Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function

GARRETT M. MORRIS,[1] DAVID S. GOODSELL,[1] ROBERT S. HALLIDAY,[2] RUTH HUEY,[1] WILLIAM E. HART,[3] RICHARD K. BELEW,[4] ARTHUR J. OLSON[1]

[1] *Department of Molecular Biology, MB-5, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037-1000*
[2] *Hewlett-Packard, San Diego, California*
[3] *Applied Mathematics Department, Sandia National Laboratories, Albuqurque, NM*
[4] *Department of Computer Science & Engineering, University of California, San Diego, La Jolla, CA*

Received February 1998; accepted 24 June 1998

ABSTRACT: A novel and robust automated docking method that predicts the bound conformations of flexible ligands to macromolecular targets has been developed and tested, in combination with a new scoring function that estimates the free energy change upon binding. Interestingly, this method applies a Lamarckian model of genetics, in which environmental adaptations of an individual's phenotype are reverse transcribed into its genotype and become heritable traits (*sic*). We consider three search methods, Monte Carlo simulated annealing, a traditional genetic algorithm, and the Lamarckian genetic algorithm, and compare their performance in dockings of seven protein–ligand test systems having known three-dimensional structure. We show that both the traditional and Lamarckian genetic algorithms can handle ligands with more degrees of freedom than the simulated annealing method used in earlier versions of AUTODOCK, and that the Lamarckian genetic algorithm is the most efficient, reliable, and successful of the three. The empirical free energy function was calibrated using a set of 30 structurally known protein–ligand complexes with experimentally determined binding constants. Linear regression analysis of the observed binding constants in terms of a wide variety of structure-derived molecular properties was performed. The final model had a residual standard error of 9.11 kJ mol^{-1} (2.177 kcal mol^{-1}) and was chosen as the new energy

Correspondence to: A. J. Olson; e-mail: olson@scripps.edu
Contract/grant sponsor: National Institutes of Health, contract/grant numbers: GM48870, RR08065

CCC 0192-8651 / 98 / 141639-24

function. The new search methods and empirical free energy function are available in AUTODOCK, version 3.0. © 1998 John Wiley & Sons, Inc. J Comput Chem 19: 1639–1662, 1998

Keywords: automated docking; binding affinity; drug design; genetic algorithm; flexible small molecule protein interaction

Introduction

A fast atom-based computational docking tool is essential to most techniques for structure-based drug design.[1,2] Reported techniques for automated docking fall into two broad categories: matching methods and docking simulation methods.[3] Matching methods create a model of the active site, typically including sites of hydrogen bonding and sites that are sterically accessible, and then attempt to dock a given inhibitor structure into the model as a rigid body by matching its geometry to that of the active site. The most successful example of this approach is DOCK,[4,5] which is efficient enough to screen entire chemical databases rapidly for lead compounds. The second class of docking techniques model the docking of a ligand to a target in greater detail: the ligand begins randomly outside the protein, and explores translations, orientations, and conformations until an ideal site is found. These techniques are typically slower than the matching techniques, but they allow flexibility within the ligand to be modeled and can utilize more detailed molecular mechanics to calculate the energy of the ligand in the context of the putative active site. They allow computational chemists to investigate modifications of lead molecules suggested by the chemical intuition and expertise of organic synthetic chemists.

AUTODOCK[6,7] is an example of the latter, more physically detailed, flexible docking technique. Previous releases of AUTODOCK combine a rapid grid-based method for energy evaluation,[8,9] precalculating ligand–protein pairwise interaction energies so that they may be used as a look-up table during simulation, with a Monte Carlo simulated annealing search[10,11] for optimal conformations of ligands. AUTODOCK has been applied with great success in the prediction of bound conformations of enzyme–inhibitor complexes,[12,13] peptide–antibody complexes,[14] and even protein–protein interactions[15]; these and other applications have been reviewed elsewhere.[16]

We initiated the current work to remedy two limitations of AUTODOCK. (i) We have found that the simulated annealing search method performs well with ligands that have roughly eight rotatable bonds or less: problems with more degrees of freedom rapidly become intractable. This demanded a more efficient search method. (ii) AUTODOCK is often used to obtain unbiased dockings of flexible inhibitors in enzyme active sites: in computer-assisted drug-design, novel modifications of such lead molecules can be investigated computationally. Like many other computational approaches, AUTODOCK performs well in predicting relative quantities and rankings for series of similar molecules; however, it has not been possible to estimate in AUTODOCK whether a ligand will bind with a millimolar, micromolar, or nanomolar binding constant. Earlier versions of AUTODOCK used a set of traditional molecular mechanics force-field parameters that were not directly correlated with observed binding free energies; hence, we needed to develop a force field that could be used to predict such quantities.

Molecular docking is a difficult optimization problem, requiring efficient sampling across the entire range of positional, orientational, and conformational possibilities. Genetic algorithms (GA) fulfill the role of global search particularly well, and are increasingly being applied to problems that suffer from combinatorial explosions due to their many degrees of freedom. Both canonical genetic algorithms[17–21] and evolutionary programming methods[22] have been shown to be successful in both drug design and docking.

In this report, we describe two major advances that are included in the new release of AUTODOCK, version 3.0. The first is the addition of three new search methods: a genetic algorithm; a local search method; and a novel, adaptive global–local search method based on Lamarckian genetics, the Lamarckian genetic algorithm (LGA). The second advance is an empirical binding free energy force field that allows the prediction of binding free energies, and hence binding constants, for docked ligands.

Methods

GENETIC ALGORITHMS

Genetic algorithms[23] use ideas based on the language of natural genetics and biological evolution.[24] In the case of molecular docking, the particular arrangement of a ligand and a protein can be defined by a set of values describing the translation, orientation, and conformation of the ligand with respect to the protein: these are the ligand's *state variables* and, in the GA, each state variable corresponds to a gene. The ligand's state corresponds to the genotype, whereas its atomic coordinates correspond to the phenotype. In molecular docking, the *fitness* is the total interaction energy of the ligand with the protein, and is evaluated using the energy function. Random pairs of individuals are mated using a process of *crossover*, in which new individuals inherit genes from either parent. In addition, some offspring undergo random *mutation*, in which one gene changes by a random amount. *Selection* of the offspring of the current *generation* occurs based on the individual's fitness: thus, solutions better suited to their environment reproduce, whereas poorer suited ones die.

A variety of approaches have been adopted to improve the efficiency of the genetic algorithm. Classical genetic algorithms represent the genome as a fixed-length bit string, and employ *binary* crossover and *binary* mutation to generate new individuals in the population. Unfortunately, in many problems, such binary operators can generate values that are often outside the domain of interest, leading to gross inefficiencies in the search. The use of real encodings helps to limit the genetic algorithm to reasonable domains. Alternative genetic algorithms have been reported[25] that employ more complicated representations and more sophisticated operators besides crossover and mutation. Some of these retain the binary representation, but must employ decoders and repair algorithms to avoid building illegal individuals from the chromosome, and these are frequently computationally intensive. However, the search performance of the genetic algorithm can be improved by introducing a local search method.[26,27]

HYBRID SEARCH METHODS IN AUTODOCK

Earlier versions of AUTODOCK used optimized variants of simulated annealing.[6,7] Simulated annealing may be viewed as having both global and local search aspects, performing a more global search early in the run, when higher temperatures allow transitions over energy barriers separating energetic valleys, and later on performing a more local search when lower temperatures place more focus on local optimization in the current valley. AUTODOCK 3.0 retains the functionality of earlier versions, but adds the options of using a genetic algorithm (GA) for global searching, a local search (LS) method to perform energy minimization, or a combination of both, and builds on the work of Belew and Hart.[27,28] The local search method is based on that of Solis and Wets,[29] which has the advantage that it does not require gradient information about the local energy landscape, thus facilitating torsional space search. In addition, the local search method is *adaptive*, in that it adjusts the step size depending upon the recent history of energies: a user-defined number of consecutive failures, or increases in energy, cause the step size to be doubled; conversely, a user-defined number of consecutive successes, or decreases in energy, cause the step size to be halved. The hybrid of the GA method with the adaptive LS method together form the so-called *Lamarckian genetic algorithm* (LGA), which has enhanced performance relative to simulated annealing and GA alone,[21,26] and is described in detail later. Thus, the addition of these new GA-based docking methods enhances AUTODOCK, and allows problems with more degrees of freedom to be tackled. Furthermore, it is now possible to use the same force field as is used in docking to perform energy minimization of ligands.

IMPLEMENTATION

In our implementation of the genetic algorithm, the chromosome is composed of a string of real-valued genes: three Cartesian coordinates for the ligand translation; four variables defining a quaternion specifying the ligand orientation; and one real-value for each ligand torsion, in that order. Quaternions are used to define the orientation[30] of the ligand, to avoid the gimbal lock problem experienced with Euler angles.[31] The order of the genes that encode the torsion angles is defined by the torsion tree created by AUTOTORS, a preparatory program used to select rotatable bonds in the ligand. Thus, there is a one-to-one mapping from the ligand's state variables to the genes of the individual's chromosome.

The genetic algorithm begins by creating a random population of individuals, where the user

defines the number of individuals in the population. For each random individual in the initial population, each of the three translation genes for x, y, and z is given a uniformly distributed random value between the minimum and maximum x, y, and z extents of the grid maps, respectively; the four genes defining the orientation are given a random quaternion, consisting of a random unit vector and a random rotation angle between $-180°$ and $+180°$; and the torsion angle genes, if any, are given random values between $-180°$ and $+180°$. Furthermore, a new random number generator has been introduced that is hardware-independent.[32] It is used in the LS, GA, and LGA search engines, and allows results to be reproduced on any hardware platform given the same seed values. The creation of the random initial population is followed by a loop over generations, repeating until the maximum number of generations or the maximum number of energy evaluations is reached, whichever comes first. A generation consists of five stages: mapping and fitness evaluation, selection, crossover, mutation, and elitist selection, in that order. In the Lamarckian GA, each generation is followed by local search, being performed on a user-defined proportion of the population. Each of these stages is discussed in more detail in what follows.

Mapping translates from each individual's genotype to its corresponding phenotype, and occurs over the entire population. This allows each individual's *fitness* to be evaluated. This is the sum of the intermolecular interaction energy between the ligand and the protein, and the intramolecular interaction energy of the ligand. The physicochemical nature of the energy evaluation function is described in detail later. Every time an individual's energy is calculated, either during global or local search, a count of the total number of energy evaluations is incremented.

