Computers in Biology and Medicine 40 (2010) 215–222

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Energy minimum theorem based on AGA, Lyapunov and force field for CADD techniques

Wen-Tsai Sung *, Yu-Feng Liu

Department of Electrical Engineering, National Chin-Yi University of Technology, No. 35, Lane 215, Section 1, Chung-Shan Road, Taiping City, Taichung County, 411 Taiwan

ARTICLE INFO

Article history:
Received 19 November 2008
Accepted 29 November 2009

Keywords:
Computer aided drug design
Minimum energy
Protein folding
Adaptive GA
Lyapunov

ABSTRACT

This paper employs the energy minimum to enhance drug docking efficiency in a computer aided drug design (CADD) system. The energy minimum application is used to enhance CADD docking performance. The proposed method is discussed in three aspects, adaptive genetic algorithms (AGA), Lyapunov stability theorem and molecular force field. As in previous researches, docking is the crucial component in drug development. The number of docking sites affects the drug docking speed. Reducing the scope of the geometric search is the key task. This paper proposes AGA to improve geometric molecular docking search efficiency. The Lyapunov stability theorem forwards the stability state identification. Protein folding intention generally finds the most appropriate stability state when the thermodynamic and molecular mechanical free energy has reached the lowest point. The AMBER force field simulation is used to discover the molecular statistical mechanics in a drug-ligand. AGA was found better in terms of processing the geometric graphic search operation. The AGA and Lyapunov algorithms were utilized to sieve out the global energy minimum approach from the numerous, raw docking sites.

1. Introduction

Virtual screening is a significant step before any new drug can established. The objective of this work is to improve virtual screening using computer technologies. Experiments on existing drugs and compounds are conducted. The simulation results are compared with previously published studies [1,2]. Because a high throughput screening is required and the number of ligand types in the compound library is large, CADD technologies are necessary to improve drug design. Maggio et al. performed CADD simulations on 10 targets [3,4]. The hit ratio was 2%–24% higher than that from the conventional strategy hit ratio, which was 0.01%–0.001%. In silico virtual screening, CADD is divided into direct (receptor-based) and indirect (ligand-based, quantitative structure-activity relationship (QSAR), pharmacophore) drug design depending on whether the target structure is known or unknown. Owing to rapid advances in structural biology and computer technology, structure-based CADD using docking techniques, virtual screening and library design, along with the target structure focusing combinatorial chemistry, have become powerful tools in the multi-step drug discovery process. This investigation adopts the energy minimum theorem to explore CADD and accelerate the molecular docking process. This solution is categorized as a direct method via docking virtual screening and in silico de Novo Ligand design. This study first employs an improved AGA to optimize the geometric search for many receptor binding sites and discover docking sites near the global energy minimum. The docking scope is restricted because binding sites are judged according to their stability in terms of Lyapunov heavy loading. The improved AGA is utilized to narrow the scope of the geometric search. The Lyapunov rule is used to determine docking sites that are stable and approach the global energy minimum. This study successfully solves the docking problem with various computer graphic technologies, system controls and bioinformatics. Computer graphics is a vital technology in modern medicinal chemistry. Previous published studies apply the WebDeGrator system to establish molecular computer modeling for the docking process [5–7].

2. Adaptive GA

The genetic algorithm is a representative of a class of methods based on heuristic random search techniques. It was proposed by John H. Holland in the early seventies and has since found application in a number of practical problems [8]. The genetic algorithm requires a considerable amount of computational time. The adaptive GA (AGA) is an attempt to speed up the algorithm via visual docking sites. In this study, the multi-threading technique is recognized as an efficient tool for transforming the

* Corresponding author. Tel.: +886 4 23924505; fax: +886 4 22466927.
E-mail address: songchen@ncut.edu.tw (W.-T. Sung).

0010-4825/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2009.11.018

genetic algorithm into a non-random form in initial populations. The key to obtaining high performance in geometric search computing is to reduce the protein cave search in the energy minimum states. According to our previous searches, we succeeded in using the Intuitive GA (IGA) to enhance docking process efficiency [1]. The strength of the IGA lies in its' ability to locate the global optimum in a multi-modal surrounding. Unfortunately, regardless how robust and efficient a genetic algorithm, the solution it provides always bears a certain measure of unreliability. The IGA can locate the global optimum with only a certain probability of success. Considerable attention has been paid to efforts to increase that probability.

Many attempts to improve the search performance of genetic algorithms have been made since the GA's first appearance in 1975. These methods or mechanisms include non-binary coding, fitness scaling, elitist strategy, extinction and immigration strategy. Various alternative ways for reproduction, crossover and mutation have greatly improved the performance of genetic algorithms, enhancing the convergence speed to prevent premature convergence.

Crossover and mutation play important roles in making genetic algorithms a powerful search technique, but improper selection of crossover and mutation rates may lead to premature convergence and local optimal solutions. Even if these parameter settings are optimal in the earlier search stages, they will often become inefficient in the later stages. Hence, adaptive probabilities for crossover and mutation were proposed to overcome these drawbacks. In Adaptive genetic algorithms (AGA), the adaptive crossover and mutation rates vary depending on the fitness of each chromosome and do not need to be specified before running the AGA. The AGA has been proven better than the traditional GA in most cases. The AGA can converge to global optimum in far fewer generations than the traditional GA [9].

The AGA is good at the optimizing function with many local optimal points and has no restriction on the form of the fitness functions. The AGA is population oriented, stores a sample replica of the function profile being optimized and provides important clues about the global structure of that function. The AGA is parallel and also global. The fitness information calculated using the fitness function from different members relies on various genetic operators, especially through selection, crossover and mutation mechanisms. In the evaluation step, the users perform decisions using the physical condition in the protein conformation. The AGA can be implemented using several threads. The main benefits that arise from multi-threading are: better program structure and efficient use of multiple processors [10] (Fig. 1).

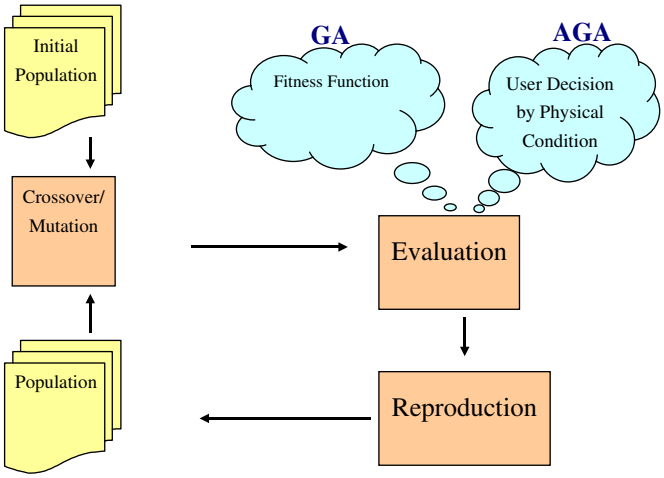

Fig. 1. The Difference between GA and AGA [11].

3. Employed AGA and Lyapunov theorem for system practice

Conventional or well-known modern GA all combine random computation with a random selection process, such as population and mutation selection processes. Therefore, traditional GA often generate incorrect results, requiring verification by selecting the average value of the time calculation product. In mutation and crossbreeding, the computation method is non-directional, resulting in an unstable product that reaches equilibrium slowly. The classical GA is frequently trapped in local optimal solutions. Another good thing obtained from adapting the genetic algorithm is that we can bypass the task of defining the parameter values, which is in most cases is left to the user. Those values are known to significantly affect the algorithm's performance. Poorly chosen parameters can cause the algorithm to not produce any relevant solutions at all. Moreover, the optimal parameter configuration is often problem dependent. This can create difficulty for the inexperienced GA user.

The AGA is therefore proposed to enhance the efficiency of the conventional GA by combining it with the uniform design concept and numerical analysis. The AGA and classical GA are different in due to the hybridization and mutation algorithms. The AGA employs all of the population in a complete group to calculate the result, whereas the traditional GA only employs a portion of the population via one possibility. The AGA's mutation is regulative [12]. A promising solution to this challenge is to employ the Lyapunov function to the docking that observes the binding site points. Through the direct Lyapunov method, the global minimum energy site is extracted from various binding sites. This approach is quite reasonable in a drug molecule in which the molecule's motion is derived from the applied energy and control theory. An alternative is to solve the eigenvalue λ of the drug docking dynamic system and determine whether λ is less than the convergence value ε from the initial state to an infinitely long time. If the docking system conforms to this condition, the system is then stable. If the convergence rate of the eigenvalue λ is directly proportional to the Lyapunov exponential function, the docking system is then Lyapunov asymptotically stable. This scheme is termed the "indirect Lyapunov method" [13].

In the Lyapunov stability theorem, the stability concept of differential equation is described first.

$$\frac{dXi}{dt} = f_i(X_1, X_2, \ldots, X_n) \quad (i = 1, 2, \ldots, n) \tag{1}$$

Suppose that in the initial condition $X_i(t_0) = X_i^0$, Eq. (1) has a solution $X_i(t)$. When the initial condition yields a small perturbation motion η_i, the initial condition then becomes $X_i'(t_0) = X_i^0 + \eta_i$ and Eq. (1) has a new result $X_i(t, \{\eta_i\})$. The stability definition follows.

The solution $Xi(t)$ of system equation (1) is stable if for each positive $\varepsilon > 0$ a positive number $\delta > 0$ exists such that the condition $|\eta_i| \le \delta$ is satisfied, for all $t \ge t_0$ invariably:

$$|X_i(t, \{\eta_i\}) - X_i(t)| < \varepsilon \quad (i = 1, 2, \ldots, n) \tag{2}$$

Conversely, the system is unstable under when this condition does not correspond.

If $Xi(t)$ is stable and the following is satisfied:

$$\lim_{t \to \infty} |X_i(t, \{\eta_i\}) - X_i(t)| = 0 \quad (i = 1, 2, \ldots, n) \tag{3}$$

The function $X_i(t)$ is termed "asymptotic stability" or "Lyapunov stability". The above system stability definition can also be described simply: If the initial condition is $X(t_0) = X_0$, then the system solution is $X(t)$. When the initial condition is a finite change (δ), the system solution equation is still sited in the finite field (ε) with the solution $X(t)$ of original initial condition, meaning that $X(t)$ is stable. When the initial condition indicates

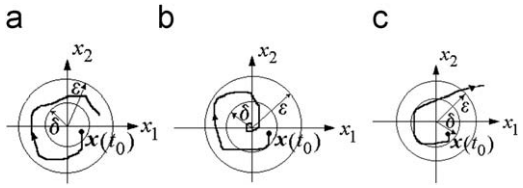

Fig. 2. Three equilibrium state and representative trajectory ($|X-X_i|<\varepsilon,\forall t>0$): (a) marginally stable, (b) asymptotically stable and (c) unstable.

a finite change, the system solution is near $X(t)$ and the solution comes back to $X(t)$ after a long time, and $X(t)$ is asymptotically stable [14] (Fig. 2).

4. AGA approach based on crossover and mutation of adaptive probabilities

An AGA with diversity-guided mutation, which combines the adaptive probabilities of crossover and mutation as proposed by Spears [15] Using homogeneous finite Markov chains, it is proven that an AGA with diversity-guided mutation and a genetic algorithm with diversity-guided mutation converge to the global optimum if they maintain the best solutions. The convergence of adaptive genetic algorithms with adaptive probabilities for crossover and mutation is studied [16].

As mentioned above, the crossover rate *pc* and mutation rate *pm* selection determines the AGA performance. In the conventional crossover and mutation mechanism, *pc* and *pm* are held as constants. Srinivas proposed to adaptively adjust the *pc*, *pm* and fitness value for an effective AGA process [17]. While a higher crossover rate can increase the generation of new individuals in the search process, it may also increase the probability for destroying superior individuals, hence leading to local optima. In order to prevent converging to local optima, the mutation rate cannot be too small, but a higher mutation rate may result in a random geometric search. Therefore, this study adaptively changes the crossover and mutation rate according to the fitness of each chromosome and the current evolution situation for approaching the global minimum energy position in binding site.

When the increase in maximum fitness in each generation slows down and the best individual is still far from excellent, the proposed approach increases the crossover and mutation rate to accelerate the convergence speed and prevent local optima. From past research, when the evolution is slack, the difference between maximum fitness and average fitness in a generation becomes small and almost held as a constant. Thus, the adaptive crossover and mutation rate can be defined according to this property as follows:

$$P_C = \begin{cases} \dfrac{k_1(f_{\max}-f')}{f_{\max}-\bar{f}}, & f' \geq \bar{f} \\ k_3, & f' \geq \bar{f} \end{cases} \tag{4}$$

$$P_m = \begin{cases} \dfrac{k_2(f_{\max}-f)}{f_{\max}-\bar{f}}, & f \geq \bar{f} \\ k_4, & f < \bar{f} \end{cases} \tag{5}$$

where $f_{\max}$ is the maximum fitness in the population, f is the average fitness of the population, f' is the greater fitness among parents in the crossover process, f is the fitness of the individual in the mutation process, k_1 and k_3 are crossover probabilities adjust parameters, k_2 and k_4 are mutation probabilities adjust parameters.

In Eqs. (4) and (5), the parents or individuals with higher fitness have a higher crossover or mutation rate. When the evolution slows down, $(f_{\max}-f)$ will become relatively small, hence the equations will increase the crossover and mutation rate to accelerate the convergence speed and avoid local optima. Furthermore, p_c (p_m) will be zero when $f'=f_{\max}$, thus the individual with the highest fitness will not be destroyed through the crossover and mutation process. When $f'<f$, reset $p_c=k_1$ to prevent *pc* or *pm* higher than 1.In this paper, the $k_1=k_3=1.0$, $k_2=k_4=0.5$ would get the best solution.

The adaptive crossover and mutation rate in equations (4) and (5) can adjust *pc* and *pm* according to the fitness of the chromosome, and when the evolution slows down, p_c and p_m will increase to accelerate the convergence speed and avoid local optima. The proposed approach has better performance than the traditional GA which keeps p_c and p_m as constants.

5. The operation and strategy in AGA

When two parents are identical, the offspring produced by them through crossover will be still the same. When the chromosomes gradually become alike or even the same after many generations of evolution, the crossover operation will gradually lose its ability to generate new chromosomes, and the evolution will stagnate. If all of the chromosomes in the population are the same, the crossover operation is then useless and the only way left to generate better chromosomes is mutation. However, mutation is not efficient enough to overcome the stagnate situation, especially for long chromosomes. Therefore, the extinction and immigration strategy has been proposed to address this difficulty.

In Fig. 3, the excellent strategy operates before the main operations of genetic algorithms, i.e. reproduction, crossover, and mutation operation, in every generation.

This process keeps the best *p*% individuals and takes them directly into the next generation without joining those operations. The new generation is formed with the surviving individuals and the new generated offspring, thus good performing individuals

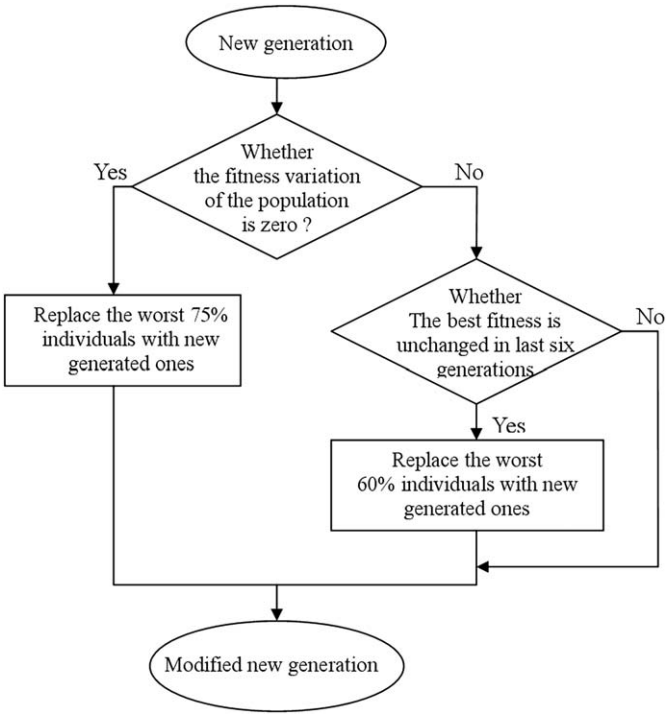

Fig. 3. The extinction and immigration strategy for this study.

will not be destroyed by the evolution process. This excellent strategy ensures that the maximum fitness will increase continuously in each generation.

Extinction eliminates a certain number of the worst chromosomes and new chromosomes are generated randomly to fill out the population. The role of extinction and immigration is similar to mutation, while it has a more dramatic effect. Extinction and immigration is applied when all chromosomes in the population are the same or the maximum fitness values remain unchanged over a certain number generations.

6. The AGA flow chart

Combining the AGA operators and strategies mentioned above with the basic structure of simple genetic algorithms, this study obtained an AGA with an excellent extinction and immigration strategy. According to the above discussions, this paper employed the operators and strategies in our study as shown in Fig. 4. After each adjustment this study makes sure that the crossover and mutation operations have the chance to work continuously. For this reason, this paper employed minimum crossover and mutation rates of 0.03. If the crossover or mutation rate is ≤ 0.005, the adjustment operation stops decreasing the probability. The refined genetic algorithm incorporating this adaptation approach is described in Fig. 4.

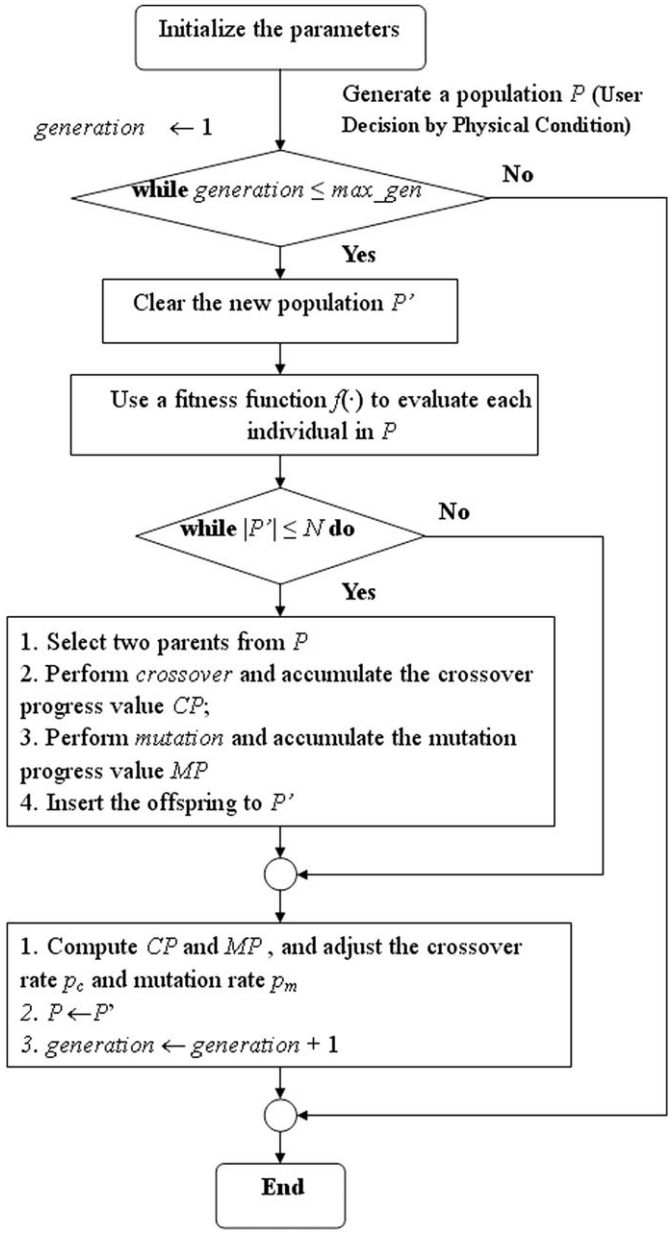

Fig. 4. Flow chart of AGA.

7. Docking accuracy and efficiency

The docking accuracy, ranking accuracy and algorithm efficiency of the docking algorithms were evaluated. A solution pose was treated as accurate if the root–mean–square–distance (RMSD) of the pose and the reference (experimentally determined) pose was below 2 Å. The docking accuracy requires one of the top 30 scored poses to be accurate, while the ranking accuracy requires the top scored pose to be accurate. For a docking algorithm to be valuable, it needs to be efficient and achieve high docking accuracy. A comparative study of eight popular docking programs was recently undertaken on a 100-complex benchmark [18].

Besides the search algorithm, the scoring function is also critical to the accuracy of the docking algorithm. The ideal scoring function would determine the binding affinities between the ligand and the receptor, which include factors such as the Van der Waals interaction, H-bonding, hydrophobicity and electrostatics. However, the degree of contribution of each factor to the scoring is well understood. Three main approaches exist for studying the scoring function, namely force-field-based, empirical-based, and knowledge-based methods. The force-field-based method approximates the score with non-bond energy terms from the well-studied force field, such as AMBER or CHARMM. The Empirical-based system applies a set of protein–ligand complexes with experimentally determined binding affinities to train the parameters in the scoring function. The knowledge-based scheme adopts the Boltzmann hypothesis with known structural database to compute the score. However, despite extensive research, no scoring function comes close to the ideal, and consensus-scoring functions are sometimes used. Significantly, most docking algorithms are capable of handling different (additive) scoring functions by interpolating the score through a grid, which can be precomputed and stored. An accurate and efficient protein-small-molecule docking algorithm is fundamental to structure-based drug design. During the drug design process, once the 3D structure of a target protein (which is believed to be responsible for the disease) is determined, new drug candidates can be identified by virtually screening a database of known compounds through small-molecule docking.

In general, the docking problem has two components: a scoring or evaluation function that can accurately discriminate (i.e., experimentally observed) docking solutions from incorrect solutions, and a search algorithm that searches the configurational and conformational space for the candidate poses measured with the scoring function. The docking problem is challenging because the scoring function, which measures the binding affinity between ligand and receptor, is not completely understood, and the search space is high-dimensional-besides six degrees of freedom (DOF) in the configuration space—both molecules are flexible and might undergo conformational changes upon binding, resulting in hundreds or thousands of DOFs in the conformational space. Performing exhaustive conformational searches during docking is thus computationally infeasible. The current most commonly utilized approach in modeling protein–small-molecule docking is to address only the conformational space of the ligand, assuming that the protein receptor is rigid. This approach is

known as the rigid-receptor–flexible-ligand docking model. This study employed the AGA based on energy minimum theorem to optimize the docking process, and adopted the Lyapunov model to lower the number of docking sites and thus shorten the docking time.

Many docking algorithms have been presented in the last 20 years. These algorithms can broadly be classified into three categories, namely stochastic search, incremental construction, and multi-conformer docking algorithms. Stochastic search algorithms include AutoDock, ICM and GOLD [19]. These algorithms are based on genetic algorithms and/or Monte Carlo-simulated annealing. The incremental construction algorithms first dissect each molecule into a set of rigid fragments based on rotatable bonds, and then incrementally assemble the fragments around the binding pocket. Some examples of this class are DOCK, FlexX and Surflex [20]. Unlike the incremental construction, multi-conformer docking algorithms separately generate a set of low-energy conformers, and then perform rigid docking for every conformer. Multi-conformer docking algorithms include FLOG and FRED (from OpenEye Scientific Software). A brief description of FRED can be found in [21]. The most efficient existing algorithms are FRED, DOCK and FlexX.

8. AGA computer simulation

Example: An integrated simulation in HIV protease inhibitors docking with HEX Software (1mt8)

As described in the above chapter, the Human Immunodeficiency Virus (HIV) protease is responsible for cleaving the polypeptide expressed by viral DNA after an HIV infection. Since this process is a step in the maturation of a virus, a reasonable drug design strategy would involve the discovery of a compound that irreversibly binds with the protease to eliminate its functionality. Such a drug would not cure the disease, since the offending DNA is still operative in the infected cells, but it does decrease the spread of the virions generated by the infected cell. This strategy can be further investigated because the protease is designed to bind with a particular peptide sequence, a drug could be designed that "looks like" that peptide sequence but does not behave like it. This strategy is "bait and trap": the protease tries but fails to cleave the drug. Furthermore, the drug binds so tightly with the protease that it cannot be released. In designing such a drug, the medicinal chemist could employ wet lab techniques to generate several compounds that are all drug candidates. They would then test these compounds in an assay experiment with the purpose of evaluating the affinity of the compounds for the protease. However, this process is expensive and should be undertaken only for compounds that exhibit some predicted high level of affinity. Bioinformatics is helpful in this case. Software can be utilized to evaluate hundreds or even thousands of "virtual compounds" stored in a database. In this evaluation, the compounds are described in terms of data structures, such as a list of coordinates of the atoms in a molecule and the list of connections between the atoms. If the protease receptor site is known; i.e. a data structure representation for the receptor is available, then the following steps can be applied:

1. Adopt an algorithm that performs a docking of the candidate compound with the protease receptor.
2. Utilize another algorithm to predict the binding affinity for the receptor/compound association.
3. Sort the results to determine which compounds give the highest affinity to observe the consistent features of the most active compounds.
4. Employ this information to design a more focussed set of compounds and repeat these steps.

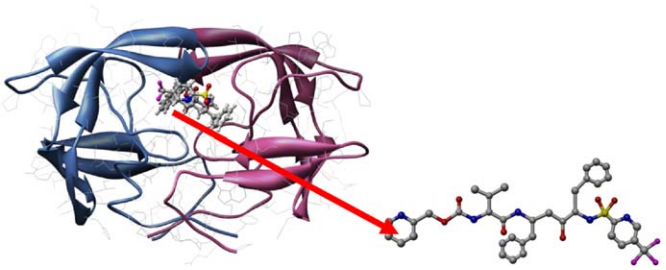

Fig. 5. Dock the drug (ACT: $C_2\ H_3\ O_2^-$) to into the 1mt8 HIV protease.

Inhibitors of this viral protease can be used to fight the HIV infection. Protease inhibitors interfere with continued infection by blocking the ability of the protease to cleave the viral polypeptide into functional enzymes. Mutations enable HIV to withstand treatments involving only one drug, resulting in the growing use of multiple-drug therapies combining a protease inhibitor with a reverse transcript inhibitor (Fig. 5).

In this example, the first step is selected 1mt8 HIV protease as the receptor. In the next step we applied AGA to the geometric search and found eight suitable ligands which were utilized for docking into the 1mt8 HIV protease. The eight-ligand chemical character was searched from the NIAID, NLM, NCBI, and NIST compound. Fig. 6 shows the DMP323 search result, and Table 1 presents the eight-ligand molecular structure.

The third step is that the ligand and receptor be regulated into experimental components. We adopted the Hex software to dock the drug ligand into the 1mt8 receptor. The docking operation steps are as follows:

(1) Open the candidate x.pdb file without Gag with the Browse button associated with the Receptor molecule.
(2) Open x.pdf file for the 3D version of the drug with the Browse button associated with the Ligand molecule.
(3) Set clustering RMSD to 2.0.
(4) Set complex type to: Protein–small ligand.
(5) Leave the optional text boxes empty.
(6) Finished.

In this step, we plug-in the AGA and Lyapunov programs to lower the number of binding site and repeat the geometric, energy and activity estimation steps. These steps have a recursive relationship. Table 2 indicates the different steps cause distinct results. Round 3 is closed to the final results which permit the benchmark.

In this example Fig. 7 indicates the process has been run for 386 generations with the program defaults. When the new generation is over 237, the best fitness is limited to 0.05, 38 scope, and the best next generation appeared. The fitness outcome is 7.41 after generating 32 generations, which is better than the classic GA in 189 generations. Table 3 presents that the AGA was the best approach in this case, since it yielded a short docking time 32% and the best fits in terms of the geometric calculation for searching the receptor binding sites.

The final step estimates whether the ligand receptor docking activity converges or diverges in λ_{max}, whether it is has affinity is an agonist to Kd,. Compared with the benchmark database does rmsd and energy have the lowest standard value, we can obtain the optimum drug candidate through these steps.

This example employed the following settings: energy (kcal/mol) and RMSD (Å); 210 runs, 536 individuals, 1.0×10^6 energy evaluations, two-point crossover (prob.=0.3) and non-uniform mutation (prob.=0.7); mean in 100 runs; percentage of conformations identified by the algorithm with RMSD < 0.5 Å; percentage of conformations found by the algorithm with

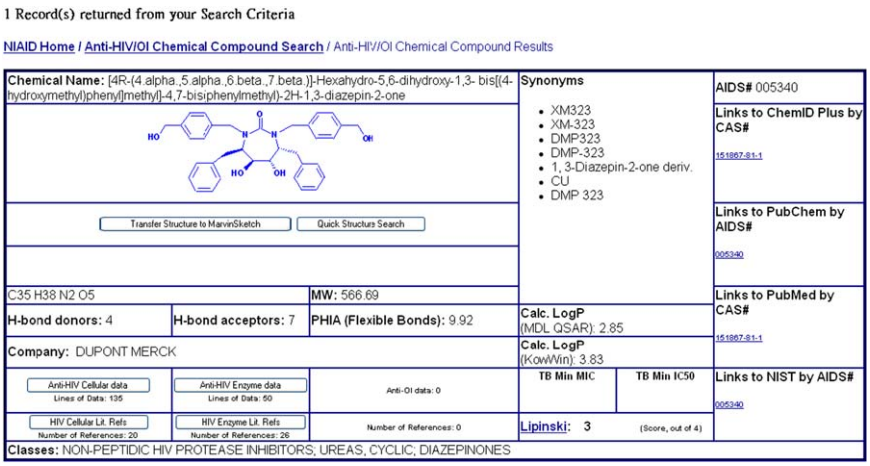

Fig. 6. Searching DMP323 and the molecular character from NIAID database.

Table 1
Molecular structure and AIDS number of the eight ligands.

Ligand	AIDS Number	Molecular Structure
Ritonavir	028478	
L-Chicoric Acid	029768	
Nelfinavir	028590	
Gallein	081066	
DMP323	005340	
Indinavir	005824	
Cefaclor	070030	
Saquinavir	000640	

RMSD < 2 Å. Standard deviations are given in parentheses. Table 4 presents the docking results.

9. Conclusions and future work

This investigation examined the geometry, energy and activity of cave structures for the receptor and binding sites of various ligands. AGA was utilized to increase the docking efficiency. The vital role of the minimum energy feature in protein folding and drug docking was then discussed. The Lyapunov theorem was employed to solve the energy problem. The drug–receptor interaction activity was observed and the protein or ligand characteristics were determined using various computer simulations [22].

The performances of the above AGA algorithms in optimizing several unimodal and multimodal functions were compared. The results show that for multi-modal functions the average AGA convergence generation with diversity guided mutation is about 63 less than that of the AGA with adaptive probabilities and a genetic algorithm with diversity-guided mutation. The AGA with diversity-guided mutation does not lead to premature convergence. It was also shown that better balance between overcoming premature convergence and quickening convergence speed can be obtained.

According to the drug–receptor interaction model, eight drug ligands were applied to dock the HIV-based protein enzymes. The 1mt8 HIV protease was used as the receptor for an Alzheimer's disease case study to perform docking. Example 1 (1mt8) depicted some computer simulation results. Some drug–receptor affinity and activity such as K_d, agonists or antagonists, hit ratio, free energy, RMSD were observed in this simulation. Based on the work in our above studies, docking efficiency was improved and production speed was increased with the AGA and Lyapunov algorithms according to energy minimum themes. This study presented a novel systemic method based on our previous work, "Study on Molecular Docking for Computer-Aided Drug Design via Lyapunov Equation and Minimum Energy". A serial literature survey and practice in computer simulation and computations contributed some significant themes in this investigation. Computer simulation example results were employed to propose and demonstrate some computation methods and theories. For instance, the AGA lowered the number of docking sites, shortened the docking time 32%, and enhanced the geometric graphic search operation after comparing four optimal geometric search methods, along with the Pegg et al., and Camila et al. methods [23]. The Lyapunov rule was adopted to determine the docking site stability. We believe that this investigation succeeded in integrating biology, information technology (IT), system engineering and chemistry into modern bioinformatics. Important research for CADD will continue in the future.

Conflict of interest statement

None declared.

Table 2
Indicated the different step caused the distinct results.

Ligand	Round						
	Round 1		Round 2		Round 3		Benchmark
	rmsd	Energy	rmsd	Energy	rmsd	Energy	Energy
Ritonavir	0.289	−106.21	0.219	−102.51	0.139	−101.11	−101.68
L-Chicoric acid	0.19	−82.9	0.12	−79.2	0.04	−77.8	−79
Nelfinavir	0.216	−89.2	0.146	−85.5	0.066	−84.1	−83.2
Gallein	0.193	−75.05	0.123	−71.35	0.043	−69.95	−70.1
DMP323	0.188	−68.48	0.118	−64.78	0.038	−63.38	−62.9
Indinavir	0.187	−58.14	0.117	−54.44	0.037	−53.04	−53.1
Cefaclor	0.202	−90.25	0.132	−86.55	0.052	−85.15	−86.2
Saquinavir	0.208	−93.38	0.138	−89.68	0.058	−88.28	−85.3

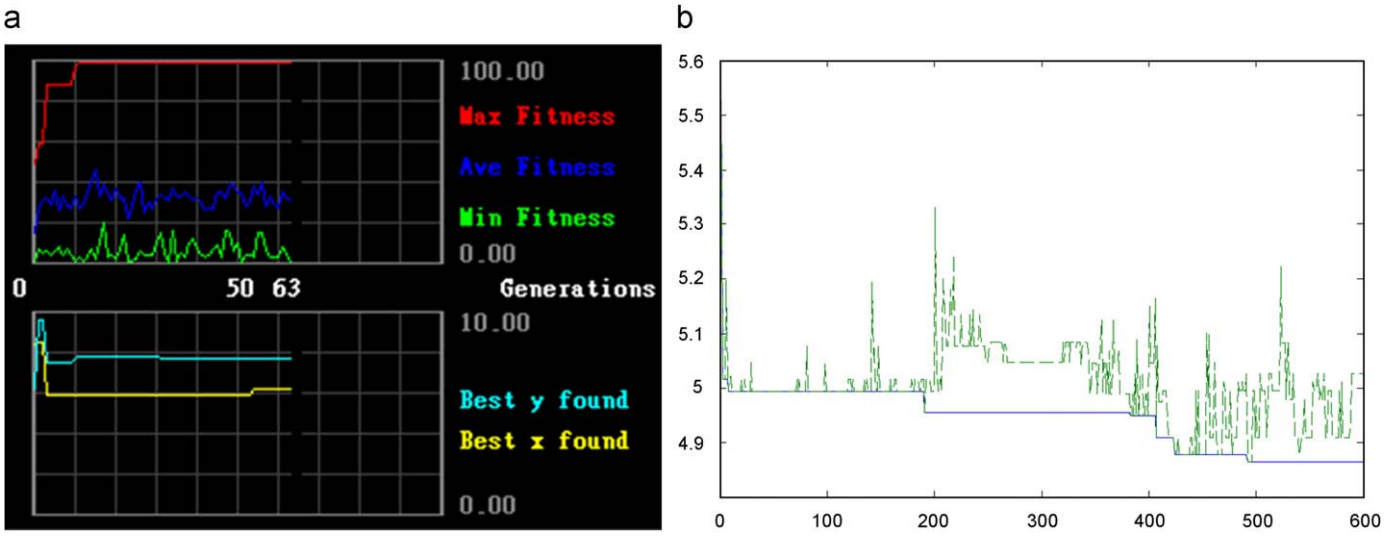

Fig. 7. The AGA and number database computation process.

Table 3
Three optimal docking methods in practice [24].

Optimal algorithms	**Basic feature**	**In this Experiment result** (1mt8 HIV protease)		
		Number of binding sites	**Docking time (h)**	**Average interaction energy for ligand (Kcal/mole)**
Traditional GA	The swift calculate, global minimum	285	0.87	−128.3
IGA	An improved GA with Intuition method	212	0.65	−144.0
AGA	An improved GA with Adaptive method	128	0.42	−96.6

Table 4
Docking results example. (1mt8 HIV protease).

Ligand	Lowest rmsd	Energy of lowest	Success ratio	Affinity (1/Kd)	With Lyapunov (%)
Ritonavir	0.09	−100.18	92	67	30
L-Chicoric Acid	0.04	−75.32	86	82	21.5
Nelfinavir	0.066	−82.18	73	74	15.6
Gallein	0.043	−68.91	89	91	18.3
DMP323	0.038	−58.37	72	54	23
Indinavir	0.037	−87.08	91	87	12
Cefaclor	0.052	−82.17	85	85	8
Saquinavir	0.058	−86.78	83	75	27

Acknowledgments

This research was supported by the National Science Council of Taiwan under Grant NSC 98-2221-E-426-007. The authors would like to thank the National Chin-Yi University of Technology, Taiwan for financially supporting this research.

References

[1] W.-T. Sung, Intuitive GA and Lyapunov to promote molecular docking efficiency for drug design, Journal of Medical and Biological Engineering 29 (1) (2009) 39–45.
[2] S.-C. Ou, C.-Y. Chung, W.-T. Sung, C.-C. Tsai, C.-C. Chien, D.-Y. Su, Virtual screening and computer-aided drug design in molecular docking via Lyapunov function, WSEAS Transactions on Biology and Biomedicine 1 (4) (2004) 384–390.
[3] S.C. Pegg, et al., A genetic algorithm for structure-based de novo design, Journal of Computer Aided Molecular Design 15 (2001) 911–933.
[4] S. Camila, et al., Agenetic algorithm for the ligand-protein docking problem, Genetics and Molecular Biology 27 (4) (2004) 605–610.
[5] S.-C. Ou, H.-Y. Chung, W.-T. Sung, C.-Y. Chung, Using directivity genetic algorithms to geometry search and conformational stability for improving molecular simulation techniques, WSEAS Transactions on Mathematics and Computers in Biology and Biomedicine 3 (4) (2006) 291–296.
[6] W.-T. Sung, Determine global energy minimum solution via Lyapunov stability theorem, International Journal of Innovative Computing, Information and Control (IJICIC) 5 (7) (2009) 1-08-22.
[7] S.-C. Ou, H.-Y. Chung, W.-T. Sung, Enhancing Compression and Encryption of Image with FPGA-Based, Multimedia Tools and Applications, vol. 28(1), Kluwer academic publishers, Dordrecht, January 2006.
[8] S.-C. Ou, W.-T. Sung, S.-J. Hsiao, Study on the wireless and virtual reality technology—based on WebDeGrator system, Journal of Internet Technology, vol. 3(4), Computer Center, National Dong Hwa University, 2002, pp. 257–266.
[9] S.-C. Ou, C.-Y. Chung, H.-Y. Chung, W.-T. Sung, C.-C. Cheng, Molecular docking for protein folding structure and drug-likeness prediction, WSEAS International Journal on Biology and Biomedicine 2 (1) (2005) 57–63.
[10] SunSoft, Solaris 2.4: Multithreaded Programming Guide, Sun Microsystems, Mountain View, California, 1994.
[11] H.S. Kim, et al., Application of interactive genetic algorithm to fashion design, Engineering Applications of Artificial Intelligence 13 (2000) 635–644.
[12] S. Handschuh, J. Gasteiger, Pharmacophores Derived from the 3D Substructure Perception, IUL Biotechnology Series 2 (Pharmacophore), 2000, pp. 431–453.
[13] S.-C. Ou, C.-Y. Chung, W.-T. Sung, C.-C. Tsai, C.-C. Chien, D.-Y. Su, Virtual screening and computer-aided drug design in molecular docking via Lyapunov function, WSEAS Transactions on Biology and Biomedicine 01 (4) (2004) 384–390.
[14] S.-C. Ou, W.-T. Sung, S.-J. Hsiao, K.-C. Fan, Interactive web-based training tool for CAD in a virtual environment, Computer Applications in Engineering Education 10 (4) (2002) 182–193.
[15] William M. Spears, Kenneth A. De Jong, Dining with GAs: operator lunch Theorem. in: Proceedings of Foundations of Genetic Algorithms, 1998.
[16] William M. Spears, Aggregating Models of Evolutionary Algorithms. in: Proceedings of the Conference on Evolutionary Computation, 2000.
[17] N. Srinivas, K. Deb, Multi-objective optimisation using non-dominated sorting in genetic algorithms, In Evolutionary Computation 2 (3) (1995) 221–248.
[18] C.-Y. Chung, S.-C. Ou, C.-C. Tsai, W.-T. Sung, Application of minimum energy to protein structure prediction and drug design, in: Presented at 2004 Chinese and Japan Molecular Design, Nanotechnology and Drug Transmission Conference IRC Biology Medical Center, 2004.
[19] R.J. Davy, I.A. Hiskens, Lyapunov functions for multi-machine power systems with dynamic loads, Technical Report no. EE9472, Department of Electrical and Computer Engineering, The University of Newcastle, IEEE Transactions on Circuits and Systems, 23(4), December 1994, pp.167–173.
[20] G. Kresse, J. Furthmüller, Efficiency of *Ab-Initio* total energy calculations for metals and semiconductors using a plane-wave basis set, Computational Material Science 6 (1996) 15–50.
[21] C.A. Laughton, A study of simulated annealing protocols for use with molecular dynamics in protein structure prediction, Protein Engineering 7 (1994) 235–241.
[22] S.-C. Ou, H.-Y. Chung, W.-T. Sung, Development of a computer aided geometric design system based on parallel architecture, WSEAS Transactions on Computers 5 (2) (2006) 278–284.
[23] C.-Y. Chung, S.-C. Ou, C.-C. Tsai, W.-T. Sung, C.-Ch. Chien, D.-Y. Su, Application of molecular docking to protein folding structure prediction and drug discovery, in: International Proteomics Conference (IPC'03), In the Grand Hotel on May 14–17, Taipei, Taiwan, 2004.
[24] H. Preben, Olesen, et al., A new class of selective GSK-3 inhibitors, Journal of Medicinal Chemistry 46 (15) (2003) 3333–3341.