Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 22 (2003) 23–30
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Analysis and optimization of structure-based virtual screening protocols (1): exploration of ligand conformational sampling techniques

Andrew C. Good [a,*], Daniel L. Cheney [b]

[a] *Structural Biology and Modeling, Bristol-Myers Squibb, P.O. Box 5100, Wallingford, CT 06492, USA*
[b] *Structural Biology and Modeling, Bristol-Myers Squibb, P.O. Box 5400, Princeton, NJ 08543-5400, USA*

Accepted 14 February 2003

Abstract

Ligand conformational flexibility has long been recognized as an important issue in virtual screening (VS). To this end, a number of different methodologies have been adapted to tackle the problem. Many of said techniques were originally designed for ligand derived pharmacophore screens, but have subsequently been fashioned for application within structure-based virtual screening (SVS). A popular adaptation is the pre-calculation of diverse ligand conformations for subsequent docking in target active sites. In this paper, we study a number of the software programs currently being used in conformer generation, analyzing their ability to regenerate known ligand binding conformations. The implications of these studies are discussed, from the perspective of VS in general and SVS in particular.

Keywords: Structure-based virtual screening; Conformer generation; Conformational sampling

1. Introduction

The core premise behind all virtual screening (VS) techniques is to screen a database of molecules computationally using structural descriptors that relate in some way to potential biological activity. Many methodologies are utilized, from two-dimensional (2D) topological and substructure similarity searches utilizing active ligand structures, to searches within the three-dimensional (3D) constraints of a target active site [1,2]. It is this exploitation of target active site data that forms the mainstay of structure-based virtual screening (SVS). An outstanding feature of this technique is that searches are not restricted to binding modes inferred from known ligands. Rather, whole active sites can be searched, permitting the discovery of ligands that bind with the target protein in hitherto unforeseen ways. This characteristic also poses a major challenge for SVS searching, one of time. If a large compound inventory is to be scanned quickly searches must be rapid. This is a significant problem for SVS, since the myriad potential binding modes to be explored create a huge search space within which to work.

Ligand flexibility forms one of the primary challenges facing all 3D VS methods, and numerous solutions to the problem have been presented over the years. Current CPU power makes a fairly exhaustive analysis of ligand conformations feasible for select ligand-based VS technology. The same cannot be said of SVS, however, where search space issues generally restrict the time available for conformational sampling.

SVS approaches to conformational flexibility fall into three primary categories. In the first, individual molecular conformations are generated independently of the active site using a variety of techniques. Conformers can either be generated “offline” and stored away for subsequent SVS searches [3,4], or generated at runtime [5]. The second approach is to dock the core fragments of molecules ahead of time before growing out or superimposing the flexible side chains [6–8], thereby mitigating the issue of conformer combinatorial explosion. The third technique undertakes full molecule conformational generation within the context of the active site of interest [9–11]. In this way, efficiency is gained by restricting conformer generation to regions of biological relevance. Because of our method of application using DOCK [12], we are particularly interested in the first category of conformer generation based on calculation independent of the site. Numerous conformational generation methods have been applied to try and cover conformational space as quickly and efficiently as possible using this technique. While the algorithms used are quite varied in approach, little attempt has been made to compare them in the context of the sampling settings currently applied within

* Corresponding author.
E-mail address: andrew.good@bms.com (A.C. Good).

1093-3263/03/$ – see front matter
doi:10.1016/S1093-3263(03)00123-2

SVS search paradigms. To gain an insight into this issue studies were devised determining the ability of said techniques to reproduce the bioactive conformation of known protein data bank [13] ligands. The result of these studies are described below.

2. Methods

The technique used to measure conformer quality is an extension of that presented by Ricketts et al. [14], which analyzed conformers generated by 3D structure builders and their relationship to crystallographic conformations. Each ligand was analyzed using multiple conformational generation techniques at multiple sampling levels. The conformers generated were superimposed onto the PDB ligand conformation, and the conformer with the lowest heavy atom rms was used as the quality measure.

A collection of 30 Ligands were selected from 237 complexes extracted from the PDB for which experimentally determined binding affinity data were available (Keske and Dixon, 1994, unpublished). These data had originally been applied in scoring function development by Gschwend et al. [15]. Ligands were selected by visual and numerical analysis to provide a reasonable degree of diversity in both functionality and flexibility. The final selection is shown in Fig. 1(a) and (b). The bioactive conformation was extracted from PDB file and atom types corrected for use in Sybyl [16]. No consideration was given to tautomeric forms (e.g. 0dfr_tmp), as these were not considered likely to effect the results. A smiles string was created for each ligand, and this was used to build a 3D structure of the ligand using CONCORD [17]. The resulting 3D structure was used as input to each conformational sampling technique. For each method, all the conformers generated were converted into a SYBYL mol2 database. The atom id mapping of the bioactive conformation was then mapped to that of the mol2 structures in the corresponding database. Finally, a SPL script incorporating the SYBYL *match* command was used to determine the conformer with the lowest rms to the bioactive conformer.