This is followed, in our implementation, by proportional *selection* to decide which individuals will reproduce. Thus, individuals that have better-than-average fitness receive proportionally more offspring, in accordance with:

$$n_o = \frac{f_w - f_i}{f_w - \langle f \rangle} \qquad f_w \neq \langle f \rangle$$

where n_o is the integer number of offspring to be allocated to the individual; f_i is the fitness of the individual (i.e., the energy of the ligand); f_w is the fitness of the worst individual, or highest energy, in the last N generations (i.e., N is a user-definable parameter, typically 10); and $\langle f \rangle$ is the mean fitness of the population. Because the worst fitness, f_w, will always be larger than either f_i or $\langle f \rangle$, except when $f_i = f_w$, then for individuals that have a fitness lower than the mean, $f_i < \langle f \rangle$, the numerator in this equation, $f_w - f_i$, will always be greater than the denominator $f_w - \langle f \rangle$, and thus such individuals will be allocated at least one offspring, and thus will be able to reproduce. AUTODOCK checks for $f_w = \langle f \rangle$ beforehand, and if true, the population is assumed to have converged, and the docking is terminated.

Crossover and *mutation* are performed on random members of the population according to user-defined rates of crossover and mutation. First, crossover is performed. Two-point crossover is used, with breaks occurring only between genes, never within a gene—this prevents erratic changes in the real values of the genes. Thus, both parents' chromosomes would be broken into three pieces at the same gene positions, each piece containing one or more genes; for instance, *ABC* and *abc*. The chromosomes of the resulting offspring after two-point crossover would be *AbC* and *aBc*. These offspring replace the parents in the population, keeping the population size constant. Crossover is followed by mutation; because the translational, orientational, and torsional genes are represented by real variables, the classical bit-flip mutation would be inappropriate. Instead, mutation is performed by adding a random real number that has a Cauchy distribution to the variable, the distribution being given by:

$$C(\alpha, \beta, x) = \frac{\beta}{\pi\left(\beta^2 + (x - \alpha)^2\right)}$$
$$\alpha \geq 0,\ \beta > 0,\ -\infty < x < \infty$$

where α and β are parameters that affect the mean and spread of the distribution. The Cauchy distribution has a bias toward small deviates, but, unlike the Gaussian distribution, it has thick tails that enable it to generate large changes occasionally.[26]

An optional user-defined integer parameter *elitism* determines how many of the top individuals automatically survive into the next generation. If the elitism parameter is non-zero, the new population that has resulted from the proportional selection, crossover, and mutation is sorted according to its fitness; the fitness of new individuals

having resulted from crossover and/or mutation is calculated as necessary. Because populations are implemented as heaps, selection of the best n individuals is efficient.

The genetic algorithm iterates over generations until one of the termination criteria is met. At the end of each docking, AUTODOCK reports the fitness (the docked energy), the state variables, and the coordinates of the docked conformation, and also the estimated free energy of binding. AUTODOCK performs the user-specified number of GA dockings, and then carries out conformational cluster analysis on the docked conformations to determine which are similar, reporting the clusters ranked by increasing energy.

LAMARCKIAN GENETIC ALGORITHM

The vast majority of genetic algorithms mimic the major characteristics of Darwinian evolution and apply Mendelian genetics. This is illustrated on the right-hand side of Figure 1 (note the one-way transfer of information from the genotype to the phenotype). However, in those cases where an inverse mapping function exists (i.e., one which yields a genotype from a given phenotype), it is possible to finish a local search by replacing the individual with the result of the local search; see the left-hand side of Figure 1. This is called the Lamarckian genetic algorithm (LGA), and is an allusion to Jean Batiste de Lamarck's (discredited) assertion that phenotypic characteristics acquired during an individual's lifetime can become heritable traits.[33]

The most important issues arising in hybrids of local search (LS) techniques with the GA revolve around the *developmental mapping*, which transforms genotypic representations into phenotypic ones.[26] The genotypic space is defined in terms of the genetic operators—mutation and crossover in our experiments—by which parents of one generation are perturbed to form their children. The phenotypic space is defined directly by the problem, namely, the energy function being optimized. The local search operator is a useful extension of GA global optimization when there are local ''smoothness'' characteristics (continuity, correlation, etc.) of the fitness function that local search can exploit. In hybrid GA + LS optimizations, the result of the LS is always used to update the fitness associated with an individual in the GA selection algorithm. If, and only if, the developmental mapping function is *invertible*, will the Lamarckian option—converting the phenotypic result of LS back into its corresponding genotype become possible.

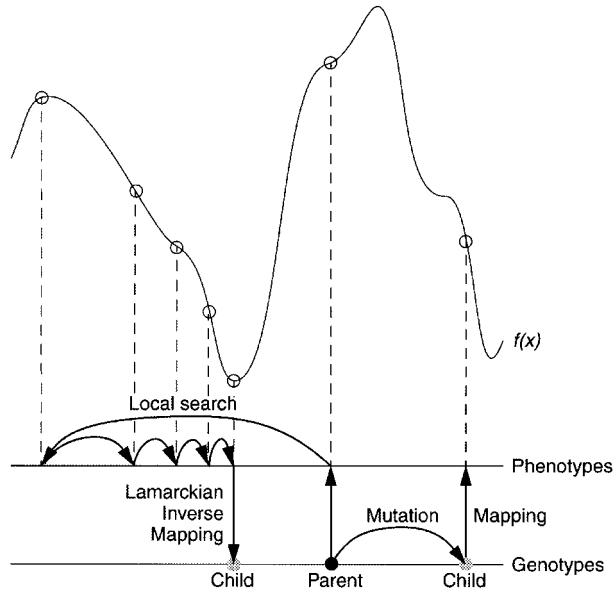

FIGURE 1. This figure illustrates genotypic and phenotypic search, and contrasts Darwinian and Lamarckian search.[27] The space of the genotypes is represented by the lower horizontal line, and the space of the phenotypes is represented by the upper horizontal line. Genotypes are mapped to phenotypes by a developmental mapping function. The fitness function is $f(x)$. The result of applying the genotypic mutation operator to the parent's genotype is shown on the right-hand side of the diagram, and has the corresponding phenotype shown. Local search is shown on the left-hand side. It is normally performed in phenotypic space and employs information about the fitness landscape. Sufficient iterations of the local search arrive at a local minimum, and an inverse mapping function is used to convert from its phenotype to its corresponding genotype. In the case of molecular docking, however, local search is performed by continuously converting from the genotype to the phenotype, so inverse mapping is not required. The genotype of the parent is replaced by the resulting genotype, however, in accordance with Lamarckian principles.

In our case, the fitness or energy is calculated from the ligand's coordinates, which together form its phenotype. The genotypic representation of the ligand, and its mutation and crossover operators, have already been described. The developmental mapping simply transforms a molecule's genotypic state variables into the corresponding set of atomic coordinates. A novel feature of this application of hybrid global–local optimization is that the

Solis and Wets LS operator searches through the genotypic space rather than the more typical phenotypic space. This means that the developmental mapping does not need to be inverted. Nonetheless, this molecular variation of the genetic algorithm still qualifies as Lamarckian, because any ''environmental adaptations'' of the ligand acquired during the local search will be inherited by its offspring.

At each generation, it is possible to let a user-defined fraction of the population undergo such a local search. We have found improved efficiency of docking with local search frequencies of just 0.06, although a frequency of 1.00 is not significantly more efficient.[26] Both the canonical and a slightly modified version of the Solis and Wets method have been implemented. In canonical Solis and Wets, the same step size would be used for every gene, but we have improved the local search efficiency by allowing the step size to be different for each type of gene: a change of 1 Å (1 Å = 10^{-10} m) in a translation gene could be much more significant than a change of 1° in a rotational or torsional gene. In the docking experiments presented here, the translational step size was 0.2 Å, and the orientational and torsional step sizes were 5°.

In the Lamarckian genetic algorithm, genotypic mutation plays a somewhat different role than it does in traditional genetic algorithms. Traditionally, mutation plays the role of a local search operator, allowing small, refining moves that are not efficiently made by crossover and selection alone. With the explicit local search operator, however, this role becomes unnecessary, and is needed only for its role in replacing alleles that might have disappeared through selection. In LGA, mutation can take on a more exploratory role. The Cauchy deviates are a compromise between radical jumps to arbitrary sections of the conformation space and detailed exploration of the local topography.

DERIVATION OF THE EMPIRICAL BINDING FREE ENERGY FUNCTION

The study of molecular structure underpins much of computational molecular biology. There are several established methods for performing molecular mechanics and molecular dynamics, notably AMBER,[34,35] CHARMM,[36] DISCOVER,[37] ECEPP,[38] and GROMOS.[39] Many of these traditional force fields model the interaction energy of a molecular system with terms for dispersion/repulsion,[40] hydrogen bonding,[41] electrostatics,[42–45] and deviation from ideal bond lengths and bond angles. These methods are excellent for studying molecular processes over time, for optimizing bound conformations, and for performing free energy perturbation calculations between molecules with a single atom change,[46] but they often require considerable investments of computer time and, unfortunately, these approaches tend to perform less well in ranking the binding free energies of compounds that differ by more than a few atoms. What is needed is an empirical relationship between molecular structure and binding free energy.

The first thoroughly established linear free energy relationship was observed by Hammett as early as 1933, and reported in 1937.[47] It was used to relate structure and reactivity of small organic molecules on a quantitative basis. Hammett was able to derive substituent constants and reaction constants that could then be used to calculate rate constants and equilibrium constants for a specific reaction of a specific compound. It could be said that Hammett's work was the forerunner of modern-day quantitative structure–activity relationships (QSAR), pioneered by Hansch and coworkers in the 1960s. Here it is assumed that the sum of the steric, electronic, and hydrophobic effects of substituents in a compound determines its biological activity; see, for example, Fujita,[48] Hansch,[49] and more recently Selassie et al.[50]

Current structure-based scoring functions seek to remedy some of the deficiencies of traditional force fields by developing empirical free energy functions that reproduce observed binding constants. Most of these approaches use an expanded ''master equation'' to model the free energy of binding, adding entropic terms to the molecular mechanics equations[51]:

$$\Delta G = \Delta G_{\mathrm{vdw}} + \Delta G_{\mathrm{hbond}} + \Delta G_{\mathrm{elec}} + \Delta G_{\mathrm{conform}} + \Delta G_{\mathrm{tor}} + \Delta G_{\mathrm{sol}}$$

where the first four terms are the typical molecular mechanics terms for dispersion/repulsion, hydrogen bonding, electrostatics, and deviations from covalent geometry, respectively; ΔG_{tor} models the restriction of internal rotors and global rotation and translation; and ΔG_{sol} models desolvation upon binding and the hydrophobic effect (solvent entropy changes at solute–solvent interfaces). This latter term is the most challenging. Most workers use variants of the method of Wesson and Eisen-

berg,[52] calculating a desolvation energy based on the surface area buried upon complex formation, with the area of each buried atom being weighted by an atomic solvation parameter. Böhm built on earlier work with the *de novo* inhibitor design program LUDI,[53] and used linear regression to calibrate a similar function against a set of 45 diverse protein–ligand complexes with published binding constants.[54] The final function predicted binding constants for a set of test complexes with a standard deviation equivalent to about a factor of 25 in binding constant: more than sufficient to rank inhibitors with millimolar, micromolar, and nanomolar binding constants. Jain devised a continuous, differentiable scoring function,[55] which is, in essence, very similar to that of Böhm, but based on non-physical pairwise potentials using Gaussians and sigmoidal terms.

We have implemented a similar approach using the thermodynamic cycle of Wesson and Eisenberg.[52] The function includes five terms:

$$\begin{aligned}\Delta G = {} & \Delta G_{\mathrm{vdW}} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \\ & + \Delta G_{\mathrm{hbond}} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\ & + \Delta G_{\mathrm{elec}} \sum_{i,j} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \\ & + \Delta G_{\mathrm{tor}} N_{\mathrm{tor}} \\ & + \Delta G_{\mathrm{sol}} \sum_{i,j} (S_i V_j + S_j V_i) \mathrm{e}^{(-r_{ij}^2/2\sigma^2)} \end{aligned} \quad (1)$$

where the five ΔG terms on the right-hand side are coefficients empirically determined using linear regression analysis from a set of protein–ligand complexes with known binding constants, shown in Table I. The summations are performed over all pairs of ligand atoms, i, and protein atoms, j, in addition to all pairs of atoms in the ligand that are separated by three or more bonds.

The *in vacuo* contributions include three interaction energy terms, used in previous versions of AUTODOCK: a Lennard–Jones 12-6 dispersion/repulsion term; a directional 12–10 hydrogen bonding term, where $E(t)$ is a directional weight based on the angle, t, between the probe and the target atom[9]; and a screened Coulombic electrostatic potential.[56] Each of these terms, including their parameterization, have already been described.[7]

A measure of the unfavorable entropy of ligand binding due to the restriction of conformational degrees of freedom is added to the *in vacuo* function. This term is proportional to the number of sp^3 bonds in the ligand, N_{tor}.[54] We investigated variants that included and excluded methyl, hydroxyl, and amine rotors.

In the development of an empirical free energy function for AUTODOCK, the desolvation term was most challenging, because AUTODOCK uses a grid-based method for energy evaluation, and most published solvation methods are based on surface area calculations. We investigated two different methods of calculating the desolvation energy term. The first of these methods was based on estimating atom-by-atom contributions to the interfacial molecular surface area between the ligand and the protein using the difference in the surface areas of the complex and the unbound protein and unbound ligand. Both the solvent-accessible and solvent-excluded surface areas were considered, being calculated with MSMS,[57] a fast and reliable program that computes analytical molecular surfaces. Unfortunately, there can be significant errors in the value of the interfacial solvent-accessible surface areas, due to the ''collar'' of accessible surface that surrounds the ligand–protein interface in the complex. We also tested seven variants of the pairwise, volume-based method of Stouten et al.[58]: this method has the advantage that it is consistent with the pre-calculated affinity grid formulation used by AUTODOCK. For each atom in the ligand, fragmental volumes of surrounding protein atoms are weighted by an exponential function and then summed, evaluating the percentage of volume around the ligand atom that is occupied by protein atoms. This percentage is then weighted by the atomic solvation parameter of the ligand atom to give the desolvation energy. The full method may be broken into four separate components: burial of apolar atoms in the ligand, burial of apolar protein atoms, burial of polar and charged atoms in the ligand, and burial of polar and charged protein atoms. Great success has also been reported in using simply the amount of hydrophobic surface area buried upon complexation as a measure of the ''hydrophobic effect,''[54] so we tested several formulations that included only the volume lost around ligand carbon atoms. The burial of polar atoms caused particular problems, as discussed in what follows. Apart from the volume-based method, we tested a simpler formulation for the solvent transfer of polar atoms; that is, a constant term corresponding to the favorable free energy of interaction of a polar atom with solvent

TABLE I.
Protein–Ligand Complexes Used to Calibrate Empirical Free Energy Function, Along with Brookhaven Protein Data Bank (PDB) Accession Codes and Binding.

Protein–ligand complex	PDB code	Log(K_i)[a]
Concanavalin A / α-methyl-D-mannopyranoside	4cna	2.00
Carboxypeptidase A / glycyl-L-tyrosine	3cpa	3.88
Carboxypeptidase A / phosphonate ZAA═P═(O)F	6cpa	11.52
Cytochrome P-450$_{cam}$ / camphor	2cpp	6.07
Dihydrofolate reductase / methotrexate	4dfr	9.70
α-Thrombin / benzamidine	1dwb	2.92
Endothiapepsin / H-256	zer6	7.22
ε-Thrombin / MQPA	1etr	7.40
ε-Thrombin / NAPAP	1ets	8.52
ε-Thrombin / 4-TAPAP	1ett	6.19
FK506-binding protein (FKBP) / immunosuppressant FK506	1fkf	9.70
D-Galactose / D-glucose binding protein / galactose	2gbp	7.60
Hemagglutinin / sialic acid	4hmg	2.55
HIV-1 Protease / A78791	1hvj	10.46
HIV-1 Protease / MVT101	4hvp	6.15
HIV-1 Protease / acylpepstatine	5hvp	5.96
HIV-1 Protease / XK263	1hvr	9.51
Fatty-acid-binding protein / C_{15}COOH	2ifb	5.43
Myoglobin (ferric) / imidazole	1mbi	1.88
McPC603 / phosphocholine	2mcp	5.23
β-Trypsin / benzamidine	3ptb	4.74
Retinol-binding protein / retinol	1rbp	6.72
Thermolysin / Leu-hydroxylamine	4tln	3.72
Thermolysin / phosphoramidon	1tlp	7.55
Thermolysin / *n*-(1-carboxy-3-phenylpropyl)-Leu-Trp	1tmn	7.30
Thermolysin / Cbz-Phe-*p*-Leu-Ala (ZFpLA)	4tmn	10.19
Thermolysin / Cbz-Gly-*p*-Leu-Leu (ZGpLL)	5tmn	8.04
Purine nucleoside phosphorylase (PNP) / guanine	1ulb	5.30
Xylose isomerase / CB3717	2xis	5.82
Triose phosphate isomerase (TIM) / 2-phosphoglycolic acid (PGA)	2ypi	4.82

[a] Adapted from Böhm.[54]

is estimated, and this is subtracted from the binding free energy.

Trilinear interpolation is used to evaluate rapidly the intermolecular dispersion/repulsion energy, the hydrogen bonding energy, the electrostatic potential, and the solvation energy of each atom in the ligand, using grid maps that have been pre-calculated over the protein for each atom type in the ligand. In AUTODOCK 3.0, we have implemented a faster method of trilinear interpolation[59] than was available in earlier versions of AUTODOCK. Both methods are mathematically equivalent. The original implementation used 24 multiplications to perform each three-dimensional trilinear interpolation, but, by cascading seven one-dimensional interpolations, the number of multiplications has been reduced to 7.

Thirty protein–ligand complexes with published binding constants were used in the calibration of AUTODOCK's free energy function (Table I), and were chosen from the set of 45 used by Böhm,[54] omitting all complexes that he modeled (i.e., using only complexes for which crystallographic structures were available). One of the limitations of these binding constant data is that the conditions under which they were determined vary, which intrinsically limits the accuracy of our best model. We converted between the inhibition constant, K_i, and the observed free energy change of binding, ΔG_{obs}, using the equation:

$$\Delta G_{obs} = RT \ln K_i$$

where R is the gas constant, 1.987 cal K^{-1} mol^{-1}, and T is the absolute temperature, assumed to be

room temperature, 298.15 K.[60] Note that this equation lacks a minus sign because the inhibition constant is defined for the dissociation reaction, $EI \rightleftharpoons E + I$,[61] whereas ΔG_{obs} refers to the opposite process of binding, $E + I \rightleftharpoons EI$; where E is the enzyme and I is the inhibitor.

To remove any steric clashes in the crystallographic complexes, each ligand was optimized using AUTODOCK's new Solis and Wets local minimization technique described earlier, but with the previously reported force field.[7] The separate contributions from the hydrogen bonding, dispersion/repulsion, electrostatic, and solvation energies were evaluated. Empirical free energy coefficients for each of these terms were derived using linear regression in the S-PLUS software package,[62] and cross-validation studies were performed. In total, 900 different binding free energy models were tested: each linear model consisted of a van der Waals term, a hydrogen bonding term (one of 6 variants), an electrostatic term, a torsional entropy term (one of 5 variants), and a desolvation term (one of 15 variants). We also investigated whether the inclusion of a constant term improved the model. Six of the seven test systems used to test the docking procedure, which were originally used to test AUTODOCK, version 2.4,[7] were also in the training set of 30 protein–ligand complexes; therefore, to validate the chosen coefficients, linear regression was repeated for the set of 24 protein–ligand complexes, excluding the 6 overlapping test systems.

TESTING DOCKING METHODS

Seven protein–ligand complexes, with a range of complexity and chemical properties, were chosen from the Brookhaven Protein Data Bank[63, 64] to compare the performance of the docking techniques (see Fig. 2). To facilitate comparison with the previous force field, we chose the same set of six test systems investigated earlier,[7] but added a harder docking problem to challenge all the search methods (see Table II). The simplest test cases were the β-trypsin/benzamidine and cytochrome $P\text{-}450_{cam}$/camphor complexes, which had small, rigid ligands. Interactions in the former are dominated by electrostatic interactions and hydrogen bonds to the substrate amidine, whereas the latter is dominated by hydrophobic interactions. McPC-603/phosphocholine and streptavidin/biotin were moderately flexible, and represented test systems having an intermediate level of difficulty. HIV-1 protease/XK263, hemagglutinin/sialic acid, and dihydrofolate reductase/methotrexate provided more difficult tests, with many rotatable bonds and diverse chemical characteristics.

We compared the performance of Monte Carlo simulated annealing (SA), the genetic algorithm (GA), and the Lamarckian genetic algorithm (LGA). The new empirical free energy function presented here was used for energy evaluation in all cases. Dockings were performed using approximately the same number of energy evaluations ($\sim$ 1.5 million), so each method could be judged given similar computational investments. The CPU time taken for a single docking varied from 4.5 to 41.3 minutes, on a 200-MHz Silicon Graphics MIPS 4400 with 128 MB of RAM, depending on the number of rotatable bonds and the number of atoms in the ligand.

At the end of a set of dockings, the docked conformations were exhaustively compared to one another to determine similarities, and were clustered accordingly. The user-defined root-mean-square positional deviation (rmsd) tolerance was used to determine if two docked conformations were similar enough to be included in the same cluster, and symmetrically related atoms in the ligand were considered. These clusters were ranked in order of increasing energy, by the lowest energy in each cluster. Ordinarily, the structure of the protein–ligand complex would not be known, so the criteria by which the dockings would be evaluated are the energies of the docked structures, and, in cases where there are several plausible, low-energy structures, the number of conformations in a conformationally similar cluster. Because one of our goals was to test the ability of the methods to reproduce known structures, we also compared the rmsd between the lowest energy docked structure and the crystallographic structure.

DOCKING- AND SEARCH-METHOD-SPECIFIC PARAMETERS

The proteins and ligands in the seven docking tests were treated using the united-atom approximation, and prepared using the molecular modeling program, SYBYL.[65] Only polar hydrogens were added to the protein, and Kollman united-atom partial charges were assigned. Unless stated otherwise, all waters were removed. Atomic solvation parameters and fragmental volumes were assigned to the protein atoms using a new AUTODOCK utility, ADDSOL. The grid maps were calculated using AUTOGRID, version 3.0. In all seven protein–ligand

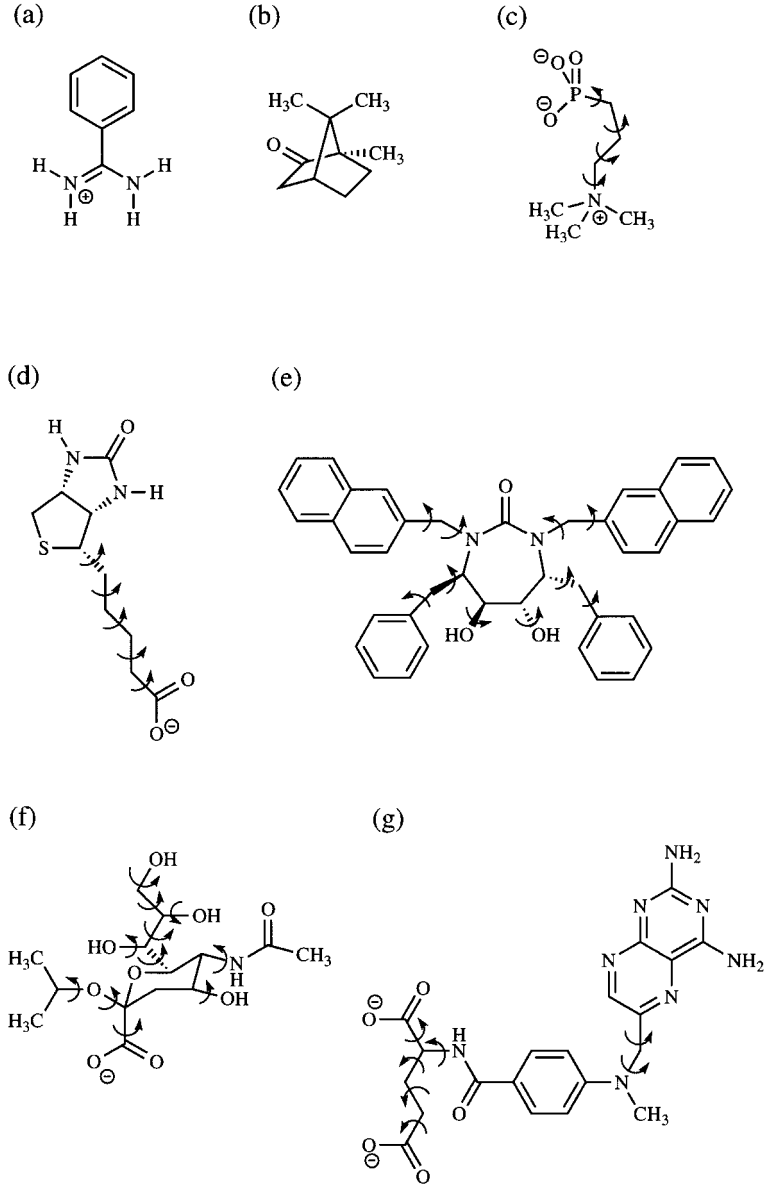

FIGURE 2. The seven ligands chosen for docking, showing the rotatable bonds as curly arrows: (a) benzamidine; (b) camphor; (c) phosphocholine; (d) biotin; (e) HIV-1 protease inhibitor XK-263; (f) isopropylated sialic acid; and (g) methotrexate. Note that two ligands, (e) and (f), contain hydroxyl rotors, which are not counted in the total number of torsional degrees of freedom; note also that cyclic rotatable bonds are excluded.

cases, we used grid maps with 61 × 61 × 61 points, a grid-point spacing of 0.375 Å, and, because the location of the ligand in the complex was known, the maps were centered on the ligand's binding site. The ligands were treated in SYBYL initially as all atom entities, that is, all hydrogens were added, then partial atomic charges were calculated using the Gasteiger–Marsili method.[66,67] AUTOTORS, an AUTODOCK utility, was used to define the rotatable bonds in the ligand, if any, and also to unite the nonpolar hydrogens added by SYBYL for the partial atomic charge calculation. The partial charges on the nonpolar hydrogens were added to that of the hydrogen-bearing carbon also in AUTOTORS.

In all three search methods, 10 dockings were performed; in the analysis of the docked conformations, the clustering tolerance for the root-mean-square positional deviation was 0.5 Å, and the crystallographic coordinates of the ligand were used as the reference structure. For all three search

TABLE II.

X-Ray Crystal Structure Coordinates Used in Docking Experiments, Their Brookhaven Protein Data Bank Accession Codes and Resolution, Number of Rotatable Bonds in the Ligand, Number of Torsional Degrees of Freedom, Total Number of Degrees of Freedom, and Energy of Crystal Structure Using the Empirical Force Field Presented Here.

Protein–ligand complex	PDB code	Resolution (Å)	Reference	Number of rotatable bonds	N_{tor} [a]	Total number of degrees of freedom	Energy of crystal structure (kcal mol^{-1})
β-Trypsin / benzamidine	3ptb	1.7	69	0	0	7	−7.86
Cytochrome P-450$_{cam}$ / camphor	2cpp	1.63	70	0	0	7	−4.71
McPC-603 / Phosphocholine	2mcp	3.1	71	4	4	11	+5.48[b]
Streptavidin / biotin	1stp	2.6	73	5	5	12	−8.86
HIV-1 protease / XK263	1hvr	1.8	75	10	8	17	−18.62
Influenza hemagglutinin / sialic acid	4hmg	3.0	76	11	7	18	−4.71
Dihydrofolate reductase / methotrexate	4dfr	1.7	79	7	7	14	−13.64

[a] N_{tor} is the number of torsional degrees of freedom used in the calculation of the predicted free energy change of binding, ΔG_{pred}. Note that this excludes rotatable bonds that only move hydrogens, such as hydroxyl, amino, and methyl groups.

[b] This energy is dominated by a large, positive contribution from C2 and O1 to the internal nonbonded energy, of +6.13 kcal mol^{-1}; these atoms are 2.26 Å apart.

methods, the step sizes were 0.2 Å for translations and 5° for orientations and torsions. These step sizes determined the amount by which a state variable could change when a move is made in simulated annealing and the relative size of mutation in the local search, whereas the α and β parameters determined the size of the mutation in the genetic algorithms, GA and LGA. The Cauchy distribution parameters were $\alpha = 0$ and $\beta = 1$. Note that in simulated annealing, random changes were generated by a uniformly distributed random number generator; in the Solis and Wets local search, by a normal distribution; and, in the genetic algorithm, by a Cauchy distribution. In the simulated annealing tests, the initial state of the ligand was chosen randomly by AUTODOCK. We used the optimal set of simulated annealing parameters that were determined from the schedule experiments described earlier.[7] These included an initial annealing temperature of 616 cal mol^{-1}, a linear temperature reduction schedule, 10 runs, 50 cycles, and a cycle-termination criterion of a maximum of 25,000 accepted steps or 25,000 rejected steps, whichever came first. The minimum energy state was used to begin the next cycle; the only exception was for 1hvr, where the initial annealing temperature was increased to 61,600 cal mol^{-1}. The maximum initial energy allowed was 0.0 kcal mol^{-1}, and the maximum number of retries was 1000, used to generate a low energy random initial state to begin each simulated annealing docking.

In the GA and LGA dockings, we used an initial population of random individuals with a population size of 50 individuals; a maximum number of 1.5×10^6 energy evaluations; a maximum number of generations of 27,000; an elitism value of 1, which was the number of top individuals that automatically survived into the next generation; a mutation rate of 0.02, which was the probability that a gene would undergo a random change; and a crossover rate of 0.80, which was the probability that two individuals would undergo crossover. Proportional selection was used, where the average of the worst energy was calculated over a window of the previous 10 generations. In the LGA dockings, the pseudo-Solis and Wets local search method was used, having a maximum of 300 iterations per local search; the probability of performing local search on an individual in the population was 0.06; the maximum number of consecutive successes or failures before doubling or halving the local search step size, ρ, was 4, in both cases; and the lower bound on ρ, the termination criterion for the local search, was 0.01.

Results and Discussion

CALIBRATION OF EMPIRICAL FREE ENERGY FUNCTION

Several linear regression models were tested for their ability to reproduce the observed binding

TABLE III.
Calibration of Empirical Free Energy Function.

Model[a]	Residual standard error	Multiple R^2	ΔG_{vdw}[b]	ΔG_{estat}	ΔG_{hbond}	ΔG_{tor}	ΔG_{solv}
A	2.324	0.9498	0.1795 (0.0263)	0.1133 (0.0324)	0.0166 (0.0625)	0.3100 (0.0873)	0.0101 (0.0585)
B	2.232	0.9537	0.1518 (0.0269)	0.1186 (0.0246)	0.0126 (0.0382)	0.3548 (0.0890)	0.1539 (0.1050)
C	2.177	0.9559	0.1485 (0.0237)	0.1146 (0.0238)	0.0656 (0.0558)	0.3113 (0.0910)	0.1711 (0.1035)

[a] Models differ in the formulation of the solvation term and the hydrogen bonding term. Model A: full volume-based solvation term and standard 10–12 hydrogen bonding, as in Eq. (1). Model B: apolar ligand atoms only in the solvation term, and standard 10–12 hydrogen bonding. Model C: apolar ligand atoms only in the solvation term, and the standard 10–12 hydrogen less the estimated average, as in Eq. (2).
[b] Values for the model coefficients, with standard deviations in parentheses.

constants of structurally characterized complexes. Table III shows the results for the three major candidates, and Figure 3 shows the correlation between the observed and the predicted binding free energies for the 30 protein–ligand complexes in the calibration set, using the chosen model (model C). Model A adds the full volume-based solvation method and the torsional restriction term to the original molecular mechanics force field. Model B simplifies the solvation method by evalu-

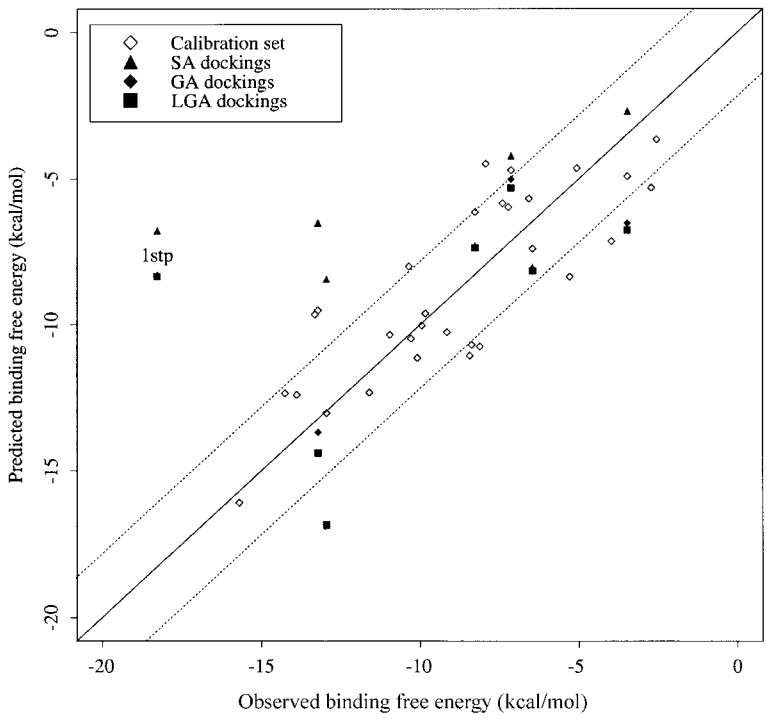

FIGURE 3. Predicted versus observed binding free energies for the calibration set and the docking tests. The solid line shows a perfect fit, and the dotted lines show one standard deviation above and below this. Hollow diamonds show the 30 protein–ligand complexes used in fitting the terms of the binding free energy function. Solid triangles show the results of the simulated annealing (SA) dockings, solid diamonds show the genetic algorithm (GA) dockings, and the solid squares show the Lamarckian genetic algorithm (LGA) dockings. Note the outlying biotin–streptavidin complex (1stp), where it is believed there are significant contributions to the binding free energy due to protein rearrangements.

ating the volume buried for only the carbon atoms in the ligand. Model C also uses only ligand carbon atoms in the desolvation calculation, and also adds a constant term to the hydrogen bonding function, modeling desolvation of polar atoms. Model C was chosen for incorporation into AUTODOCK 3.0, based on its better overall statistics, and on criteria discussed in what follows. The form of this free energy function is:

$$\begin{aligned}\Delta G = {} & \Delta G_{\mathrm{vdW}} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \\ & + \Delta G_{\mathrm{hbond}} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} + E_{\mathrm{hbond}} \right) \\ & + \Delta G_{\mathrm{elec}} \sum_{i,j} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \\ & + \Delta G_{\mathrm{tor}} N_{\mathrm{tor}} \\ & + \Delta G_{\mathrm{sol}} \sum_{i_C,j} S_i V_j \mathrm{e}^{(-r_{ij}^2/2\sigma^2)} \end{aligned} \quad (2)$$

where E_{hbond} is the estimated average energy of hydrogen bonding of water with a polar atom, and the summation in the solvation term is performed over all pairs consisting of only carbon atoms in the ligand, i, and atoms of all types, j, in the protein. Note that the internal or intramolecular interaction energy of the ligand is *not* included in the calculation of binding free energy; during docking, however, internal energy *is* included in the total docked energy, because changes in ligand conformation can affect the outcome of the docking, so this must be taken into consideration. We looked at linear regression models that did include the internal energy, and found that adding this term did not improve the model. The assumption made is that the internal energy of the ligand in solution and in the complex are the same. The energies used and reported by AUTODOCK should be distinguished: there are *docked energies*, which include the intermolecular and intramolecular interaction energies, and which are used during dockings; and *predicted free energies*, which include the intermolecular energy and the torsional free energy, and are only reported at the end of a docking. Because the intermolecular energy grid maps include the desolvation term, dockings using the new, empirical force field in AUTODOCK version 3.0 may be qualitatively different from results found using earlier versions.

Three coefficients, for dispersion/repulsion, electrostatics, and loss of torsional freedom were very stable in the linear regression analysis, with consistent coefficient values in different formulations and reasonable standard deviations. In our best model, dispersion/repulsion energies, with parameters taken from AMBER,[34] were weighted by a factor of 0.1485, yielding an energy of about -0.2 kcal mol^{-1} for the most favorable atom–atom contacts. Electrostatics, modeled with a screened Coulomb potential,[56] were weighted by a factor of 0.1146, yielding an energy of about -1.0 kcal mol^{-1} for an ideal salt bridge. In the torsional restriction term, each torsional degree of freedom requires 0.3113 kcal mol^{-1}.

The major differences between models occurred with the interaction of the hydrogen bonding term and the desolvation term. Hydrogen bonding is modeled with a directional 12–10 potential.[68] We encountered a major problem when calibrating this hydrogen bonding function. Because the test set included only natural enzyme–ligand complexes, optimized by millions of years of evolution, hydrogen bonding groups in the ligands are nearly always paired with the appropriate hydrogen bonding group in the protein. Thus, the number of hydrogen bonds that the ligand forms in the complex and the number it forms with solvent when free in solution are approximately the same; that is, there is little change in the free energy of hydrogen bonding, and $\Delta G_{\mathrm{hbond}}$ was evaluated to be approximately zero. Unfortunately, this provides no information on the cost of burying a hydrogen bonding group without forming a bond with the protein, and our data set did not include cases to evaluate this. Of course, the volume-based solvation method should account for this—the unfavorable polar contribution to the solvation energy should compensate for the favorable 12–10 hydrogen bonding energy. The linear regression, however, consistently returned coefficients that set the hydrogen bonding energy and desolvation energy to nearly zero, and increased the dispersion/repulsion term to compensate (see Model A in Table III). We chose an alternative formulation to resolve this problem.

We obtained the best results by separating the desolvation of polar atoms from the volume-based calculation. We assumed that the extent of hydrogen bonding in the complexes was roughly the same as the extent of hydrogen bonding in solution. The calculated hydrogen bonding energy, using the directional 12–10 hydrogen bonding function, was divided by the maximal number of possible hydrogen bonds, counting two for each oxygen atom and one for each polar hydrogen. For

the 30 complexes used in calibration, it was found that 36% of the maximum possible hydrogen bonding sites were actually utilized. Values of E_{hbond} ranging from 36% to 100% of the maximal well depth of 5 kcal mol^{-1} had little effect on the success of the formulation (data not shown), and a value of 36% was chosen. Optimized weights yielded an ideal hydrogen bonding energy in the complex of -0.328 kcal mol^{-1}, and the estimated average energy of each hydrogen bond in solution of -0.118. Because hydrogen bonding was modeled by this difference, the typical hydrogen bonding free energy of a complex was approximately zero, but there was a penalty of about 0.2 kcal mol^{-1} for oxygen and nitrogen atoms that did not form hydrogen bonds, driving the simulation toward docked conformations with maximal hydrogen bonding. We are currently exploring an appropriate data set for evaluating this formulation more rigorously.

The desolvation for carbon atoms in the ligand was evaluated using two different classes of atom type, aliphatic and aromatic, as in the original study.[58] The desolvation term was weighted by a factor of 0.1711 in our final empirical free energy force field, so a typical aliphatic carbon atom yields an energy of about -0.2 kcal mol^{-1} upon binding.

We cross-validated the free energy model in two ways. First, we investigated the influence of each member of the training set on the final coefficients of the model, by removing each one from the training set and calculating the coefficients from the remaining 29 complexes. We found that none of them had a strong effect on the final values of the coefficients.

We also performed a second kind of cross-validation of the free energy model, by performing Solis and Wets local search using AUTODOCK and the new free energy function, starting from the x-ray crystallographic conformations of each inhibitor in 20 HIV-1 protease-inhibitor complexes, to compare the resulting optimized conformations' predicted free energy change of binding, $\Delta G_{\mathrm{binding}}$, with the experimentally determined values. These protease inhibitors were quite different, having from 7 to 28 torsions, and widely different side chains—charged, polar, and hydrophobic, and constituted a diverse test set. As can be seen from the results in Table IX, the correlation was very good, with an overall rmsd between the experimental and calculated values of $\Delta G_{\mathrm{binding}}$ of 1.92 kcal mol^{-1}.

The final form of the free energy function may seem overparameterized, with additional weighting parameters added to a previously optimized parameterization. We retained the molecular mechanics formulation, however, specifically for its ability to model the distance dependence of each energetic term. This distance dependence (and angular dependence in hydrogen bonding) is essential for finding valid docked conformations, but the amount and resolution of the available protein–ligand data do not support a full re-pameterization of the functions.

DOCKING EXPERIMENTS

Because we are comparing different search methods, it is important to ensure that the methods are treated equally. It is therefore important that each search method be allowed approximately the same number of energy evaluations in a docking. The number of energy evaluations in a docking depends on the termination criteria, and because it is not possible to predict how many accepted or rejected steps the stochastic SA method will make at a given temperature, the number of evaluations varies in SA. The range was from 1.19×10^6 to 2.33×10^6, depending on the protein–ligand test system, even though the same parameters were used for the number of cycles, accepted steps and rejected steps. In the case of the GA dockings, the population was 50 and the number of generations was 27,000, which gave a total of 1.35×10^6 energy evaluations in a docking; thus, the GA dockings were terminated by reaching the maximum number of generations. In the case of the LGA dockings, 6% of the population underwent Lamarckian local search, each search consisting of 300 iterations and each iteration using an extra energy evaluation. Thus, the LGA dockings, even with the same population size and number of generations as the GA, were terminated by reaching the maximum number of energy evaluations, 1.50×10^6.

The results of the simulated annealing, genetic algorithm, and the Lamarckian genetic algorithm docking experiments are summarized in Tables IV, V and VI, respectively. The lowest energy docked structure found by each method is compared with the crystal structure of the ligand in Figure 4. The predicted change in free energy upon binding, ΔG_{pred}, for the lowest energy found by LGA is shown in Table VII, along with the experimentally observed change in free energy upon binding, ΔG_{obs}. In addition, the breakdown of the energy of the lowest energy docked conformation is shown, in terms of the intermolecular interaction energy,

TABLE IV.
Results of Simulated Annealing Dockings.[a]

PDB code	Number of clusters	Number in rank 1	Energy (kcal mol^{-1}) and rmsd (Å)						Number of energy evaluations
			Lowest energy	rmsd of lowest energy	Mean energy		Mean rmsd		
3ptb	5	6	−8.03	0.21	−7.84	(0.08)	0.50	(0.17)	2.01×10^6
2cpp	4	6	−7.29	0.81	−7.22	(0.03)	0.91	(0.30)	2.33×10^6
2mcp	10	1	−4.09	0.88	70.89	(2.10×10^2)	5.40	(4.80)	1.85×10^6
1stp	10	1	−8.48	1.27	−7.71	(0.66)	1.24	(0.35)	2.00×10^6
1hvr	10	1	−11.77	1.15	1.12×10^5	(3.36×10^5)	6.13	(2.61)	1.19×10^6
4hmg	10	1	−2.59	3.77	6.99×10^4	(1.52×10^5)	6.20	(2.94)	1.55×10^6
4dfr	10	1	−8.73	4.83	6.13×10^2	(1.96×10^3)	5.04	(1.74)	1.30×10^6

[a] The parameters used were 10 runs, 50 cycles, and a cycle-termination criterion of 25,000 accepted steps or 25,000 rejected steps, whichever came first. The rmsd conformational clustering tolerance was 0.5 Å, calculated from the ligand's crystallographic coordinates. Standard deviations given in parentheses.

ΔG_{inter} the intramolecular energy, ΔG_{intra}, and the torsional free energy, ΔG_{tor}. These results are discussed case-by-case in what follows; ''crystallographic rmsd'' refers to the root-mean-square positional deviation of a given conformation from the crystallographic coordinates.

β-Trypsin / Benzamidine (3ptb)

The recognition of benzamidine by β-trypsin, which binds tightly in the specificity pocket of trypsin, is chiefly due to the polar amidine moiety and the hydrophobic benzyl ring.[69] The amidine moiety was treated as being protonated. It was assumed that delocalization of the π-electrons of the benzene ring extended to the π-system of the amidine, and thus the ligand was treated as a rigid body. All three search methods succeeded in finding lowest energy conformations that were also the ones with the lowest crystallographic rmsd. In this case, the method that found the docked structure with the lowest energy was GA (Table V), but that found by the LGA method (Table VI) was practically the same. The mean of the final docked

TABLE V.
Results of Genetic Algorithm Dockings.[a]

PDB code	Number of clusters	Number in rank 1	Energy (kcal mol^{-1}) and rmsd (Å)						Number of energy evaluations
			Lowest energy	rmsd of lowest energy	Mean energy		Mean rmsd		
3ptb	2	9	−8.17	0.32	−7.72	(1.35)	1.50	(3.39)	1.35×10^6
2cpp	4	7	−7.36	0.93	−6.65	(2.11)	2.18	(3.42)	1.35×10^6
2mcp	10	1	−5.17	0.85	−3.61	(0.95)	5.26	(2.98)	1.35×10^6
1stp	7	4	−10.09	0.75	−8.42	(1.82)	2.96	(3.04)	1.35×10^6
1hvr	7	4	−21.41	0.82	−11.09	(9.79)	2.79	(1.97)	1.35×10^6
4hmg	9	2	−7.60	1.11	−5.72	(1.77)	2.32	(1.43)	1.35×10^6
4dfr	10	1	−16.10	0.95	−10.24	(3.95)	4.39	(2.37)	1.35×10^6

[a] The parameters used were 10 runs a population size of 50, and a run-termination criterion of a maximum of 27,000 generations or a maximum of 1.5×10^6 energy evaluations, whichever came first. Note that, in this case all runs terminated after the maximum number of generations was reached, which equals the product of the population size and the number of generations. The rmsd conformational clustering tolerance was 0.5 Å, calculated from the ligand's crystallographic coordinates. Standard deviations are given in parentheses.

TABLE VI.
Results of Lamarckian Genetic Algorithm Dockings.[a]

PDB code	Number of clusters	Number in rank 1	Energy (kcal mol^{-1}) and rmsd (Å)						Number of energy evaluations
			Lowest energy	rmsd of lowest energy	Mean energy		Mean rmsd		
3ptb	1	10	−8.15	0.45	−8.15	(0.00)	0.46	(0.01)	1.50×10^6
2cpp	1	10	−7.36	0.93	−7.36	(0.00)	0.93	(0.00)	1.50×10^6
2mcp	6	2	−5.54	1.05	−4.15	(0.15)	1.10	(0.07)	1.50×10^6
1stp	1	10	−10.14	0.69	−10.06	(0.05)	0.66	(0.06)	1.50×10^6
1hvr	2	9	−21.38	0.76	−19.11	(6.92)	0.85	(0.35)	1.50×10^6
4hmg	3	7	−7.72	1.14	−7.54	(0.19)	1.18	(0.12)	1.50×10^6
4dfr	2	7	−16.98	1.03	−16.90	(0.07)	0.98	(0.07)	1.56×10^6

[a] The parameters used were 10 runs, a population size of 50, and a run-termination criterion of a maximum of 27,000 generations or a maximum of 1.5×10^6 energy evaluations, whichever came first. Because local search also uses energy evaluations, the total number of energy evaluations for the LGA method was greater than that for the GA method, using the same population size and maximum number of generations; in the LGA dockings, the runs terminated because the maximum number of energy evaluations was exceeded. The rmsd conformational clustering tolerance was 0.5 Å, calculated from the ligand's crystallographic coordinates. Standard deviations are given in parentheses.

energy across the ten dockings was lowest for LGA, followed by SA, and finally GA. This is reflected in a comparison of the mean rmsd of the docked conformation from the crystallographic structure for each of the methods: GA had the highest mean rmsd, followed by SA, and on average, the LGA produced conformations with the lowest crystallographic rmsd. Thus, considering their average performance, the best search method at finding the lowest energy and the lowest rmsd was the LGA. The predicted binding free energy, ΔG_{pred}, of the lowest docked energy structure obtained using the LGA method was −8.15 kcal mol^{-1} (Table VII), whereas the observed value, ΔG_{obs} was −6.46 kcal mol^{-1}: this is within the estimated error of the model.

Cytochrome P-450$_{\text{cam}}$ / Camphor (2cpp)

Camphor binds to the monooxygenase cytochrome P-450$_{\text{cam}}$ such that the 5-*exo* C—H bond is hydroxylated stereospecifically. The active site is deeply sequestered within the enzyme, and the crystal structure of the complex does not possess an obvious substrate access channel.[70] This buried active site presents a more challenging docking problem than 3ptb. Once bound, however, the substrate is ''tethered'' by a hydrogen bond that is donated from the Tyr-96 hydroxyl to the carbonyl oxygen of camphor, while the subtle complementarity of the pocket and the hydrophobic skeleton of camphor help to position the rest of the substrate. The lowest energy found was −7.36 kcal mol^{-1}, found by both the GA (Table V) and LGA methods (Table VI); SA's lowest energy was −7.29 kcal mol^{-1} (Table IV), which is practically the same. All methods found the crystallographic structure, SA succeeding in 9 of 10 dockings, GA in 7 out of 10 dockings, and LGA in all of the dockings (with success, once again, being measured as having a crystallographic rmsd of less than 1 Å). In all three search method cases, the lowest energy cluster was the most populated, with 6, 9, and 10 members using SA, GA, and LGA, respectively. The predicted binding free energy, ΔG_{pred}, of the lowest docked energy structure, was −7.36 kcal mol^{-1} using the LGA method (see Table VII), whereas the observed value, ΔG_{obs}, was −8.27 kcal mol^{-1}—once again, this was within the estimated error of the model.

McPC-603 / Phosphocholine (2mcp)

Antibody molecules bind their target antigens with exquisite specificity, having close complementarity between antigen and antibody surfaces, hydrogen bonding, van der Waals, and electrostatic interactions. Phosphocholine binds to Fab McPC-603,[71] and is an example of recognition is predominantly electrostatic in character, primarily due to the influence of Arg H52.[72] There is little conformational change in the side chains of Fab McPC-603 upon binding, as indicated from the unbound crystal structure. We allowed all four

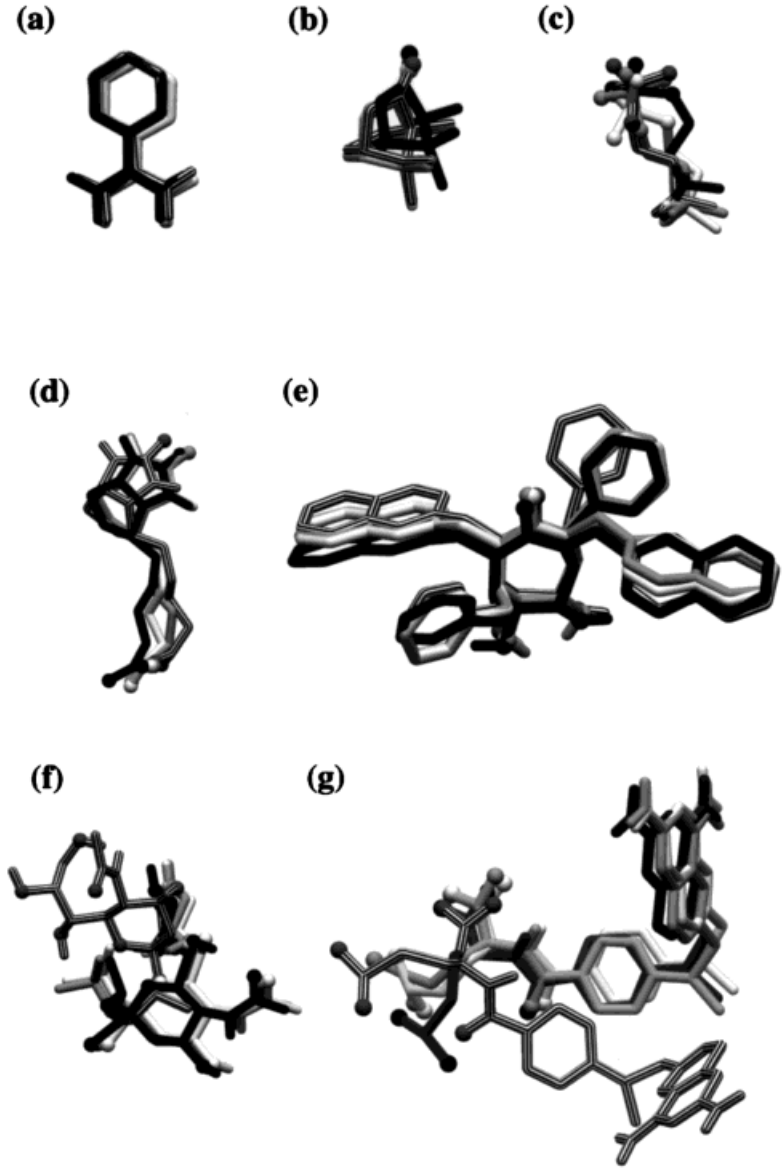

FIGURE 4. A comparison of the lowest energy structure found by each search method and the crystal structure. The latter is shown in black. The simulated annealing results are rendered with a striped texture, the genetic algorithm results are shaded gray, and the Lamarckian genetic algorithm results are white. Oxygen atoms are shown as spheres; other heteroatoms are not shown. Note that simulated annealing failed in the last two test cases, 4hmg and 4dfr, but both the genetic algorithm and the Lamarckian genetic algorithm succeeded.

bonds to rotate during docking. The energy of the crystal structure was positive, due in most part to a large, positive internal energy dominated by C2 and O1 being too close (2.26 Å); this could be improved if local minimization had been performed on the crystal structure before docking. The lowest energy found by each of the three search methods were -4.09, -5.17, and -5.54 kcal mol^{-1}, using SA, GA, and LGA, respectively. Unlike 3ptb and 2cpp, these differences in energy were more significant. Both SA and GA found 10 different clusters, whereas LGA found 6 clusters. The mean energy of the 10 dockings was $+70.89$, -3.61, and -4.15 kcal mol^{-1} for SA, GA, and LGA, respectively. Thus, on average, the LGA performed best in finding the lowest energy docked structure. Furthermore, the mean rmsd from the crystallographic coordinates was 5.40, 5.26, and 1.10 Å for SA, GA, and LGA, respectively, indicating that LGA also reproduced the crystal structure

TABLE VII.
Comparison of Predicted Free Energy of Binding, ΔG_{pred}, of Lowest Energy Docked Structure Obtained Using Lamarckian Genetic Algorithm, and Observed Free Energy of Binding, ΔG_{obs}.[a]

	Energy (kcal mol^{-1}) and rmsd (Å)							
PDB code	Lowest energy	rmsd of lowest energy	ΔG_{inter}	ΔG_{intra}	ΔG_{tor}	ΔG_{pred}	ΔG_{obs}	$(\Delta G_{pred} - \Delta G_{obs})$
3ptb	−8.15	0.45	−8.15	0.00	0.00	−8.15	−6.46	−1.69
2cpp	−7.36	0.93	−7.36	0.00	0.00	−7.36	−8.27	+0.91
2mcp	−5.54	1.05	−6.57	+1.03	+1.25	−5.32	−7.13	+1.81
1stp	−10.14	0.69	−9.90	−0.24	+1.56	−8.34	−18.27	+9.93[b]
1hvr	−21.38	0.76	−19.34	−2.04	+2.49	−16.85	−12.96	−3.89
4hmg	−7.72	1.14	−8.93	+1.21	+2.18	−6.75	−3.48	−3.27
4dfr	−16.97	1.03	−16.57	−0.40	+2.18	−14.39	−13.22	−1.17

[a] ΔG_{inter} is the intermolecular interaction energy between the ligand and the receptor, ΔG_{intra} is the intramolecular interaction energy of the ligand, and ΔG_{tor} is the torsional free energy change of the ligand upon binding.
[b] This large discrepancy may be due to neglect of the conformational rearrangements of streptavidin upon binding biotin, which are neglected in the docking simulation and binding free energy calculation.

most often. The predicted binding free energy, ΔG_{pred}, of the lowest docked energy structure was −5.32 kcal mol^{-1} using the LGA method (see Table VII), whereas the observed value, ΔG_{obs}, was −7.13 kcal mol^{-1}—this was also within the estimated error of the model.

Streptavidin / Biotin (1stp)

One of the most tightly binding noncovalent complexes is that of streptavidin/biotin, with an experimentally observed dissociation constant, K_d, of $10^{-15}M$. Comparison of the *apo* form and the complex[73] shows that the high affinity results from several factors, including formation of multiple hydrogen bonds and van der Waals interactions between the biotin and the protein, in addition to the ordering of surface polypeptide loops of streptavidin upon binding biotin. The method that found the lowest energy was LGA, at −10.14 kcal mol^{-1}, although GA was not significantly different, followed by SA with −8.48 kcal mol^{-1}. The method with the lowest mean energy was LGA at −10.06 kcal mol^{-1}, then GA with −8.42 kcal mol^{-1}, and finally SA with −7.76 kcal mol^{-1}. The method that found the crystallographic complex coordinates most often was LGA, having a mean rmsd of 0.66 Å, then SA at 1.24 Å, and finally GA with 2.96 Å. At the rmsd tolerance chosen for these experiments, 0.5 Å, SA found 10 different conformational clusters, GA found 7 clusters (the most populated was rank 1, with 4 members), and LGA found 1 cluster.

It was not possible to include the entropic effects of the flexible surface loops of streptavidin in the docking of biotin, although they make significant contributions to the binding free energy as revealed by a recent set of experiments involving an atomic force microscope.[74] It was found that the unbinding forces of discrete complexes of streptavidin with biotin analogs were proportional to the enthalpy change of the complex formation but independent of changes in the free energy, which indicates that the unbinding process is adiabatic and that entropic changes occur after unbinding. This may help to explain why the predicted binding free energy of the streptavidin/biotin complex (ΔG_{pred}) −10.14 kcal mol^{-1}, underestimated the magnitude of the observed value (ΔG_{obs}) −18.27 kcal mol^{-1} (Table VII).

HIV-1 Protease / XK263 (1hvr)

HIV-1 protease inhibitors prevent the maturation of virions of HIV, and are a major target for computer-assisted drug design in the development of AIDS therapies. Substrates and inhibitors of HIV-1 protease are typically extended peptides or peptidomimetics, with a dozen or more freely rotatable bonds and, as such, they present a challenging target for automated docking techniques. In addition, considerable protein motion is expected in the flaps upon binding, to allow the continuous polypeptide to reach the active site. However, most docking methods use a rigid protein target, and explicit modeling of the opening

and closing of the flaps is not performed: thus, the ligand must ''thread'' its way into the active site. The cyclic urea HIV-protease inhibitor, XK-263, has 10 rotatable bonds, excluding the cyclic urea's flexibility. All three search methods found solutions near to the crystal structure[75]: interestingly, the lowest docking energy found by SA was -11.77 kcal mol^{-1} and had an rmsd of 1.15 Å from the crystal structure, whereas GA and LGA found much lower energies but were still near to the active site, having crystallographic rmsd values of 0.82 Å and 0.76 Å, respectively. The lowest docking energy found overall was -21.41 kcal mol^{-1}, and was found using GA, although that found by LGA was practically the same. The predicted binding free energy, ΔG_{pred}, of the lowest energy structure was -16.85 kcal mol^{-1} using LGA, whereas the observed value, ΔG_{obs}, was -12.96 kcal mol^{-1}. The larger discrepancy between the predicted and observed values may be due to the entropic contributions of protein side chain and flap conformational rearrangements, or may be due to other low-energy conformational states of the cyclic urea moiety of XK-263, which are neglected in our calculations.

Influenza Hemagglutinin / Sialic Acid (4hmg)

The recognition of sialic acid by influenza hemagglutinin is chiefly mediated through hydrogen bonding: sialic acid has five hydroxyls, three in the glycerol group, one carboxylate, a cyclic ether oxygen, and an acetamido group, with a total of 11 rotatable acyclic bonds. We used the crystal structure of Weis et al.,[76] although the low resolution meant that the overall coordinate error was approximately 0.35–0.40 Å, which, in itself, presents a potential challenge in the docking tests. We modeled an isopropylated derivative of sialic acid to mimic part of an adjacent six-membered ring that would normally be present in this complex, but was not seen due to disorder: this introduced an extra rotatable bond, giving a total of 11 torsions. Furthermore, in these tests, we used the crystal conformation of the six-membered ring, although normally we would use several of the lowest energy conformations of the ring system and dock these separately.

This was one of two cases where simulated annealing failed to find a docking that was near the crystal structure: the lowest energy structure found had an rmsd of 3.77 Å from the crystallographic structure, and the docking with the lowest crystallographic rmsd was 2.36 Å. The mean energy of all 10 SA dockings was very high (6.99×10^4 kcal mol^{-1}); only 4 of the 10 SA dockings found negative energies. Both GA and LGA, however, succeeded in finding conformations near (≤ 1.5 Å rmsd) the crystal conformation. The lowest energy found was by LGA, and was -7.72 kcal mol^{-1}. This structure had a crystallographic rmsd of 1.14 Å, and had a predicted binding free energy, ΔG_{pred}, of -6.75 kcal mol^{-1}; the observed binding free energy, ΔG_{obs}, for the sialic acid–hemagglutinin complex was -3.48 kcal mol^{-1}. The difference in predicted and observed binding free energies may be due to the structural differences between the isopropylated derivative that was docked and sialic acid itself.

Dihydrofolate Reductase / Methotrexate (4dgr)

Methotrexate is an antimetabolite that attacks proliferating tissue selectively induces remissions in certain acute leukemias[77]; however, dangerous side effects of methotrexate in normal cells continue to make DHFR an important target in computer-assisted anticancer drug design.[78] We used the crystal structure of *E. coli* dihydrofolate reductase complexed with methotrexate[79] to investigate a more challenging docking problem. We assumed that waters 603, 604, and 639, which mediate hydrogen bonding between the inhibitor and the protein, were conserved biowaters, and included them in the protein structure in our grid calculations. Ideally, these should be predicted, and recently a method based on a *k*-nearest-neighbors classifier and a genetic algorithm called *Consolv* was reported to do just this.[80]

This is one of the two test cases where simulated annealing failed: the lowest energy structure that it found had an rmsd of 4.83 Å from the crystal structure. This could be because the final docked conformation in simulated annealing is arrived at after a series of continuous steps, and if the route to the active site is blocked, the docking will tend to fail before the ligand reaches the active site. Note that, in the case of the camphor–cytochrome P-450$_{\text{cam}}$ docking, the random initialization loop was able to find initial states that were inside the binding pocket, but in this case the dockings failed to start near the active site.

The lowest energy found was -16.98 kcal mol^{-1}, and was found using LGA: this structure had an rmsd from the crystal structure of 1.03 Å.

The predicted binding free energy, ΔG_{pred}, of the lowest docked energy structure was -14.39 kcal mol^{-1} using LGA, whereas the observed value, ΔG_{obs}, was -13.22 kcal mol^{-1}. This finding was within the estimated error of the model.

JUDGING SEARCH METHODS

To evaluate the new search methods, and to compare them with the earlier search method of simulated annealing, we addressed the following questions: Which search method is most *efficient*? That is, which finds the lowest energy in a given number of energy evaluations? Which search method is most *reliable*? That is, which method finds the most conformations similar to that of the lowest energy? Finally, which search method is most *successful*? That is, which finds the crystallographic conformation most often after a given number of dockings? Furthermore, because these comparisons were carried out using the new, empirical free energy force field, these tests also represent an evaluation of the force field itself, and, if the global minimum of the force field is unable to reproduce observed crystallographic structures, its usefulness will be limited. Because it is very difficult to determine the global minimum of such a complex function, we cannot answer this question definitively; however, we can report the lowest energy found by any of the methods and its structural similarity to that of the crystal structure.

If we calculate statistics across all seven protein–ligand test systems for each search method, we obtain a quantitative estimate of relative performance of each search method (see Table VIII). If, in each of the seven test systems, we assume that the lowest docked energy found by any method is the effective global minimum energy, and then calculate the difference between this energy and all of the docked energies found by each search method, we can then calculate the mean and standard deviation of this difference energy for each search method. Ideally, the mean and standard deviation of this value would be zero. The mean of this difference energy was lowest for LGA (0.40 kcal mol^{-1}), followed by GA (3.41 kcal mol^{-1}), and finally SA (2.62×10^5 kcal mol^{-1}): the very high mean difference energy for SA is indicative of the cases in which this method failed to escape a local minimum, where the ligand was partially or wholly trapped within the protein. Hence, in answer to the first question, the Lamarckian genetic algorithm, LGA, is the most efficient search method.

In terms of how often the structure with lowest energy was found, LGA performed best: the mean of the number of docked structures in rank 1 was 78% for LGA, 40% for GA, and 24% for SA. The mean of the number of clusters found was lowest for LGA (2.29), followed by GA (7.00), and finally SA (8.43). Hence, the most reliable search method was LGA.

In comparing the relative success of each search method in reproducing the crystallographic structure, considering the crystallographic rmsd across all 10 dockings in each of the 7 test systems, the mean rmsd was lowest for LGA (0.88 Å, standard

TABLE VIII.
Statistical Comparison of Three Search Methods in AutoDock 3.0 Across all Seven Test Systems.[a]

				Energy (kcal mol^{-1}) and rmsd (Å)				
Search method	Statistic	Number of clusters	Number in rank 1	Difference from effective global minimum energy	rmsd of lowest energy	Mean energy	Mean rmsd	Number of energy evaluations
SA	Mean	8.43	2.43	2.62×10^5	1.85	2.61×10^5	3.63	1.75×10^6
	SD	2.70	2.44	1.40×10^5	1.74	4.60×10^4	2.61	4.15×10^5
GA	Mean	7.00	4.00	3.41	0.82	-7.64	3.06	1.35×10^6
	SD	3.06	3.06	5.31	0.25	2.59	1.32	0.00
LGA	Mean	2.29	7.86	0.40	0.86	-10.47	0.88	1.50×10^6
	SD	1.80	2.91	2.62	0.24	5.47	0.25	0.00

[a] The search methods are simulated annealing (SA), genetic algorithm (GA), and Lamarckian genetic algorithm (LGA). The mean and standard deviation (SD) for each criterion is shown. The effective global minimum energy for each of the seven test systems is the lowest docked energy found by any method for that test system. For each of the 10 dockings, the difference between the final docked energy and this effective global minimum energy was calculated; the mean and standard deviation was calculated across all 7 test systems, which was repeated for each search method.

TABLE IX.
Results of Cross-Validation of Free Energy Function Using Local Search on 20 HIV-1 Protease-Inhibitor Complexes.

PDB code	Experimental $\Delta G_{binding}$ (kcal / mol)	Calculated $\Delta G_{binding}$ (kcal / mol)
1hvs	−14.04	−10.95
1hvk	−13.79	−11.60
1hvi	−13.74	−12.39
7hvp	−13.11	−12.19
1hps	−12.57	−11.80
1hpv	−12.57	−8.24
4phv	−12.51	−14.36
1hef	−12.27	−9.52
1hiv	−12.27	−13.02
1hvl	−12.27	−10.35
8hvp	−12.27	−9.36
1aaq	−11.62	−9.68
1htg	−11.58	−13.13
9hvp	−11.38	−10.54
1hih	−10.97	−11.43
1heg	−10.56	−8.60
1sbg	−10.56	−10.35
1htf	−9.31	−8.21
1hbv	−8.68	−9.75
1hte	−7.69	−7.28

deviation 0.25 Å), followed by GA (3.06 Å, standard deviation 1.32 Å), and finally SA (3.63 Å, standard deviation 2.61 Å). These average results indicate that, of the three search methods, LGA will find the crystallographic structure most often. Thus, the answer to the last question, ''Which method is most successful?,'' is LGA.

In two different cases, 4hmg and 4dfr, the simulated annealing method failed to reproduce the corresponding crystal structure, although it succeeded with 1hvr (see Fig. 4). This is important because methotrexate has 7 rotatable bonds, and would be expected to be solvable using our rule-of-thumb that SA succeeds in problems with 8 torsions or less; however, the HIV-1 protease inhibitor XK-263, has 10 rotatable bonds, and was successfully docked using SA. Thus, the degree of difficulty of a docking problem is not as simple as how many rotatable bonds there are; other factors, such as the nature of the energy landscape, clearly play an important role.

It could be said that the crystallographic rmsd of the lowest energy structure found by any of the search methods is an estimate of the quality of the force field, although this is complicated by the fact that the search method itself must determine a docking near to the global minimum, an unknown state. We can calculate the energy of the ligand in the crystal structure using the new force field (see Table II), which we assume to be near the global minimum, but, unfortunately, the crystal structure may contain frustrations and bad contacts. This appears to be the case in 2mcp, where a close contact between C2 and O1 causes a positive total energy to be calculated for the crystal structure. In all cases, the lowest energy found, considering all the search methods, was lower than that of the corresponding crystal structure.

The crystallographic rmsd of the lowest energy (found by any search method) for each of the protein–ligand test systems were all within 1.14 Å, or less, of the crystal structure. This suggests that the force field's global minimum in each of the protein–ligand cases was near to the crystal structure, if we accept the assumption that the crystal structure was near to or at the global minimum, and that the lowest energy found was near to the global minimum. In some cases, dockings were found that had lower crystallographic rmsd values but slightly higher energies than the lowest energy found. All of the lowest crystallographic rmsd values were 0.89 Å or less, indicating that low-energy structures found by the force field were very similar to the corresponding crystal structure.

Conclusion

AUTODOCK is a software package of general applicability for automated docking of small molecules, such as peptides, enzyme inhibitors, and drugs, to macromolecules, such as proteins, enzymes, antibodies, DNA, and RNA. New search methods have been introduced and tested here, using a new, empirical binding free energy function for calculating ligand–receptor binding affinities.

We have shown that, of the three search methods tested in AUTODOCK (simulated annealing, genetic algorithm, and Lamarckian genetic algorithm), the most efficient, reliable, and successful is the Lamarckian genetic algorithm LGA. We defined efficiency of search in terms of lowest energy found in a given number of energy evaluations; reliability in terms of reproducibility of finding the

lowest energy structure in independent dockings, as measured by the number of conformations in the top ranked cluster; and success in terms of reproducing the known crystal structure. Simulated annealing failed to reproduce the crystal structures for the influenza hemagglutin–sialic acid complex (4hmg) and the dihydrofolate reductase–methotrexate complex (4dfr). However, both the genetic algorithm and the Lamarckian genetic algorithm methods succeeded. Thus, the introduction of the LGA search method extends the power and applicability of AUTODOCK to docking problems with more degrees of freedom than could be handled by earlier versions.

The predicted binding affinities of the lowest energy docked conformations, using the LGA method and the new empirical free energy function, were within the standard residual error of the force field in four of the seven cases (3ptb, 2cpp, 2mcp, and 4dfr), and reasonably close in two other cases (1hvr and 4hmg). The large discrepancy between the predicted and the observed binding affinity of biotin for streptavidin (1stp), even though the crystal structure was successfully reproduced, may be due to the large free energy change that accompanies conformational changes in the protein upon binding, in particular the surface loops. This remains a limitation of the method, because protein motion is not modeled and successfully predicting such large-scale protein conformational changes is difficult. The AUTODOCK method works well when there is little change between the *apo* and ligand-bound forms of the protein, even if the protein undergoes significant conformational changes during binding.

AUTODOCK predicts the binding affinity using one conformation of the ligand–protein complex. A new class of models for predicting receptor–ligand binding affinities has been reported recently that considers not just the *lowest* energy state of the complex, but the *predominant states* of the binding molecules.[81] These approaches are grounded in statistical thermodynamics, and combine a modest set of degrees of freedom with aggressive conformational sampling to identify the low-energy conformations of the complex and the free molecules. AUTODOCK version 3.0 currently performs extensive conformational sampling, information that could be incorporated into the calculation of the binding affinity. We are studying how the search methods can be modified such that statistical thermodynamics calculations can be performed while the docking proceeds, to improve the calculation of the binding affinity.

Availability

More information about AUTODOCK and how to obtain it can be found on the World Wide Web at: http://www.scripps.edu/pub/olson-web/doc/autodock.

Acknowledgments

The authors thank Dr. Bruce S. Duncan and Dr. Christopher Rosin for their helpful comments and suggestions. This work is publication 10887-MB from The Scripps Research Institute.

References

1. J. M. Blaney, and J. S. Dixon, *Perspect. Drug Discov. Design*, **1**, 301 (1993).
2. I. D. Kuntz, E. C. Meng, and B. K. Shoichet, *Acc. Chem. Res.*, **27**, 117 (1994).
3. R. Rosenfeld, S. Vajda, and C. DeLisi, *Annu. Rev. Biophys. Biomol. Struct.*, **24**, 677 (1995).
4. I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, and T. E. Ferrin, *J. Mol. Biol.*, **161**, 269 (1982).
5. B. K. Shoichet and I. D. Kuntz, *Prot. Eng.*, **6**, 723 (1993).
6. D. S. Goodsell and A. J. Olson, *Prot. Struct. Func. Genet.*, **8**, 195 (1990).
7. G. M. Morris, D. S. Goodsell, R. Huey, and A. J. Olson, *J. