.n$ nodes with the lowest energy values were inserted into a short list from which they were picked one by one using a simple algorithm. First, the node with the lowest energy value was selected. Then the Euclidean distances between the chosen node and the rest of the members of the short list were computed and added to their field energy values to obtain a simple scoring s_{ij} for every node i and algorithm step j (eq 2)

$$s_{i0} = E_i s_{ij} = s_{i(j-1)} + d_{ik} \qquad (2)$$

where E_i is the normalized energy for node i, and d_{ik} is the normalized distance between node i and node k, the one selected at the $(j$-$1)$ step.

The node with the best scores is selected, and the procedure is repeated, accumulating into the nodes scores both the nodes energy values and the distances to the selected nodes, until n nodes were finally extracted. Notice that the first nodes were selected mainly on the basis of their energy values, but as the method progresses, the weight of the energy value on the scoring decreases while the weight of the distance with the previously selected nodes grows.

In our tests, the described algorithm produces results that are an excellent compromise between a node selection highly concentrated on the field minima and a sparsely dispersed one. Figure 3 shows some examples of the hot spots selected by AMANDA for some druglike compounds.

It is worth emphasizing that the AMANDA algorithm guarantees to select at least one node for every atom of the molecule for which the list of prefiltered nodes is not empty (and thus, for every atom likely to participate in relevant interactions). The total number of nodes selected for a whole molecule is not fixed and depends on the number of atoms and groups able to produce relevant interactions. If the molecule contains no such atoms, the list of hot spots could be empty. AMANDA contains only two adjustable parameters (the energy cutoff value and the value of the weight a in eq 1) which can be tuned to the requirement of a certain specific application but that, once fixed, does not need to be adjusted for every series of compounds. This makes the algorithm particularly suitable for nonsupervised work and for the analysis of series containing highly diverse structures. From a computational point of view, the algorithm is simple to implement and runs very fast since it does not include clustering or iterative optimization steps. In the applications described here, the AMANDA algorithm was run using in-house written ANSI C programs, but it can be easily implemented in any programming language.

Hot Spots Validation. In order to be useful for drug design, the hot spots must depict the regions around a ligand which are more likely to participate in noncovalent bonding interactions with its receptor. Therefore, the quality of a hot spots extraction algorithm depends on how complete and accurate this picture is. In order to validate the method we decided to generate hot spots for a collection of ligands for which experimentally obtained ligand−receptor complexes are available. This allows comparing the position of the hot spots with the position of corresponding functional groups present at the receptor binding site. For example, in a complex between a glucose molecule and a glucose receptor, the hot spots obtained from an H-bond acceptor MIF should
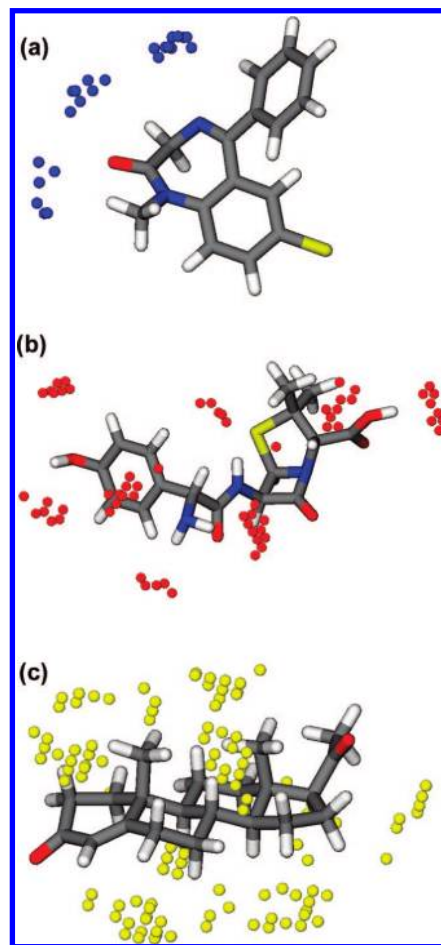

Figure 3. Hot spots selected by AMANDA from different types of GRID computed MIF. (a) H-bond donor probe (N1) on a molecule of diazepam. (b) H-bond acceptor probe (O) on a molecule of ampicilin. (c) Hydrophobic probe (DRY) on a molecule of progesterone.

overlap the receptor binding site atoms that participate in H-bonds with the ligand. When carrying out this comparison two kinds of errors are possible: finding hot spots where no interaction can exist (false positive, lack of specificity) and not finding hot spots where interactions do exist (false negative, lack of sensitivity). Therefore, the quality of any such algorithm can be quantified in terms of sensitivity and specificity using eqs 3 and 4, respectively

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \qquad (3)$$

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \qquad (4)$$

where TN represents True Negative; TP represents True Positive; FN represents False Negative; and FP represents False Positive.

However, the application of these equations first requires the definition of the criteria for deciding when a node and a binding site atom are near enough to consider that the first overlaps the second (TP), when a node is not overlapping a binding site atom (FP), and when the absence of a node can be considered correct (TN) or a mistake (FN). Before entering into details of the method used here we must say that no computational method can be expected to produce a perfect measurement of the specificity and sensitivity and that some errors are unavoidable. For example, Figure 4a (PDB entry 1swg) illustrates a common situation in which a ligand exposes some polar groups to the solvent, but the
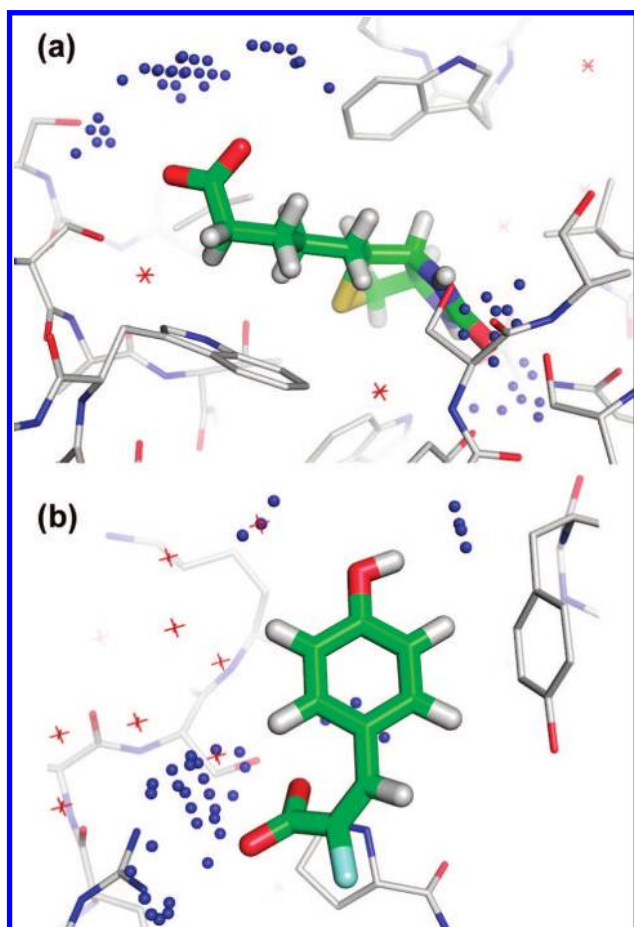
Figure 4. (a) Example of a mistake due to the lack of explicit solvent molecules in crystallographic complexes. All the nodes in the upper part of the figure were incorrectly identified as false positives due to the lack of explicit water molecules in the solvent exposed face of the ligand (PDB entry 1swg). (b) Example of a mistake due to the equivalent binding positions of a phenolic hydroxyl group. The three nodes on the upper left neatly overlap a water molecule, but the four nodes on the upper right side overlap no atom and were incorrectly labeled as false positive (PDB entry 1gyy).

Figure 5. Scheme of the computational procedure used to define interacting and noninteracting regions in a ligand−receptor complex.

crystal does not include enough explicit water molecules to match all of the ligand H-bonds. In this case, the hot spots are incorrectly accounted as FP. Another common source of mistakes is the presence in the ligands of polar groups with rotatable hydrogen atoms, since sophisticated MIF computation methods like GRID take into account all the hydrogen accessible positions for computing the MIF values. In the example shown in Figure 4b (PDB entry 1gyy), the phenol group could establish H-bond interactions with an H-bond donor group at any of the two alternative and equivalent orientations represented by the two clusters of blue points but only at one of these positions. Indeed, one of the clusters overlaps a water molecule present in the complex but the other does not, and these nodes are counted (again incorrectly) as false positive results. For all these reasons, the results reported herein are useful to compare the quality of different methods, but their interpretation in absolute terms must be made with care.

In this work and in order to obtain an objective quantification of the hot spots quality we developed an ad-hoc computational procedure for the analysis (Figure 5). The procedure starts from a MIF obtained only for the ligand. Since the analysis must be focused on the ligand neighbor-
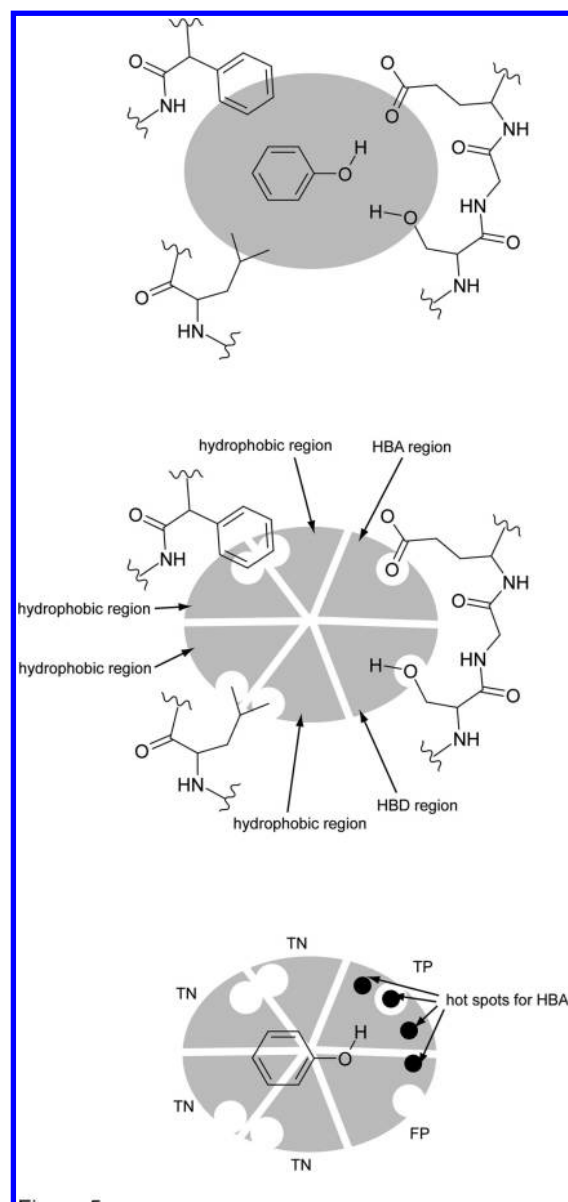
hood, the MIF nodes located farther than 4 Å from any ligand atoms were discarded. Then, the remaining nodes were assigned to the nearest receptor atom, thus defining a certain number of regions, every one linked to a single heavy atom of the receptor binding site. With respect to a certain probe, these regions can be labeled as "interacting" if the receptor atom can make the same kind of interaction (hydrophobic, H-bond donor, or H-bond acceptor) than the considered probe and is closer to the ligand than a certain cutoff distance (4.2 Å for hydrophobic interactions and 3.2 Å for polar probes). Otherwise, the region is labeled as "noninteracting". Once all the regions were labeled, they can be compared with the hot spots: interacting regions represent a true positive (TP) when they contain at least one hot spot point or a false negative (FN) when they do not. Conversely, noninteracting regions represent a false positive (FP) when they contain at least one hot spot point or a true negative (TN) when they do not. The total number of these four kinds of regions, for a certain complex, probe and hot spots method can be

accounted and used to compute the specificity and sensitivity provided by the hot spots description in this particular case.

GRIND Computation. The computation of GRIND includes a hot spot analysis step, which was described in detail in the original article.[7] The procedure starts with a GRID generated MIF. As in AMANDA, the field is prefiltered by removing all the nodes with energies over prespecified thresholds (see Table 1). In this case the purpose is not to remove nonspecific interactions, like in AMANDA, but only removing very small values, and therefore the thresholds applied were much less strict (less negative) than in AMANDA. Then, all the remaining nodes were subjected to an iterative optimization procedure with the aim to obtain a set of nodes of a prefixed size (usually 100 or 150 nodes and constant for all the molecules in the series) with the highest possible score. The optimization is carried out using a modified version of the Fedorov interchange algorithm,[9] and the score given to a certain node set is computed as the weighted average of two criteria: the sum of the node energies and the sum of their mutual Euclidean distances, being this weight a parameter provided by the user. In this study the GRIND hot spots were obtained using the original algorithm,[7] as implemented in the ALMOND software.[10] The weight was set to 50%, and the number of nodes was set to 150 in all instances, unless specified.

Minima Computation. Field minima were extracted from GRID MIF using the programs MINIM and FILMAP included in the GRID 22 package.[8] These programs were designed as tools for neutralizing formal charges in biological molecules, by inserting a number of counterions at the MIF minima positions. They work in tandem; MINIM selects nodes in the MIF that have lower energy values than a certain threshold and which are surrounded in all directions by nodes with higher energy. FILMAP tries to populate optimally these candidate positions with a certain number of counterions. In this study, the MINIM program was configured to obtain minima under the following energy cutoff values (DRY, −0.25 Kcal/mol; O, −1.0 Kcal/mol; N1, −4.1 Kcal/mol), and the interpolation option was set to yes. FILMAP was configured to select from the candidate positions a maximum number of minima equal to one-fifth of the total number of ligand atoms. As in the previous case, these precise parameters were selected after visual inspection of the minima extracted for a large series of druglike compounds.

Comparative Analysis in QSAR Applications. The performance of the GRIND original method and a mixed GRIND-AMANDA method was compared by reinvestigating a collection of published GRIND studies.[7,11−13] We started by collecting the structures either by retrieving the original files from the authors or by rebuilding them, using the published procedures. The GRIND original models and the new GRIND-AMANDA models were built in parallel using our own in-house developed software, which reproduces exactly the GRIND published algorithm and makes use of the latest version of GRID program for computing the MIF. In all instances, the models were built using exactly the same GRIND parameters (probes, number of nodes, weights, etc.) reported in the original references. When the authors reported the use of FFD variables selection, it was carried out using program GOLPE 4.5,[14] using the same settings described in the original references. Despite our efforts, in no case we were able to reproduce exactly the q^2 and r^2 reported values,

and small differences (either slightly higher or lower values) were obtained. This could be a consequence of the use of slightly different conformations for the compounds of the series studied, changes in the GRID force field parametrization, and details on the implementation of the TIP probe which we were unable to reproduce exactly in our software. In any case, the observed differences should not affect the results for this analysis, since its only purpose is to obtain a comparison between the models obtained with the GRIND original and the GRIND-AMANDA methods.

Computation Speed Analysis. Special versions of AL-MOND and our in-house AMANDA software were prepared. In these versions, the system time was recorded at the beginning and at the end of the hot spots computation, and the time spent was shown. Both programs were compiled in the same computer with the same compiler and optimization settings. Such versions were used to compute GRIND and AMANDA hot spots in a realistic setting, using a set of 4200 compounds extracted from the DrugBank database[15] which contains only marketed drugs or druglike compounds. Three MIF were obtained for each compound using program GRID with probes DRY, O, and N1. All the computations were carried out in a Pentium 4 1.8 GHz workstation with Linux Centos 3.0 operating system.

Statistical Analysis. Statistical analysis (mean, median, Pearson correlation coefficient, and t-Student tests) were carried out using SPSS 12.0 software.

RESULTS AND DISCUSSION

AMANDA Algorithm. Figure 3 shows the resulting hot spots obtained with the AMANDA method (described in detail in the Methods section) from MIF computed on some drug structures. It should be noticed how the total number of nodes extracted is not fixed, and, for example, the description of H-bond donor regions around diazepam (Figure 3a) requires much less nodes than the description of hydrophobic regions around progesterone (Figure 3c). It must be also noticed how every single region is described by a handful of nodes (usually 5−10 nodes) which are conveniently scattered to cover most of the region. Moreover, when a single atom produces several regions, the algorithm guarantees that every one of them is conveniently represented in the results.

The AMANDA algorithm was specifically developed for applications in drug design. The aim was to obtain a method meeting the following specific requirements: (i) the hot spots must represent every single atom which could contribute to the binding affinity for a certain receptor, in order to identify all pharmacophoric relevant features; (ii) the method must work unsupervised, producing suitable results for structurally diverse compounds without requiring any specific adjustment; and (iii) the method must be fast enough to process a large number the compounds in a reasonable time. In all these three aspects, AMANDA works very well and outperforms other previously published methods.

With respect to the first aspect, in AMANDA the hot spots are extracted atom-wise, picking the more relevant nodes from a list obtained for every atom. Provided that the prefiltering levels are set up correctly, this method guarantees an exhaustive description of the ligand interactions. In this respect, AMANDA behaves much better than methods based
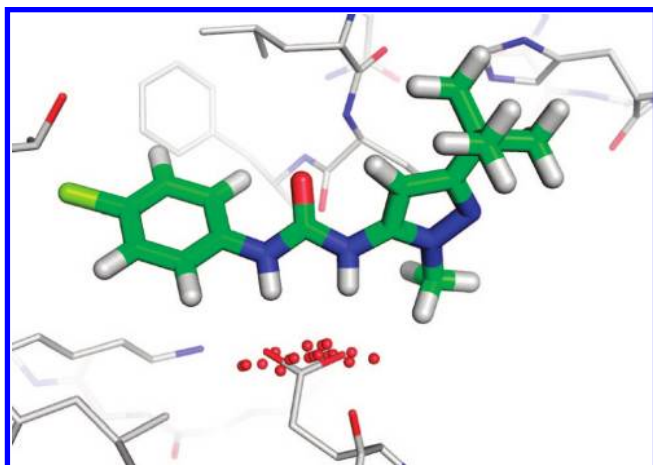
Figure 6. Result of AMANDA hot spots analysis on a MIF obtained using probe O on the binding site of a p38 MAP kinase (PDB entry 1kv1). The nodes clearly overlap the carboxylic group of a glutamic acid side chain, able to act as an H-bond acceptor.

only on field values like the extraction of minima, in which the presence of a function producing a strong interaction in the ligand produces hot spots only around this point, thus neglecting other functionalities that produce weaker (but maybe important) interactions. During the development of the GRIND hot spot method, the inadequacy of a node selection based only on the field values was recognized at an early stage, and the selection criteria was enriched by adding the mutual distances between all the nodes extracted. However, this solution has two drawbacks: (i) it does not solve the problem when the functional groups are close in the space and (ii) it is computationally demanding because it requires optimizing the whole point selection according to the aforementioned criteria using an iterative algorithm which must compute repeatedly, among other things, all the mutual distances between the selected points (e.g., 11100 distances for a set of 150 nodes).

With respect to the ability of the methods to work unsupervised and to produce good results for structurally diverse compounds, in the GRIND method the user must define explicitly in every run the total number of nodes to extract per molecule, usually 100 or 150. This is a major problem for the analysis of series containing structurally diverse compounds, because no single value can fit perfectly both small and large compounds and a compromise value must be found by trial and error, often requiring several runs. On the contrary, the set of hot spots obtained by AMANDA adjust their size automatically to the compound, containing more or less nodes according to how many ligand atoms are able to produce interactions and how intense are these interactions. A typical example of this situation is when the molecule under study is unable to produce specific interactions with a certain probe (for example, unable to donate H-bonds); AMANDA yields an empty list of nodes (the expected solution) in contrast with GRIND which produces the requested number of nodes in any case.

With respect to the speed of computations, AMANDA is a fast and efficient method that requires no iterative optimization algorithm, unlike GRIND, neither clustering algorithms, unlike some minima based methods.[6] In order to obtain an objective comparison between the AMANDA and GRIND computing speed we carried out a test using

ad-hoc versions of the respective programs which compute the time spent in the hot spots analysis step. Both methods were applied on a series of 4200 druglike compounds (obtained from the DrugBank database,[15] please see the Methods section for details) for which three MIF were computed. The results show that the average time required for extracting the hot spots from a single MIF was 443 ms for GRIND and 0.9 ms for AMANDA. Therefore, in a realistic test AMANDA runs more than 500-fold faster than GRIND, something not surprising if we bear in mind that GRIND uses a complex iterative optimization procedure while AMANDA is far simpler. No comparison was carried out for the minima method described here, because the computations were carried out in a multistep procedure which do not offer the possibility of a fair comparison.

AMANDA Validation. Computational methods must always be validated by applying them to practical problems. With this aim we decided to apply the method to the ligand structure of a large series of ligand−receptor complexes for which the experimentally determined structure is available. The use of these structures would allow running the hot spots methods on the ligands and then comparison of the position of the MIF nodes selected with the position of actual receptor atoms responsible of the ligand binding. This evaluation can be carried out by simple visual inspection, but when the collection of complexes is large the task can become cumbersome and it is then preferable to apply a computational method able to quantify objectively the quality of the hot spots results in terms of specificity and sensitivity. Here we will show a few examples of the application of the aforementioned hot spots methods on the complexes studied, illustrating some of the more interesting aspects that characterize every method, and then we will report the statistical analysis of the quality test results obtained for AMANDA, GRIND, and minima methods, using the hot spots validation procedure described in the Methods section.

The set used for this analysis was extracted from the refined set of 800 complexes present in the PDBbind database,[16] by removing only two entries (1h1s and 1ydr) which presented some problems for being imported into GRID. The PDBbind database was chosen because it contains well characterized complexes, with only one, noncovalently bound, ligand structure and the structures are readily available from the Internet (www.pdbbind.org) even if some structures of the ligands could not be considered to be druglike. The series contains ligands of very different sizes and polarity, including very large and very polar compounds (e.g., polysaccharides) the description of which was challenging for MIF-based methods. The ligand structures of the remaining 798 complexes were imported into GRID where we computed 3 MIF with the probes DRY, N1, and O, representing hydrophobic interactions, interactions with H-bond donor groups, and interactions with H-bond acceptor groups, respectively. The computations were carried out using default settings but for a grid step of 0.5 Å. The resulting 2394 (798 × 3) MIF obtained were submitted to hot spots analysis using AMANDA, GRIND, and field minima analysis, as described in the Methods section. Before entering into a detailed comparison of the results obtained for the different methods we can observe that in general the results were rather good. The representation of the hot spots results together with the receptor structure showed numerous instances of nearly
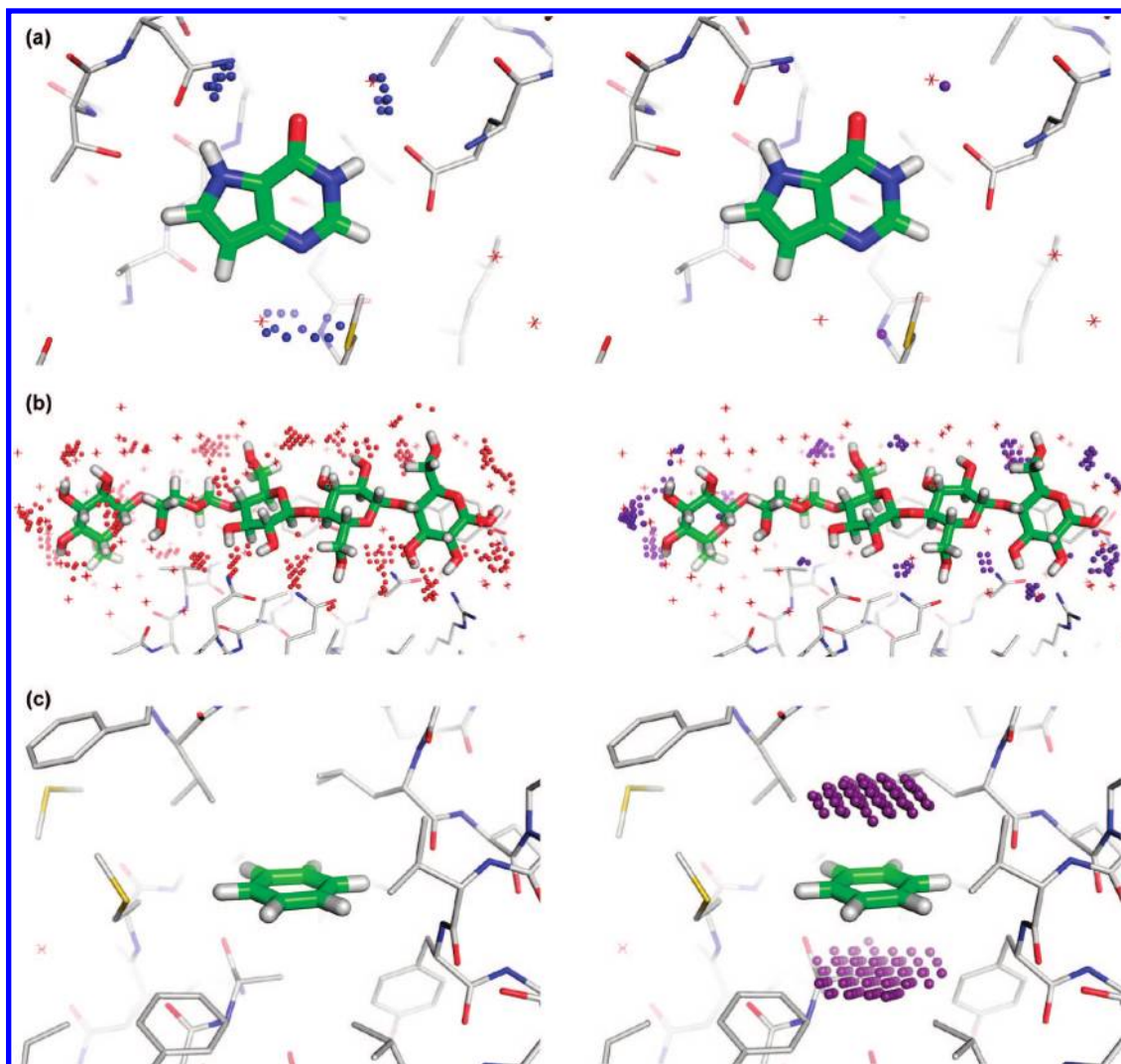
DEVELOPMENT AND VALIDATION OF AMANDA

J. Chem. Inf. Model., Vol. 48, No. 9, 2008 **1819**

Figure 7. (a) The minima results (right) computed on the N1 MIF represented here (a xanthine derivative complexed with *M. tuberculosis* purine nucleoside phosphorylase, PDB entry 1i80) overlap well the water and the arginine on the upper part but miss the water molecule on the lower part, while the AMANDA results (left) overlap all the relevant atoms. (b) The GRIND results (right) obtained on a O probe MIF for the oligosaccharide represented (PDB entry 1gu3) miss many water molecules and receptor atoms, while the AMANDA solution (left) is more complete. (c) N1 MIF obtained on the ligand of PDB entry 1l83. The GRIND results (right) extract 150 nodes even if no specific interaction is present, while the AMANDA list (left) is empty, as expected.

Table 2. Sensitivity and Specificity Values Obtained for a Set of 2394 MIF Analysis Using Diverse Hot Spots Methods

		minima	AMANDA	GRIND150	GRIND100
sensitivity	mean	0.037	0.694	0.722	0.663
	median	0.000	0.750	0.750	0.667
specificity	mean	0.995	0.803	0.769	0.812
	median	1.000	0.833	0.792	0.833

perfect overlapping of MIF nodes on top of relevant receptor atoms. Figure 6 shows an example of these results, in which the hot spots extracted from the O probe MIF using AMANDA overlaps nearly perfectly the carboxylic oxygen atoms of a glutamic acid residue present at the binding site of a p38 MAP kinase (PDB entry 1kv1).

The detailed visual inspection of the results obtained with the different methods for a sample of the complexes showed some consistent trends. The minima method produces very few nodes. The results from this method usually overlap precisely the atoms involved in the interactions, but they often fail to identify important receptor atoms when the MIF regions are large (e.g., the regions produced by DRY probe)

or when the ligand atoms can produce alternative orientations. This is the situation shown in Figure 7a (PDB entry 1i80) which shows a xanthine derivative bound to *M. tuberculosis* purine nucleoside phosphorylase. The hot spots obtained from the analysis of the N1 probe MIF with the minima method (right) overlap perfectly the nitrogen of an interacting arginine and a water molecule but fail to overlap the second water molecule shown at the bottom of the figure by more than 2 Å. On the contrary, the result of the AMANDA method on the same MIF contains more nodes per region, some of which overlap perfectly the involved atoms.

GRIND results are often rather similar to AMANDA results, with two interesting exceptions. In large ligands with many interactions like the oligosaccharide shown in Figure 7b (PDB id 1gu3), the GRIND method (right) runs out of nodes due to the preset fixed limit and fails to represent many of the relevant receptor atoms, while AMANDA adapts the number of nodes to the ligand characteristics and produces a rather complete description of all the receptor atoms involved in the ligand binding. This fixed number of nodes
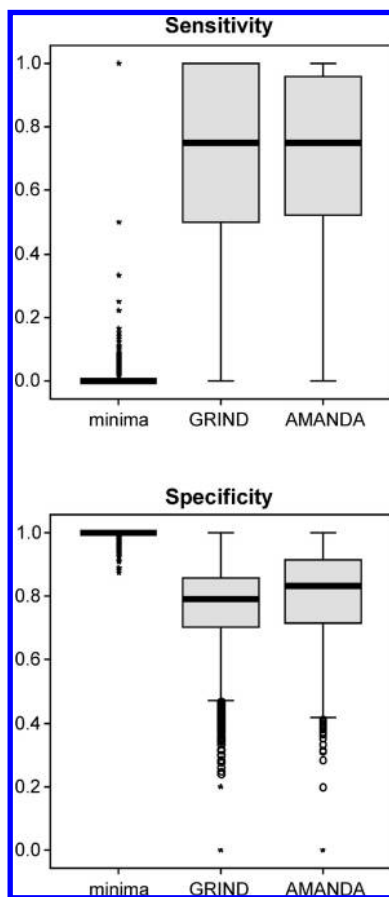
Figure 8. Box plot representing the sensitivity and specificity values obtained for a set 2394 MIF, using the method minima, GRIND, and AMANDA.

is also responsible of the GRIND problem illustrated in Figure 7c, representing a mutant of T4 lysozyme stabilized by buried benzene (PDB id 1l83). In this case, the binding site contains no polar atoms, but the GRIND analysis of the N1 MIF produces 150 nodes anyhow, that obviously do not overlap any relevant atom in the receptor. On the contrary, AMANDA produces again a correct result, by yielding an empty list of nodes.

Apart from the visual inspection of the complexes, the results obtained with the three methods were evaluated as described in the Methods section in order to quantify the sensitivity and specificity of the results. The quality of results were summarized in Table 2 and depicted in Figure 8. With respect to the sensitivity, both AMANDA and GRIND produce a more comprehensive description of the binding site than the minima method (in agreement with the example shown in Figure 7a) and therefore show much better results. The mean of the sensitivity is slightly higher for GRIND (0.72) than for AMANDA (0.69), even if the medians are identical (0.75). This is not surprising, because the results of the GRIND method, with the chosen number of nodes (150), describe the most relevant interactions for most compounds, with the only exception of very large compounds like the example represented in Figure 7b.

With respect to the specificity, the minima method yields the best results, because the few nodes extracted from field minima nearly always overlap a relevant atom. GRIND and AMANDA perform also well, but AMANDA results were significantly better than GRIND results (mean value of 0.80,

compared with 0.77, $p < 0.001$ using paired t-Student test) as shown in Figure 8. A good example of the specificity problems of the GRIND method was already shown in Figure 7c, where the constraint of producing always 150 nodes produces misleading results.

It is also interesting to report the correlation observed between the specificity and the number of potential ligand−receptor interactions (here represented by the total number of regions). As expected, due to the fixed number of nodes, the GRIND results exhibit a small but significant positive interaction between the number of regions and the specificity (Pearson correlation coefficient $r = 0.28$, $p < 0.001$) since the specificity obtained with this method is poor for compounds with few regions and better for larger compounds with more regions. On the contrary, the quality of the solution provided by AMANDA is rather stable, and the specificity does not show any statistically significant correlation with the number of regions ($r = -0.03$, $p = 0.21$). In order to further test the relationship between the quality of the GRIND results and the number of nodes setting, the analysis was repeated using 100 nodes. In this case, the specificity of the solution provided is much higher (0.81) and even slightly better than AMANDA (0.80), but the sensitivity is much lower than the previous case (0.66) and shows a statistically significant negative correlation with the number of regions ($r = -0.11$, $p < 0.001$) which indicates that the 100 nodes included in the solution are not enough for representing properly the largest ligands.

In summary, the comparison of the hot spots results in this set of complexes shows that the minima method has an unacceptably low sensitivity (median = 0.000) even if the specificity is very high. Both AMANDA and GRIND methods exhibit a good compromise between sensitivity and specificity, but AMANDA specificity is significantly higher and, what is more important, its results are more stable and are not dependent on choosing an appropriate number of nodes for every ligand.

3D QSAR Application. As previously stated, the AMANDA algorithm has many potential applications in the drug discovery process and is not limited to obtain molecular descriptors. However, in order to illustrate its suitability and further validate the method, we carried out an additional comparison by using the algorithm to build GRIND-like 3D QSAR models and evaluating the quality of the models obtained. With this aim, we have implemented the GRIND original method in ad hoc developed software, in which the hot spot selection can be run using either the originally reported algorithm or the new AMANDA algorithm described here. This software was used to reinvestigate a sample of eight previously published GRIND models,[7,11−13] for which the GRIND original and the GRIND-AMANDA models were obtained and compared as described in the Methods section. The results are shown in Table 3.

A first inspection of the results shows that the quality of the models obtained with GRIND-AMANDA, in terms of r^2 and q^2, is rather similar to the original models. Even if the values obtained in this comparison suggest a slight improvement (average q^2 was 0.65 for GRID-AMANDA and 0.62 for GRID original), the differences do not indicate a large improvement in this aspect. However, it should be noticed that in no one of the series studied the default values of nodes and weights were used for developing the original

Table 3. Results of the Reinvestigation of Eight GRIND Applications Using Both the GRIND Original and the GRIND-AMANDA Methods

series	probes	nodes[a]	weight[b]	GRIND original			GRIND-AMANDA			FFD
				LV	r^2	q^2	LV	r^2	q^2	
M1[7]	N1	150	70	2	0.88	0.66	2	0.83	0.67	yes
M2[7]	DRY, O, N1	120	35	3	0.90	0.78	3	0.89	0.75	yes[c]
M3[11]	DRY, N1	150	50	4	0.84	0.67	4	0.85	0.66	yes
M4[12]	DRY, O, N1	500	50	3	0.77	0.30	2	0.67	0.40	yes
M5[12]	DRY, O, N1	150	50	3	0.63	0.41	3	0.64	0.42	no
M6[13]	O, N1, TIP	120	50	2	0.96	0.90	2	0.97	0.90	yes
M7[13]	DRY, O, N1, TIP	200	25	3	0.92	0.68	5	0.97	0.71	yes[c]
M8[13]	DRY, O, N1	200	25	4	0.94	0.56	4	0.96	0.68	yes[c]

[a] Number of nodes. [b] Weight of field values in the GRIND filtering algorithm. [c] Some correlograms were manually excluded from the analysis.
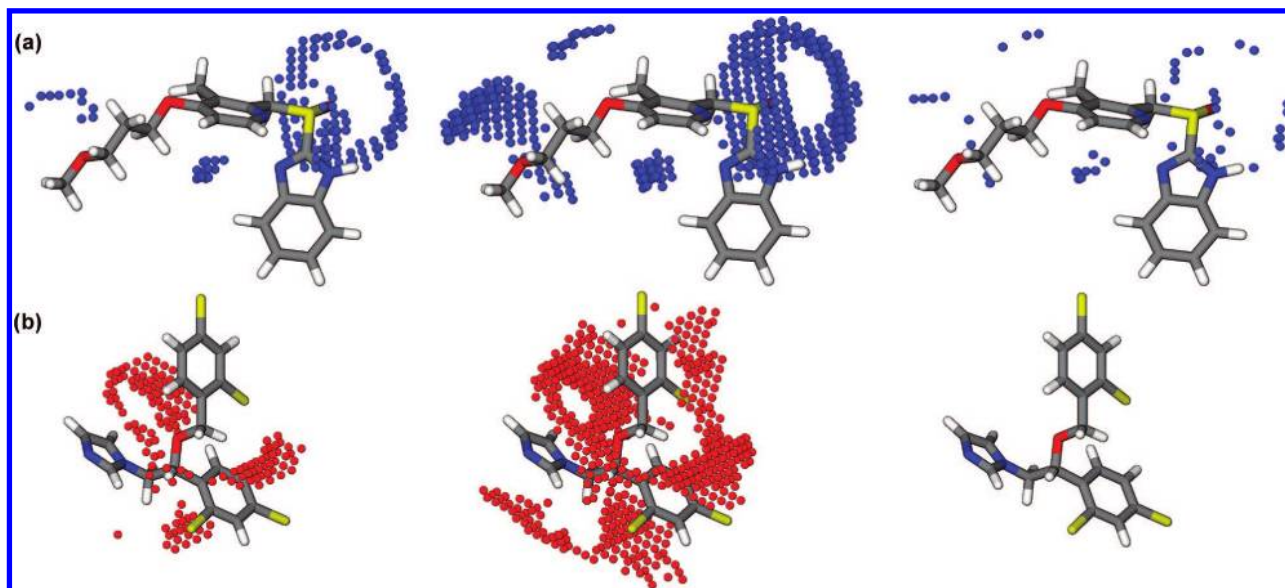

Figure 9. (a) Hot spots selected for S-rabeprazole and N1 probe with (from left to right) the GRIND original method and 150 nodes 50% weight, 500 nodes 25% weight and AMANDA. The description of one of the ether functions was missing in the 150−50 nodes setting and required to tune up the parameters to 500−25, while it was correctly described in the unmodified AMANDA description. (b) Hot spots selected for miconazole and O probe with (from left to right) the GRIND original method and 150 nodes 50% weight, 500 nodes 25% weight and AMANDA. Despite the fact that the compound has no H-bond donor, all the GRIND derived hot spot maps incorrectly show a large number of selected nodes. Conversely, the AMANDA results correctly reproduce an empty map.

GRIND models. This indicates that the final values of these parameters were found by means of several "trial and error" runs. On the contrary, the AMANDA method required no adjustment and was able to produce equally good or slightly better results in the first run. This is a very important advantage, since, in our experience, finding optimum settings requires between 5 and 10 runs and careful inspection of the results obtained by a trained chemist. Moreover, it is impossible to estimate how many QSAR models were discarded and remain unpublished due to the difficulty of finding adequate setting for the GRIND original method.

It is noteworthy that some of the largest improvements were obtained for the CYP2C9 M4 series.[12] In this application, the series was selected to contain highly dissimilar compounds, with molecular weights ranging from 100 to 1000. This is the typical situation in which it is impossible to find a fixed number of nodes which fits well with all the compounds in the series. Indeed, the authors reported the problem and identified the inability of the GRIND original method to provide a suitable representation of some ether functions unless the number of nodes was fixed to 500 and the field weight was set to 25%. On the contrary, the GRIND-AMANDA method was able to provide a good description

for the ether functions without any kind of adjustment, as it is shown in Figure 9a. Apart from the impact that this limitation of the GRIND original algorithm could have on the r^2 or q^2 of the models, its failure to identify pharmacophoricaly relevant groups seriously compromises the interpretability of the models. In the case of GRIND-AMANDA, the variables are present and provide a far more realistic model of these groups than the original method.

Another aspect related to the interpretability of the models is the fact that the GRIND original algorithm always produces the requested number of hot spots, irrespectively of the presence or not in the compound of chemical groups able to produce specific interactions. For example, Figure 9b represents the hot spots extracted from the O MIF for one of the compounds of the M4 series (miconazole), which has no H-bond donor group. It can be seen how the GRIND original method incorrectly highlights a lot of nodes, while the GRIND-AMANDA method produces an empty map. This kind of problem is very serious indeed and is likely to produce very dangerous mistakes in series containing a few compounds which lack some functions (H-bond donors or H-bond acceptors). In more congeneric series, the problem could be solved removing whole correlograms from the

analysis. The fact that for some of the studied series the authors decided to remove correlograms (M2,[7] M7,[13] and M8[13]) suggest that this problem was fairly common and had required additional trial and error runs for obtaining a model of reasonable quality. In this respect it is worth noting that the GRIND-AMANDA method was less sensitive to the presence of additional probes or correlograms and, for example, the use of the DRY probe in M6 increases slightly the predictive quality of the model (q^2 0.91) and in M8 the use of the standard probe set and of all the correlograms with no further selection yields a better model (q^2 0.75). More than these slight increases of the quality indexes it is relevant that AMANDA often produces the best models with standard settings and minimum adjustment.

CONCLUSION

AMANDA, a new algorithm for the extraction of highly relevant regions (hot spots) from MIF, was described. This algorithm was specially developed for producing suitable results for small molecules design in the field of drug discovery, since the extracted hot spots represent every single region of the MIF with a potential relevance in the interaction between this small molecule and the receptor. At the same time, this new algorithm requires no per series adjustment and can run unsupervised on large collections of compounds, being much faster than the other methods reviewed here. The method was thoroughly described and would be simple to implement in any programming language without effort. In addition, by adjusting the two tunable parameters, AMANDA can be easily adapted to other uses, like protein description.

The suitability of AMANDA results in the field of drug discovery was tested by comparing the hot spots extracted by this method with the position of real functional groups at the binding site, in a large number of complexes. The agreement between AMANDA results and the experimentally obtained complexes was first inspected visually for many complexes and then objectively quantified both in absolute terms and by comparing it with other alternative methods used to extract hot spots. The results clearly indicate that AMANDA results provide a reasonable description of the small molecules ability to interact with the receptors, thus capturing the most relevant information present in the MIF, and compare favorably with the results obtained using other methods, producing better quality indexes and more stable results.

The exercise of validation presented here, apart from being useful for comparing methods, is useful to illustrate the limits of the MIF concept. A MIF represents in the best case the potential of a ligand to interact with a receptor, but when alternative modes of interaction are accessible only one of them will resolve into an actual interaction being impossible to decide beforehand which one will interact and which will not. On the other hand, the structures of the complexes presented here represent only a single instance of the binding abilities of both the ligand and the receptor. The fact that the ligand is able to bind the receptor indicates that the binding is thermodynamically favorable but not that the ligand exhausts the ability of the receptor to establish interaction and vice versa. Therefore, if in complex A no receptor atom is complementary to a certain ligand func-

tionality, this does not mean that in complex B the same ligand functionality could not bind a receptor atom. This means that the quality indexes computed here are probably underestimated and should be used with caution.

An automatic hot spots analysis method with the properties of AMANDA has many potential uses in drug discovery. In order to test this assumption in practical drug design applications we used the AMANDA algorithm coupled with the MACC2[7] transform to obtain GRIND-like molecular descriptors and compared their performance with those obtained with the original GRIND method. In our tests, the new algorithm produced models of similar or slightly better predictive ability than the original ones but requires no trial and error runs in order to adjust any parameter and the hot spots lack two kinds of defects typical of the GRIND original hot spots, thus making the results easier to interpret. All in all, in our tests the application of AMANDA to replace the GRIND original hot spot algorithm produced slightly better models, easier to interpret and with less effort. Therefore, in conclusion, AMANDA offers the possibility of using hot spots as a universal replacement of MIF, allowing the use of the high quality description provided by MIF in a fast and simple way without any of the inconveniences traditionally associated with their practical application.

Supporting Information Available: 3D structures of the compounds included in the data sets M1, M2, M6, M7 and M8 (Table 3) and tables with the experimental, calculated, and predicted activity values obtained with the GRIND original and GRIND-AMANDA methods for the same data sets. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *7*, 849–857.

(2) *Molecular Interaction Fields. Applications in Drug Discovery and ADME Prediction*; Cruciani, G., Ed.; Wiley-VCH: Weinheim, 2006.

(3) Zavodszky, M. I.; Sanschagrin, P. C.; Kuhn, L. A.; Korde, R. S. Distilling the essential features of a protein surface for improving protein-ligand docking, scoring, and virtual screening. *J. Comput.-Aided Mol. Des.* **2002**, *12*, 883–902.

(4) Ciulli, A.; Williams, G.; Smith, A. G.; Blundell, T. L.; Abell, C. Probing Hot Spots at Protein-Ligand Binding Sites: A Fragment-Based Approach Using Biophysical Methods. *J. Med. Chem.* **2006**, *16*, 4992–5000.

(5) Cheeseright, T.; Mackey, M.; Rose, S.; Vinter, A. Molecular Field Extrema as Descriptors of Biological Activity: Definition and Validation. *J. Chem. Inf. Model.* **2006**, *2*, 665–676.

(6) Ermondi, G.; Anghilante, C.; Caron, G. A combined in silico strategy to describe the variation of some 3D molecular properties of β-cyclodextrin due to the formation of inclusion complexes. *J. Mol. Graphics Modell.* **2006**, *3*, 296–303.

(7) Pastor, M.; Cruciani, G.; McLay, I.; Pickett, S.; Clementi, S. GRid-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors. *J. Med. Chem.* **2000**, *17*, 3233–3243.

DEVELOPMENT AND VALIDATION OF AMANDA

J. Chem. Inf. Model., Vol. 48, No. 9, 2008 **1823**

(8) *GRID v22*; Molecular Discovery Ltd.: London, U.K., 2006.

(9) Fedorov, V. *Theory of Optimal Experiments*; Academic Press: New York, 1972.

(10) *ALMOND 3.3.0*; Molecular Discovery Ltd.: London, U.K., 2004.

(11) Benedetti, P.; Mannhold, R.; Cruciani, G.; Pastor, M. GBR compounds and mepyramines as cocaine abuse therapeutics: chemometric studies on selectivity using grid independent descriptors (GRIND). *J. Med. Chem.* **2002**, *8*, 1577–1584.

(12) Afzelius, L.; Masimirembwa, C. M.; Karlen, A.; Andersson, T. B.; Zamora, I. Discriminant and quantitative PLS analysis of competitive CYP2C9 inhibitors versus non-inhibitors using alignment independent GRIND descriptors. *J. Comput.-Aided Mol. Des.* **2002**, *7*, 443–458.

(13) Fontaine, F.; Pastor, M.; Sanz, F. Incorporating molecular shape into the alignment-free Grid-Independent Descriptors. *J. Med. Chem.* **2004**, *11*, 2805–2815.

(14) *GOLPE 4.5.0*; Multivariate Infometric Analysis: Perugia, Italy, 2004.

(15) Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res. 2008, Database issue, D901-6.

(16) Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind Database: Collection of Binding Affinities for Protein-Ligand Complexes with Known Three-Dimensional Structures. *J. Med. Chem.* **2004**, *12*, 2977–2980, The database is fully accessible at http://www.pdbbind.org.