Selection-Insertion Schemes in Genetic Algorithms for the Flexible Ligand Docking Problem

Camila S. de Magalhães, Helio J.C. Barbosa, and Laurent E. Dardenne

Laboratório Nacional de Computação Científica
Av. Getúlio Vargas 333
25651 070 Petrópolis RJ, BRAZIL
{camilasm,hcbm,dardenne}@lncc.br

Abstract. In this work we have implemented and analyzed the performance of a new real coded steady-state genetic algorithm (SSGA) for the flexible ligand-receptor docking problem. The algorithm employs a grid-based methodology, considering the receptor rigid, and the GROMOS classical molecular force field to evaluate the energy function. In the implementation we used the restricted tournament selection (RTS) technique in order to find multiple solutions and also introduced a new variation of this technique with an insertion criterion based in the root-mean-square-deviation (RMSD) between the ligand structures. The SSGA was tested in docking four HIV1 protease-ligand complexes with known three-dimensional structures. All ligands tested are highly flexible, having 12 to 20 conformational degrees of freedom. The implemented docking methodology was able to dock successfully all flexible ligands tested with a success ratio higher than 90 % and a mean RMSD lower than 1.3 Å with respect to the corresponding experimental structures.

1   Introduction

With the increasing amount of available biological structures obtained by experimental techniques (*e.g.* X-ray and NMR), ligand-receptor docking approaches have been very important and useful tools in structure-based rational drug discovery and design [1]. The docking problem is a difficult optimization problem involving many degrees of freedom, and the development of efficient docking algorithms and methodologies would be of enormous benefit in the design of new drugs [2]. For a protein/receptor with known three-dimensional structure, the ligand-protein docking problem basically has two main challenges: (i) development of an algorithm that efficiently searches a very complex energy landscape in order to predict the conformation and orientation of a potential drug molecule relative to the protein active site; (ii) prediction of ligand-receptor binding affinities [3] (*i.e.*, development of a computationally viable free-energy evaluation model) in order to correctly discriminate between different binding modes of the same ligand and/or to find the best drug candidate in a set of ligands. In this work, we focus our attention in the first aspect.

During the ligand binding in the protein active site, the protein and the ligand undergo conformational changes. One of the major problems in molecular docking is how to deal with protein and ligand flexibility, taking into account hundreds of thousands of degrees of freedom in the two molecules. In the last few years, several docking programs have been developed [4,5,6]. Some of them treat the receptor and the ligand as rigid body molecules considering only the ligand translational and orientational degrees of freedom [7]. Other docking algorithms also include the ligand flexibility considering the ligand conformational degrees of freedom [8,9]. In the two docking classes above, the protein structure is fixed in the position of the experimental structure. Docking large, highly flexible ligands is still a challenge to even the most sophisticated current docking algorithms, and adding the receptor flexibility remains a major challenge [10,11,12]. Genetic algorithms, with different strategies, have been shown to be a promising search algorithm for the ligand-protein docking problems [9,13,14]. In this paper, a non-generational, also referred as steady-state genetic algorithm (SSGA) [15] is adopted in association with a grid-based methodology, considering the receptor rigid, and the GROMOS [16] classical molecular force field to evaluate the energy function. A ligand conformation is represented by a chromosome constituted by real valued genes representing ligand translational, rotational and conformational degrees of freedom. The individuals are evaluated by a fitness function (*i.e.*, total interaction energy between the protein and the ligand molecule plus the intramolecular ligand energy). Individuals in the population are selected to reproduction in accordance with their fitness (a lower energy means a higher fitness), and suffer mutation or crossover operations, to generate new individuals. We implemented the SSGA using the restricted tournament selection (RTS) [17] technique in order to find multiple solutions, and we also introduced a new variation of this technique with an insertion criterion based in the root-mean-square-deviation (RMSD) between the ligand structures. The algorithm performance is tested in four HIV1 protease-ligand complexes with known three-dimensional structures. In all four tested complexes the receptor structure is assumed to be rigid. All ligands tested are highly flexible, having 12 to 20 conformational degrees of freedom (*i.e.*, dihedral angles considered flexible). The HIV1 protease enzyme is an important molecular target in the development of drugs against the AIDS virus. Additionally, the HIV1-protease enzyme has an enclosed active site and this fact, together with the large conformational flexibility of the ligands considered, make our test suite a good challenge to docking algorithms. The basic idea which guided us to introduce the RTS technique is the fact that with the increasing number of ligand conformational degrees of freedom the ligand-protein energy landscape becomes so complex that a great number of good minima solutions can be found closer to the global one. Moreover, the correct experimental structure may be no longer the actual global minimum in our model, due to approximations inherent to the energy function/model employed. Algorithms adapted to find several minima solutions can be a good choice to overcome or minimize those problems.

2   Methods

2.1   Genetic Algorithms

Genetic algorithms (GAs) are inspired in Darwin's theory of evolution by natural selection and are powerful tools in difficult search and optimization problems [18, 19]. GAs work with a population of individuals where each individual represents a possible solution for the problem to be solved. In the ligand-protein docking problem a candidate solution specifies the position of the ligand with respect to the protein. In a steady-state GA (SSGA) there is no separation between consecutive generations of the population since each offspring is created and immediately tested for insertion in the population. The SSGA is stopped when the maximum number of objective function evaluations allowed is reached.

2.2   The Solution Representation

In the implemented SSGA the individual chromosome has three genes representing the ligand translation, four genes representing the ligand orientation and the other genes represent the ligand conformation. The translational genes are the X, Y, Z reference atom coordinates (usually the closest atom to the ligand center of mass). The rotational genes are a quaternion [20] constituted by a unit vector and a rotation angle. The conformational genes are the ligand dihedral angles (one gene to each dihedral angle). The number of degrees of freedom n is thus 7 (translational plus rotational genes) plus the number of dihedral angles of the ligand molecule. To each chromosome –a real vector in R^n– a unique ligand spatial structure (coordinates for all atoms) is associated. The distance between two structures/solutions will be denoted by RMSD, which is defined by the root-mean-square deviation between the coordinates of all atoms of those structures. This can be seen as a distance between two solutions in phenotype space whereas the euclidean distance between two chromosomes is the corresponding measure in the genotype space. In fact, due to the different ranges of the translational, rotational, and conformational genes, a weighted euclidean norm was adopted here in the genotype space.

2.3   The Fitness Function

The fitness function is the ligand-protein interaction energy plus the ligand intramolecular energy, that are evaluated using the GROMOS classical force field [16], implemented in the molecular mechanics/dynamics THOR program [21], which is given by:

$$\sum_{protein}\sum_{ligand}\left\{p\left(\frac{A_{ij}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}\right)+\frac{q_iq_j}{D(r_{ij})r_{ij}}\right\}+$$

$$\sum_{ligand}\sum_{ligand}\left\{\frac{A_{ij}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}+\frac{q_iq_j}{D(r_{ij})r_{ij}}\right\}+\sum_{dihedrals}\gamma_k(1+\cos(\omega_k\theta_k-\theta_{0k}))\quad(1)$$

where r_{ij} is the distance between the atoms i and j; A_{ij} and B_{ij} are the Lennard-Jones parameters; q_i and q_j are atomic charges, and D is a sigmoidal distance-dependent dielectric function [22]; γ_k is the energy constant associated with a chemical bond rotation, θ_k is the torsion angle, ω_k is the periodicity, and θ_{0k} is the phase angle.

The first term of the equation corresponds to the van der Waals and electrostatic interactions between the protein and the ligand molecule, and the last two terms correspond to the ligand intramolecular energy.

We have introduced in eq. (1) a multiplier (p) in the protein-ligand van der Walls interaction energy term which varies with the counter of function evaluations ($neval$) according to:

$$p=\frac{neval}{c\times maxeval}\quad(2)$$

where $maxeval$ is the maximum number of function evaluations, and c was set to 0.5. The van der Waals interaction term is thus slowly introduced. As a result, ligand conformations that initially have bad van der Waals contacts with the protein have a chance to improve them in the following generations.

The ligand-protein docking problem involves millions of energy evaluations, and the computational cost of each energy evaluation increases with the number of atoms (typically thousands) of the ligand-protein complex. To reduce the computational cost we implemented a grid-based methodology where the protein active site is embedded in a 3D rectangular grid centered in the protein active site. At each point of the grid the electrostatic interaction energy and the van der Waals terms for each ligand atom type are pre-computed and stored taking into account all the protein atoms. The protein contribution at a given point is obtained by tri-linear interpolation in each grid cell.

Each individual in the initial population is placed in the grid center and then the translational genes are perturbed according to a Cauchy distribution. In this way, they are generated with higher probability near the grid center but still permitting individuals far from it. The Cauchy distribution is given by:

$$C(\alpha,\beta,x)=\frac{\beta}{\pi(\beta^2+(x-a)^2)}\quad(3)$$

where $a=0$ and $\beta=0.75$ are the Cauchy distribution parameters used. Genes corresponding to angles (dihedrals and rotationals) in degrees are randomly generated in $[0,360]$, and those corresponding to the rotational unit vector in $[-1,1]$.

2.4   The Genetic Operators

The genetic reproduction operators adopted are listed below:

- The two-point crossover (2-X) operator generates two offspring by exchanging the genes between two randomly chosen cut points in the parents chromosomes.
- The simulated binary crossover (SBX), which assigns more probability for offspring to remain closer to their parents than away from them, generates two offspring as described in [23].
- The non-uniform mutation (NUM) operator [24], when applied to an individual x_i at generation gen, mutates a randomly chosen variable x_i^j according to

$$x_i^j \leftarrow \begin{cases} x_i^j + \Delta(gen, b^j - x_i^j) & \text{if } \tau = 0 \\ x_i^j - \Delta(gen, x_i^j - a^j) & \text{if } \tau = 1 \end{cases} \tag{4}$$

where a^j and b^j are respectively the lower and upper bounds for the variable x^j, τ is randomly chosen as 0 or 1, and the function $\Delta(gen, y)$ is defined as

$$\Delta(gen, y) = y(1 - r^{(1 - \frac{gen}{maxgen})^\beta}) \tag{5}$$

with r randomly chosen in $[0, 1]$ and the parameter β set to 2. It is clear that this operator reduces the amplitude of the perturbations as the number of generations increases.

2.5   Selection and Insertion Schemes

Due to the high modality of the fitness landscape for the docking problem, a critical issue is the maintenance of useful population diversity in order to permit the investigation of several high fitness regions in parallel and reduce the chances of convergence to low quality local optima. Among the techniques proposed to deal with high modality landscapes [25], *fitness sharing*, introduced by Holland [18] and enhanced by Goldberg & Richardson [26], has the drawback of requiring knowledge about the search space (such as distance between optima) in order to set the dissimilarity threshold. *Crowding*, introduced by De Jong [27] and enhanced by Mahfoud [28] insert new offspring in the population replacing similar ones. We are particularly interested in the idea of restricted tournament selection (RTS) proposed by Harik [17] which nicely blends with our SSGA.

In this work, we have tested three selection-insertion schemes: (i) rank-based selection [15] of parents with replacement of the worst individual in the population, (ii) restricted tournament selection (RTS) [17], and (iii) a new modified RTS scheme.

In the RTS scheme, parent individuals are selected randomly from the population and the new offspring generated is placed in the population replacing the closest existing individual found in a tournament of size w, provided that the new individual is better than the winner of the tournament. The metric used was euclidean norm weighted so that all genes have the same influence in spite

of their different ranges. It is important to point out that $w = 1000$, does not mean that all individuals in the population will take part in the tournament. As the selection is random, one individual can be drafted more than once or not be drafted at all. We have also implemented a new modified RTS scheme where two tournaments are made. In the first (resp. second) tournament $w1$ (resp. $w2$) individuals that are better (resp. worse) than the new offspring are drafted. The winner of the first tournament, $CBetter$, is the closest individual (in the genotype space) to the new offspring, among the $w1$ individuals drafted in the first tournament. The winner of the second tournament, $CWorse$, is the closest individual (in the genotype space) to the new offspring, among the $w2$ individuals drafted in the second tournament.

The offspring is then inserted in the population in the following way:

– If the new offspring is closer to $CWorse$ than $CBetter$, then $CWorse$ is replaced by the newly generated offspring
– Else, If the RMSD between the new offspring and $CBetter$ is greater than 2.0 Å, then $CWorse$ is replaced by the new offspring. Otherwise, the new offspring is discarded.

In both cases the new offspring replaces $CWorse$. The modified RTS scheme uses information both from genotype space (chromosome) and the phenotype space (RMSD of all atoms coordinates). The criterion RMSD ≤ 2.0 Å is used to avoid an offspring insertion when a very similar and better individual already exists in a particular region of the search space. In this work we used $w1=w2$. If $w1 = w2 = 100\%$ is used, it means that the tournament size $w1$ is equal to the number of individuals that are better than the offspring, and the tournament size $w2$ is equal to the number the individuals that are worse than it.

3   Results

We have tested the SSGA on four HIV1 protease-ligand complexes. The experimental structures were obtained from the Protein Data Bank (PDB). The number of dihedral angles/torsions, total number of degrees of freedom (dimension) and the PDB file code, for each ligand molecule, are shown in Table 1. The structures and the dihedral angles of the four ligands tested are shown in Figure 1.   The grid is centered in the protein active site, with 23 Å of dimension

Table 1. HIV-1 protease ligands complexes tested

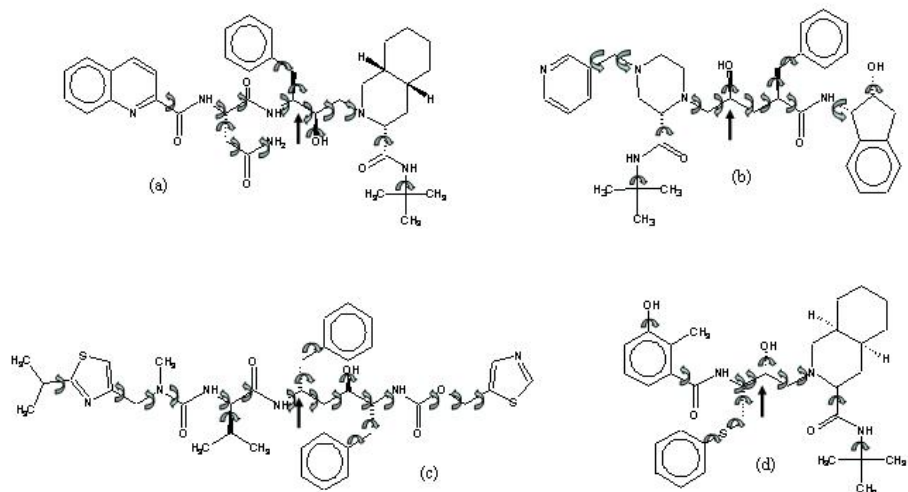| Ligand	Torsions	Dimension	PDB ID
NELFINAVIR	12	19	1ohr
INDINAVIR	14	21	1hsg
SAQUINAVIR	15	22	1hxb
RITONAVIR	20	27	1hxw

Fig. 1. Structural formulæ of HIV1 protease ligands and dihedral angles considered: (a) Saquinavir; (b) Indinavir; (c) Ritonavir; (d) Nelfinavir. Arrow: reference atom

in each direction, and a spacing of 0.25 Å. We are interested in the performance of the SSGA in identifying the experimental binding mode of the ligand molecule in the protein active site. To each ligand 30 independent runs were performed. The CPU time for each SSGA run varied from 10 to 13 minutes on a 2.0 GHz Pentium 4 with 256 MB of RAM. The algorithm success is measured by the RMSD between the crystallographic structure (from the corresponding PDB file) and the structure found by the algorithm. A structure with a RMSD ≤ 2.0 Å is classified as docked and that is considered a good result. A structure with a RMSD ≤ 2.5 Å is classified as partially docked, but for large ligands, with more than 15 dihedral angles, that is still considered a good result. The success ratio is the number of structures found with RMSD ≤ 2.0 Å in 30 runs. For the three selection-insertion schemes tested, we used a population of 1000 individuals, 1.000.000 energy evaluations, and probability of 0.15 for two-point crossover, 0.15 for SBX crossover, and 0.7 for non-uniform mutation. The flexible ligand docking results using linear ranking selection are shown in Table 2. We tested the standard RTS using a tournament size $w = 500$ and $w = 1000$. The docking results using standard RTS are shown in Table 3. The modified RTS was tested with a tournament size $w1 = w2 = 50\%$ and $w1 = w2 = 100\%$. The docking results using the modified RTS with 50% and 100% tournament sizes are shown in Table 4.

Table 2. Docking results using linear rank selection

Ligand	Lowest Energy[a]	Mean Energy[a]	Mean RMSD (Å)	Success Ratio (%)
NELFINAVIR	−57.53	−2.43	5.776	6.6
INDINAVIR	−62.97	32.96	6.049	3.3
SAQUINAVIR	−65.12	−25.17	4.764	13.3
RITONAVIR	−87.69	−7.99	5.305	10.0

[a] kcal/mol

Table 3. Docking results using standard RTS and two tournament sizes (w)

Ligand	w	Lowest Energy[a]	Mean Energy[a]	Mean RMSD (Å)	Success Ratio (%)
NELFINAVIR	500	−58.13	−52.98	1.101	83.3
	1000	−58.12	−53.02	1.427	73.3
INDINAVIR	500	−62.78	−51.16	2.573	63.3
	1000	−62.87	−51.78	1.765	76.7
SAQUINAVIR	500	−65.75	−58.26	1.225	83.3
	1000	−65.67	−62.11	0.726	86.7
RITONAVIR	500	−107.27	−84.26	2.583	50.0
	1000	−105.69	−71.91	3.137	6.7

[a] kcal/mol

Table 4. Docking results using the modified RTS and two tournament sizes (w)

Ligand	w	Lowest Energy[a]	Mean Energy[a]	Mean RMSD (Å)	Success Ratio (%)
NELFINAVIR	50%	−58.12	−50.97	2.009	63.3
	100%	−58.12	−55.55	0.394	93.3
INDINAVIR	50%	−62.87	−54.13	2.147	73.3
	100%	−62.79	−53.58	2.204	76.7
SAQUINAVIR	50%	−65.69	−61.64	1.031	90.0
	100%	−65.72	−61.44	0.824	93.3
RITONAVIR	50%	−107.11	−88.21	2.291	50.0
	100%	−107.34	−89.56	2.167	53.3

[a] kcal/mol

3.1   RMSD Analysis

In the results shown in Tables 3 and 4, the success ratio was the number of times that the lowest energy structure found by the algorithm corresponds to the respective crystallographic structure (RMSD $\leq$ 2.0 Å). In many cases, we found structures with a higher energy, but with a lower RMSD than the best

(minimum energy) solution. The results shown in the RMSD analysis use the same final population employed in the previous energy analysis. In the RMSD analysis the best solution in each run is the solution with lowest RMSD relative to the experimental structure, and not the structure with the lowest energy, as done in the previous energy analysis. The RMSD analysis of docking results using the standard RTS are shown in Table 5. The RMSD analysis of docking results using the modified RTS are shown in Table 6.

Table 5. RMSD analysis of docking results using the standard RTS and two tournament sizes (w)

Ligand	w	Lowest Energy[a]	Mean Energy[a]	Mean RMSD (Å)	SR[b] (%) ≤ 2.0 Å	SR[b] (%) (2.0,2.5] Å
NELFINAVIR	500	−58.10	−45.85	0.693	90.0	0.0
	1000	−58.09	−48.68	0.657	86.7	6.7
INDINAVIR	500	−62.61	−10.99	1.103	83.3	10.0
	1000	−62.24	−41.63	0.734	93.3	6.7
SAQUINAVIR	500	−65.28	−50.22	0.649	100.0	0.0
	1000	−65.24	−56.86	0.483	100.0	0.0
RITONAVIR	500	−106.95	−54.35	1.447	80.0	17.0
	1000	−102.29	−33.88	1.809	66.7	23.3

[a] kcal/mol
[b] Success Ratio

Table 6. RMSD analysis of docking results using the modified RTS and two tournament sizes (w)

Ligand	w	Lowest Energy[a]	Mean Energy[a]	Mean RMSD (Å)	SR[b] (%) ≤ 2.0 Å	SR[b] (%) (2.0,2.5] Å
NELFINAVIR	50%	−58.11	−49.57	1.126	73.3	6.7
	100%	−58.08	−54.96	0.296	96.7	0.0
INDINAVIR	50%	−62.71	−50.83	1.111	83.3	0.0
	100%	−62.42	−47.48	0.840	90.0	3.3
SAQUINAVIR	50%	−65.32	−60.31	0.877	93.3	3.3
	100%	−65.28	−58.45	0.574	96.7	3.3
RITONAVIR	50%	−106.84	−83.66	1.530	73.3	10.0
	100%	−106.71	−81.48	1.265	76.7	20.0

[a] kcal/mol
[b] Success Ratio

4   Discussion

The results show that with the implementation of the RTS technique we obtained a substantial increase in the algorithm performance. Comparing the results obtained using the linear rank selection (LRS) and using the standard RTS (energy analyses), we found that the mean success ratio increased from 8.4% (Table 2) to 70.1% (tournament size of 500 individuals, Table 3) and to 60.9% (tournament size of 1000 individuals, Table 3). These results indicate the standard RTS technique as a promising methodology for flexible ligand docking problems. The modified RTS (with an insertion criterion based on the RMSD between the ligand conformations) also shows a good performance to find structures close to the experimental structure with the lowest energy. The success ratio obtained using the modified RTS (tournament size of 50%, Table 4) ranges from 50% to 90% with a mean success ratio of 69.2%, and (using a tournament size of 100%, Table 4) from 53% to 100% with a a mean success ratio of 79.2%. Using the standard RTS the results show that a smaller tournament size produces a better result, while in the modified RTS the use of a tournament size of 100% showed to be the best choice. The modified RTS shows a slightly better performance than the standard RTS to find solutions closer to the experimental one and with better mean RMSD.

Considering all solutions in the final population (see RMSD analysis section), for both standard and modified RTS, we observe an increase in the success ratio regarding the experimental structure. The mean success ratio (including all ligands) obtained using the standard RTS are 88.4% and 86.7%, using a tournament size of 500 individuals (Table 5) and a tournament size of 1000 individuals (Table 5), respectively. Using the modified RTS, the mean success ratios obtained are 80.9% and 90.0%, using a tournament size of 50 % (Table 6) and a tournament size of 100% (Table 6), respectively. Once more the best results were obtained using a tournament size of 500 individuals for the standard RTS, and a tournament size of 100% for the modified RTS. Analyzing the results shown in Tables 5 and 6 we observe that RITONAVIR (the largest and most flexible ligand tested) shows the greater increase in performance when we consider the RMSD $\leq$ 2.5 Å criterion for computing the success ratio. In fact, for larger and highly flexible ligand molecules a RMSD $\leq$ 2.5 Å from the experimental structure can be considered a good result. Applying this criterion for all ligands, we observe that the mean success ratios are 95% and 96.7% for the standard RTS (tournament size of 50%, Table 5) and for the modified RTS (tournament size of 100%, Table 6) respectively. For all ligands a success ratio greater than 90% was obtained.

The results obtained in this work showed that the implementation of a multisolution RTS technique can be a very valuable approach in the flexible docking problem when dealing with highly flexible ligand molecules which are usually associated with a very complex energy hypersurface. Moreover, there are important advantages in applying a multisolution strategy in this type of problem. Usually the ligand-docking strategies approximate the real problem in the following points: (i) the absence of important factors associated with the

ligand-receptor energy function (*e.g.*, entropic and solvatation effects); (ii) the absence of explicit water molecules which can intermediate ligand-receptor hydrogen bonds and; (iii) the receptor is usually considered rigid or partially rigid. Following a multisolution docking strategy, the best distinct solutions can be used as starting points in more sophisticated and computationally expensive strategies (*e.g*, explicit solvent molecular dynamics simulations). Secondly, in real world drug design research projects, the ligand to be docked is only a drug prototype which will be probably modified several times in order to account for several chemical and pharmacological properties (*e.g*, toxicity, metabolic stability, synthetic tractability, etc.). In this sense, finding and analyzing several ligand-receptor binding modes can increase the possibilities of successful improvements in a drug prototype molecule. A more specific analysis considering the relation between the final population diversity (number and quality of solutions) and the SSGA/RTS (standard and modified) docking parameters is under progress and will be reported elsewhere.

References

1. Gane P.J., Dean P.M.: Recent Advances in Structure-Based Rational Drug Design. Current Opinion in Structural Biology **10** (2000) 401–404
2. Marrone T.J., Briggs J.M., McCammon J.A.: Structure-based Drug Design. Annu Rev Pharmacol Toxicol **37** (1997) 71–90
3. Wang R., Lu Y., Wang S.: Comparative Evaluation of 11 Scoring Functions for Molecular Docking. J. Med. Chem. **46** (2003) 2287–2303
4. Diller D.J., Verlinde C.L.M.J.: A Critical Evaluation of Several Global Optimization Algorithms for the Purpose of Molecular Docking. J Comp Chem 20 **16** (1999) 1740–1751
5. McConkey B.J., Sobolev V., Edelman M.: The Performance of Current Methods in Ligand-Protein Docking. Current Science 83 **7** (2002) 845–855
6. Brooijmans N., Kuntz I.D.: Molecular Recognition and Docking Algorithms. Annu. Rev. Biophys. Biomol. Struct. **32** (2003) 335–373
7. Ewing T.J.A., Kuntz I.D.: Critical Evaluation of Search Algorithms for Automated Molecular Docking and Database Screening. J Comp Chem **18** (1997) 1175–1189
8. Jones G., Willett P., Glen R.C., Leach A.R., Taylor R.: Development and Validation of a Genetic Algorithm for Flexible Docking. J Mol Biol **267** (1997) 727–748
9. Morris G.M., Goodsell D.S., Halliday R.S., Huey R., Hart W.E., Belew R.K., Olson A.J.: Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. J. Comp. Chem. 19 **14** (1998) 1639–1662
10. Carlson H.A., McCammon J.A.: Accommodating Protein Flexibility in Computational Drug Design. Molecular Pharmacology **57** (2000) 213–218

11. Wong C.F., McCammon J.A.: Protein Flexibility and Computer-Aided Drug Design. Annu. Rev. Pharmacol. Toxicol. **43** (2003) 31–45
12. Osterberg F., Morris G.M., Sanner M.F., Olson A.J., and Goodsell D.S.: Automated Docking to Multiple Target Structures: Incorporation of Protein Mobility and Structural Water Heterogeneity in AutoDock. PROTEINS: Structure, Function and Genetics **46** (2002) 34–40
13. Hart W.E., Rosin C., Belew R.K., Morris G.M.: Improved Evolutionary Hybrids for Flexible Ligand Docking in AutoDock. Optimization in Computational Chemistry and Molecular Biology. (eds.) C.A. Floudas & P.M. Pardalos (2000) 209–229
14. Thomsen R.: Flexible Ligand Docking Using Evolutionary Algorithms: Investigating the Effects of Variation Operators and Local Search Hybrids. BioSystems **72** (2003) 57–73
15. Whitley D.: The GENITOR Algorithm and Selective Pressure. Proc. of the Third Int. Conf. on Genetic Algorithms and their Applications. (eds) J.D. Schaffer, Morgan Kaufmann, San Mateo, CA (1989)
16. van Gunsteren W.F., Berendsen H.J.C.: Groningen Molecular Simulation (GROMOS) Library Manual. Biomos, Groningen (1987)
17. Harik G.R.: Finding Multimodal Solutions Using Restricted Tournament Selection. Proc. of the Sixth Int. Conf. on Genetic Algorithms. (eds.):Larry Eshelman, Morgan Kaufmann, San Francisco, CA (1995) 24–31
18. Holland J.H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, MI (1975)
19. Goldberg D.E.: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, New York (1989)
20. Maillot P.G.: In Graphics Gems. A. S. Glassner, (ed.): Academic Press, London, (1990) 498
21. Pascutti P.G., Mundim K.C., Ito A.S., Bisch P.M. Polarization Effects on Peptide Conformation at Water-membrane Interface by Molecular Dynamics Simulation. J. Comp. Chem. 20 **9** (1999) 971–982
22. Arora, N., Jayaram B.: Strength of Hydrogen Bonds in α-helices. J. Comp. Chem. 18 **9** (1997) 1245–1252
23. Deb K., Beyer H.: Self-adaptive Genetic Algorithms with Simulated Binary Crossover. Technical report no. CI-61/99, University of Dortmund, Department of Computer Science/LS11 (1999)
24. Michalewicz Z.: Genetic Algorithms + Data Structures = Evolution Programs. Springer-Verlag, New York (1992)
25. Sareni B., Krähenbühl" L.: Fitness Sharing and Niching Methods Revisited. IEEE Transactions on Evolutionary Computation 2 **3** (1998) 97–106
26. Goldberg D., Richardson J.: Genetic Algorithms with Sharing for Multimodal Function Optimization. Proc. of the Second Int. Conf. on Genetic Algorithms (eds.): J. J. Grefenstette, Lawrence Erlbaum Associates, Hillsdale, New Jersey (1987) 41–49
27. De Jong K.: Analysis of Behavior of a Class of Genetic Adaptive Systems. PhD Thesis, The University of Michigan (1975)
28. Mahfoud S.W.: Crowding and Preselection Revisited, Parallel Problem Solving from Nature 2. (eds.): Reinhard Männer and Bernard Manderick, North-Holland, Amsterdam (1992) 27–36