A variety of different conformational generation techniques were studied [2]:

(1) DOCK. Conformers were generated running DOCK in flexible mode with orientation sampling on a dummy target, where all results and conformers were accepted. The clash_overlap term (level at which atom–atom contact results in conformation rejection) was set to 0.7 times the van der Waals radius. To test various sampling levels the conformation_cutoff_factor was set to 3, 5 and 10 (number of conformers generated per bond deemed rotatable by DOCK, henceforth known as DOCK 3, DOCK 5 and DOCK 10 searches, respectively) in separate runs. The flex.defn was also modified somewhat in order to add and improve some of the rules used in torsion sampling. It should be noted that the in-house version of DOCK contains a number of minor changes in the code to prevent the program becoming trapped for long periods while attempting to generate conformers for heavily substituted molecules. These changes have little effect on the underlying algorithm, however.

(2) CONFIRM. CONFIRM [18] running in both “BEST” and “FAST” modes with default settings. Calculations were undertaken at maximum sampling levels of 5 and 10 conformers per single bond, plus 100 total (henceforth known as FAST 5, FAST 10, FAST 100, BEST 5, BEST 10 AND BEST 100, respectively). For the BEST 10 scenario the initial 3D structures were generated by using CATALYST [18] rather than CONCORD. This test was undertaken to see what effect a different 3D starting conformer would have on the final result.

(3) CONFORT. CONFORT [19] with rough (0.10 kcal) convergence, diverse conformer selection, acyclic bond search only, and maximum sampling at 5/10 conformers per single bond and 500 total (henceforth known as CONFORT 5, CONFORT 10 AND CONFORT 500, respectively).

(4) OMEGA. OMEGA [20] using a rms_cutoff of 1.0, and gps_energy_window set to 5.0. Maximum sampling was set to 100 conformers per structure (henceforth known as OMEGA 100).

(5) MACROMODEL [21]. Conformational generation was conducted using the Monte Carlo multiple minimum method (MCMM) with the GB-SA continuum solvent model for water and either MMFF [22], OPLS [23], AMBER [24], or MM3 [25] force fields (in that order of preference) depending on availability of parameters (force field with the most available parameters selected). A final energy cutoff of 4.8 kcal was used, consistent with Lijlefors’ findings regarding strain energies of bound conformations using similar force field and continuum models [26]. Runs were conducted so as to use as many MC steps as needed to produce the target number of conformations (5/10 per rotatable bond). The setting MCSS and MCNV were used to invoke use-directed Monte Carlo searching, i.e. the conformation least used is serves as a starting structure for the current Monte Carlo step. Duplicate conformations were removed using the default value of 0.25 Å rms.

3. Results

Sampling quality was measured in a number of ways. In the primary comparison, the average internal ranking and heavy atom rms deviation seen across the 30 ligands tested were compared. Heavy atom rms values were calculated based on the conformer closest to the bioactive conformation. Internal ranking was determined using the relative rank of the closest conformer of each technique to the bioactive conformer as calculated by heavy atom rms. The technique with the closest conformer, thus has rank 1, the

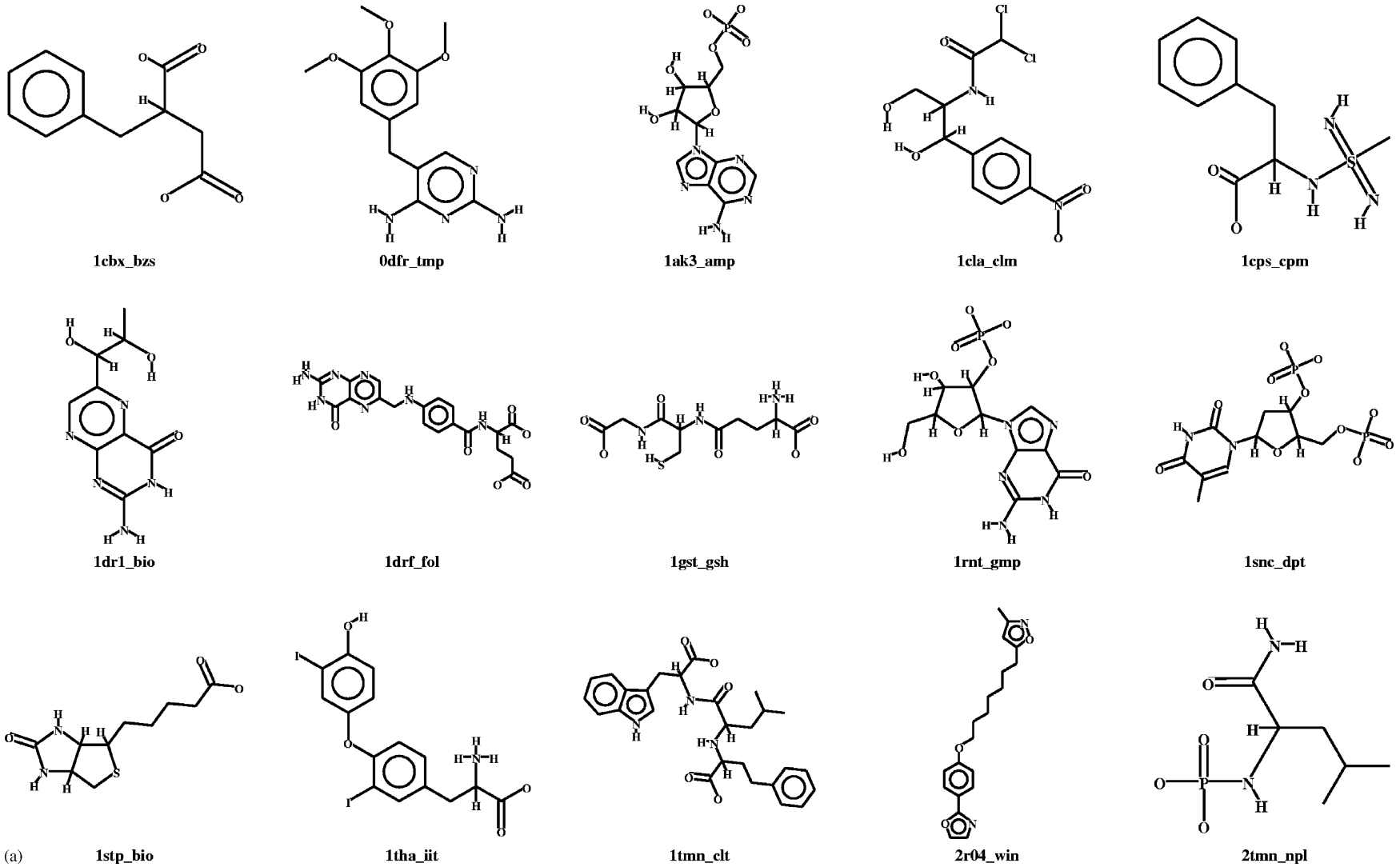
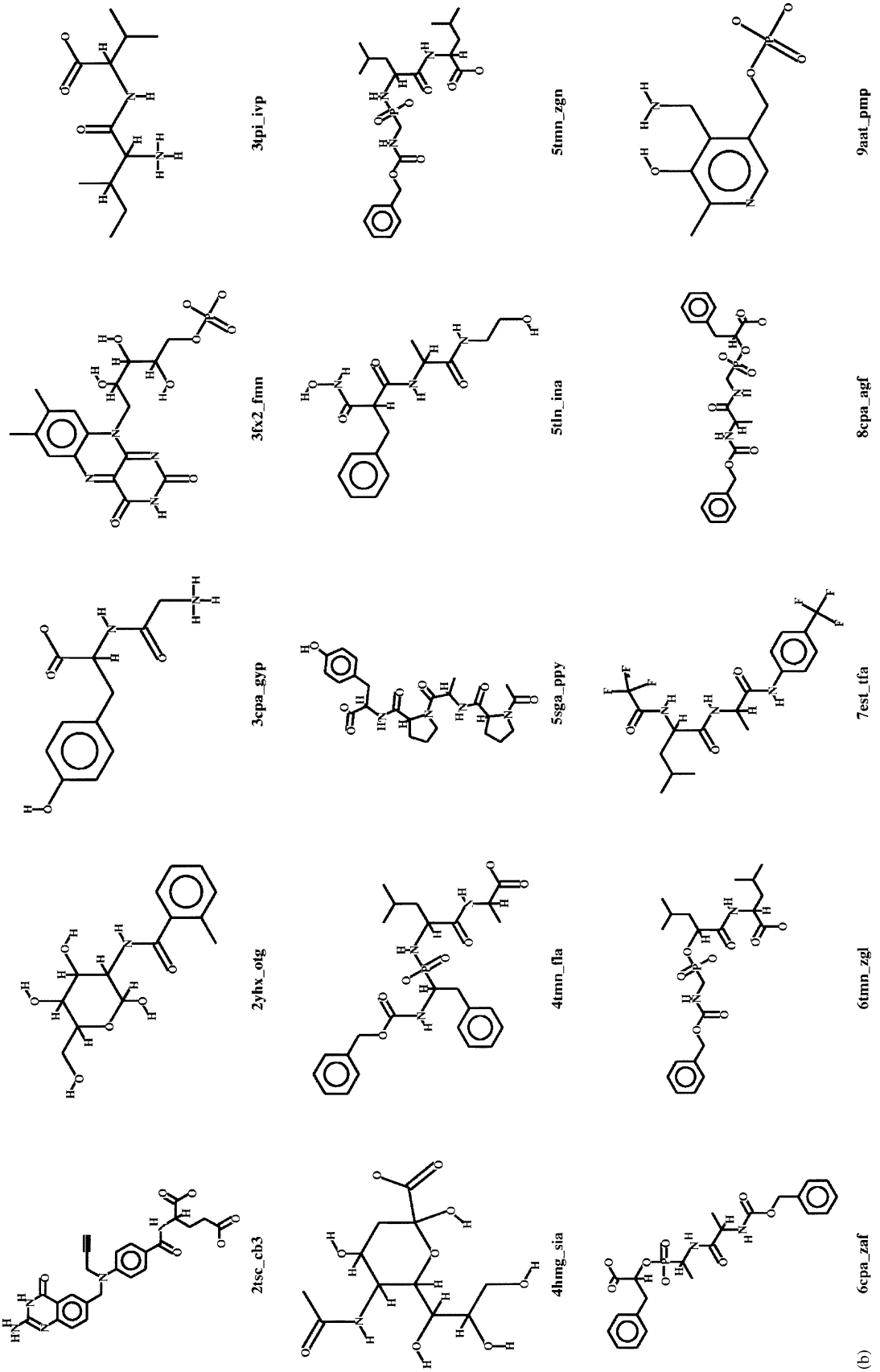


Fig. 1. Ligands analyzed in study. Name under each ligand refers to the PDB entry from which it was derived (first three letters), together with the ligand id (second three letters).



(b)

Fig. 1. (*Continued*).

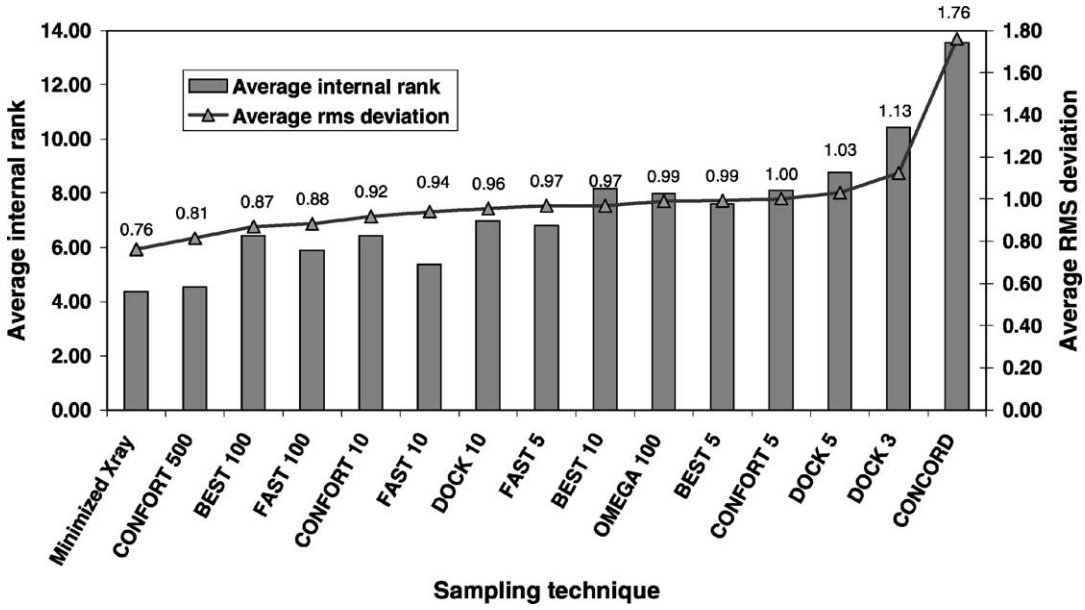


Fig. 2. Relative performance of the sampling techniques, based on average heavy atom rms deviation to ligand bioactive conformation and average internal rank of the rms deviations.

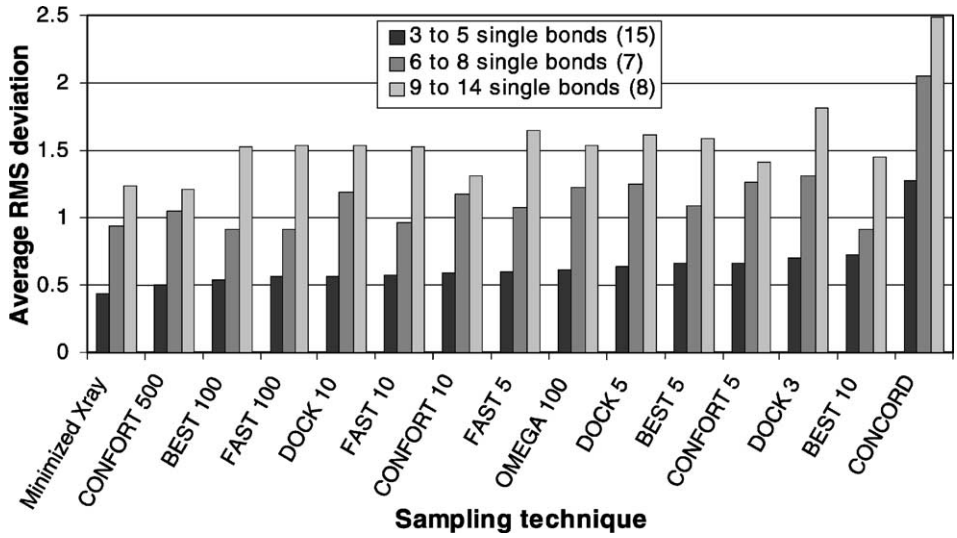


Fig. 3. Average heavy atom rms deviation performance (Å) broken down by ligand flexibility. Number of molecules associated with each flexibility group shown in brackets.

next rank 2, etc., down to 15 (the number of conformational techniques–sampling level combinations compared excluding MACROMODEL and including the CONCORD and minimized X-ray structures (TRIPOS force field minimized [16]). These ranks were averaged across all ligands to determine the average internal rank. The results are shown in Fig. 2.

The ligands were then divided up by their relative flexibility and the average rms across "flexibility group" was determined for each technique. The data for these calculations are shown in Fig. 3.

To test the improvement in performance on increased sampling, the average number of conformers generated for select techniques were determined and compared the average heavy atom rms results. Results are displayed in Fig. 4.

4. Discussion

As one would hope and expect, Figs. 2 and 3 illustrate the utility of conformational sampling. Comparison of any technique with the CONCORD conformer clearly illustrates

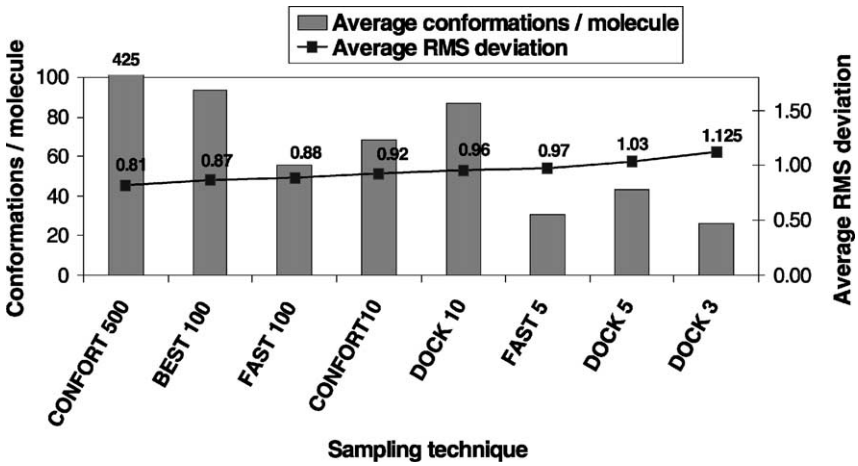


Fig. 4. Average heavy atom rms deviation performance (Å) for selected techniques compared with average number of conformers generated per molecule.

an improved ability to determine conformers with increased resemblance to the known binding orientation. Also, as one would expect the minimized X-ray structure on average bears the closest resemblance to the bioactive orientation. Interestingly, however, the average rank is not unity, illustrating that simple minimization can cause significant deviation. Modification of the force field can affect this, since changing from the TRIPOS to the MMFF force field within SYBYL reduces the average rms to 0.71. Generally, increasing the sampling improves the performance somewhat (Fig. 4). The most extreme example is the CONFORT 500 case, where average rms and rank are approaching that of the minimized X-ray structure. Given that CONFORT is using the TRIPOS force field in minimization, this sampling level is likely approaching its optimum performance. This comes at a hefty sampling level of 425 conformers per molecule, however, and in terms of optimum efficiency, it would appear that the FAST 5 run comes out on top. This sampling level showed good rms performance (0.97) with an average sampling of only 31 conformers per molecule. It is interesting to note that in general the FAST runs were found to perform at least as well as their BEST equivalents. In general, FAST was found to undertake calculation ∼50 times faster than BEST and at a comparable speed to the other techniques. It should be noted, however, that we found ∼10% of molecules processed in our databases could not be conformationally expanded using the FAST algorithm. All of the molecules from this set that were tested could be processed successfully using the BEST technique. The relatively poor performance of the BEST 10 run (which used a CATALYST [18] generated 3D starting structure) is also interesting in that it illustrates the potential variation that can occur when different initial conformers are used. DOCK sampling, which used the simplest conformer generation methods, was found to perform slightly less well than the other techniques although the differences were small. At sampling levels typically used in virtual screening calculations (all but CONFORT 500), rms deviations were found to break down according to the flexibility of the molecules analyzed. Compounds with less than five single bonds typically had rms deviations in the region of 0.5 Å. This was found to increase to ∼1.0 Å up to eight bonds, and 1.5 Å for 9–14 bonds (see Fig. 3). Fig. 5 illustrates some visual examples of rms deviations. At 0.6 Å, the

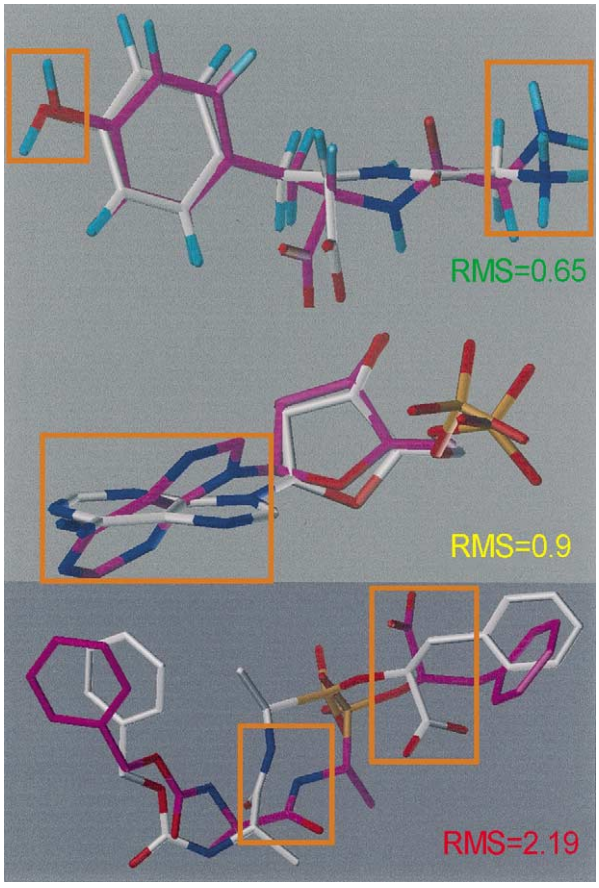


Fig. 5. Examples of conformers generated closest to bioactive conformer. These are provided to give a visual feel for the meaning of an rms deviation (Å).

molecules line up well, with all hydrogen bonding moieties in the same region of space. Even at this low rms, however, there is some deviation in hydrogen positions, making accurate calculation of highly directional descriptors (e.g. hydrogen bonding measures incorporating explicit hydrogens) a tricky proposition. As the rms increases to 0.9, general alignment is still good, but now significant deviation is possible in the exact positioning of larger portions of the molecule. Such imperfections should not pose too much problem for fuzzy descriptors, but those of a more exacting nature (for example, force fields incorporating an 'R_{12}' repulsion term) may struggle. When the rms increases further, the general shape and disposition of major functional groups remain reassuringly similar. The directionality of hydrogen bonding groups now exhibits significant breakdown, however, with amide flips and functional group inversion now possible.

The above observations are reinforced by the results of our studies using MACROMODEL. These calculations were designed to be most exacting of the conformational sampling techniques used. The resulting average rms deviation was 0.91, which is respectable but not outstanding. Only five molecules were found to have greater than five duplicates per structure which is generally thought to be a good empirical indicator of convergence (Schrödinger, private communication). This provides a further indication that the sampling levels currently used in many virtual screening calculations are often not sufficient to fully sample space. Additional evidence to this effect is forthcoming from the fact that even for closely related ligands (2tsc_cb3/1drf_fol and 6cpa_zaf/8cpa_agf) which contain only subtle steric differences, notable differences in rms performance can result (2tsc_cb3/1drf_fol = 1.72 versus 2.56 for OMEGA 100, 6cpa_zaf/8cpa_agf = 2.19 versus 1.78 for FAST 5). These results illustrate that at the sampling levels being forced on conformational search algorithms, even subtle changes can have a significant effect on the nature of the selected conformers.

The authors acknowledge the somewhat limited nature of the data set analyzed here. Analysis of more recent PDB entries would significantly increase the diversity of ligands available. Given the generally simple nature of the ligands tested this is likely a best case scenario, however, since more exotic functionality will likely only provide a stiffer test. Indeed, analogous studies, though focussed primarily on lead optimization conformational analysis using alternative quality measures and different ligand data sets, yielded broadly similar conclusions [27,28].

5. Conclusions

These studies have a number of potential implications for virtual screening search paradigms. First, current sampling levels provide a practical ceiling above which techniques begin to break down from the perspective of SVS calculations (at ∼eight single bonds). Even at levels below this cutoff, conformers, although bearing a striking resemblance to the bioactive conformation, still exhibit some deviations. This raises potential concerns for scoring function calculations where descriptors are highly directional or sensitive in nature. It also gives pause for thought when we consider the number of scoring functions being developed based on maximizing the score for the bioactive orientation. For such functions, search paradigms capable of higher sampling in biologically relevant space are likely to hold more potential, given the need to map the bioactive conformation closely. It also underscores the importance of post-docking minimization and refinement. With current sampling methods, the number of top scoring candidates for refinement is quite large; but this number may well decrease as sampling improves. While a potentially major issue for SVS, conformational quality for more flexible molecules is such that less exacting pharmacophore and fuzzy shape-based VS calculations should still be possible. Given the deviation in hydrogen bonding functionality observed, however, the use of site point vectors in such calculations would seem best avoided in favor of larger distance—positional tolerances. Further, at these sampling levels, there would appear to be little disadvantage in choosing simpler conformational sampling methods, given the limited difference in average rms deviation across the myriad techniques.

Development of sampling methodology continues apace: the program versions used in these calculations have all been superseded. In addition, no attempt was made to adjust the internal parameters of each method to optimize performance. Further efforts from either perspective would likely yield incremental improvements in performance. Nevertheless, the sampling restrictions inherent in current technique application may ultimately limit utility for more flexible molecules, particularly from the perspective of SVS calculations.

References

[1] W.P. Walters, M.T. Stahl, M.A. Murko, Virtual screening—an overview, Drug Discovery Today 3 (1998) 160–178.
[2] Virtual screening for bioactive molecules, in: H.J. Böhm, G. Schneider (Eds.), Methods and Principles in Medicinal Chemistry, Wiley-VCH, New York, 2000, p. 10.
[3] S.K. Kearsley, D.J. Underwood, R.P. Sheridan, R.T. Mosley, Flexibases: a way to enhance the use of molecular docking methods, J. Comput. Aided Mol. Des. 8 (1994) 565–582.
[4] B.E. Thomas, D. Joseph-McCarthy, J.C. Alvarez, Pharmacophore-based molecular docking, in: O.F. Guner (Ed), Pharmacophore Perception, Development, and Use in Drug Design, International University line La Jolla, USA, 2000, pp. 351–367.
[5] T.J. A Ewing, I.D. Kuntz, Critical evaluation of search algorithms for automated molecular docking and database screening, J. Comput. Chem. 18 (1997) 1175–1189.
[6] S. Makino, I.D. Kuntz, Automated flexible ligand docking method and its application for database search, J. Comput. Chem. 18 (1997) 1812–1825.
[7] D.M. Lorber, B.K. Shoichet, Flexible ligand docking using conformational ensembles, Protein Sci. 6 (1998) 938–950.
[8] M. Rarey, S. Wefing, T. Lengauer, Placement of medium-sized molecular fragments into active sites of proteins, J. Comput. Aided Mol. Des. 10 (1996) 41–54.

[9] G. Jones, P. Willett, R.C. Glen, A.R. Leach, R. Taylor, Development and validation of a genetic algorithm for flexible docking, J. Mol. Biol. 267 (1997) 727–748.
[10] C.A. Baxter, C.W. Murray, B. Waszkowycz, J. Li, R.A. Sykes, R.G.A. Bone, T.D.J. Perkins, W. Wylie, New approach to molecular docking and its application to virtual screening of chemical databases, J. Chem. Inf. Comput. Sci. 40 (2000) 254–262.
[11] M. Schapira, B.M. Raaka, H. Samuels, R. Abagyan, Rational design of novel nuclear hormone receptor antagonists, Proc. Natl. Acad. Sci. 97 (2000) 1008–1013.
[12] DOCK, developed and distributed by the Kuntz group, Department of Pharmaceutical Chemistry, 512 Parnassus, University of California, San Francisco, CA 94143-0446, USA (URL: http://www.cmpharm.ucsf.edu/kuntz).
[13] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucl. Acid Res. 28 (2000) 235–242.
[14] E.M. Ricketts, J. Bradshaw, M.M. Hann, F. Hayes, N. Tanna, D.M. Ricketts, Comparison of conformations of small molecule structures from the Protein Data Bank with those generated by Concord, Cobra, ChemDBS-3D, and Converter and those extracted from the Cambridge Structural Database, J. Chem. Inf. Comput. Sci. 33 (1993) 905–925.
[15] D.A. Gschwend, A.C. Good, I.D. Kuntz, Molecular docking towards drug discovery, J. Mol. Recognit. 9 (1996) 175–186.
[16] Sybyl, developed and distributed by Tripos Inc., 1699 South Hanley Road, Suite 303, St Louis, MO 63144, USA (URL: http://www.tripos.com).
[17] CONCORD 3.0, distributed by Tripos Inc. (see ref. [16]).
[18] CONFIRM, part of the CATALYST 3D (version 3.9) database searching and Pharmacophore hypothesis software, developed and distributed by Accelrys, 9685 North Scanton Road, San Diego, CA 92121, USA (URL: http://www.accelrys.com).
[19] R. Balducci, R.S. Pearlman, CONFORT: A rational conformation analysis tool. Am. Chem. Soc. 217 (1999) COMP-011 (Abstract Paper).
[20] OMEGA (version 1.0), developed and distributed by Openeye Scientific Software (URL: http://www.eyesopen.com).
[21] MACROMODEL 6.5, developed and distributed by Schrödinger Inc. (URL: http://www.schrodinger.com).
[22] T.A. Halgren, Merck molecular force field. Part I: Basis, form, scope, parameterization, and performance of MMFF94, J. Comput. Chem. 17 (1996) 490–519.
[23] W.L. Jorgensen, J. Tirado-Rives, The OPLS [optimized potentials for liquid simulations] potential functions for proteins, J. Am. Chem. Soc. 110 (1988) 1657–1666.
[24] S.J. Weiner, P.A. Kollman, A.D. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta Jr., P. Weiner, A new force field for molecular mechanical simulation of nucleic acids and proteins, J. Am. Chem. Soc. 106 (1984) 765–784.
[25] N.L. Allinger, Y.H. Yuh, J.H. Lii, Molecular mechanics: the MM3 force field for hydrocarbons. Part 1, J. Am. Chem. Soc. 111 (1989) 8551–8566 (Schrödinger implementation differs from the original by use of a Coulomb's law treatment of electrostatics and torsional barrier treatment of conjugation).
[26] T. Liljefors, J. Bostrom, P.-O. Norrby, Conformational energies of protein-bound ligands, vol. 42, in: Proceedings of the Alfred Benzon Symposium on Rational Molecular Design in Drug Research, 1998, pp. 246–263.
[27] D.H. Robertson, M. Jalaie, D.J. Cummins, M. Vieth, Validation study of conformer generators using PDB ligand structures, Am. Chem. Soc. 221 (2001) CINF-087 (Abstract Paper).
[28] J. Bostrom, Reproducing the conformations of protein-bound ligands: a critical evaluation of several popular conformational searching tools, J. Comput. Aided Mol. Des. 15 (2001) 1137–1152.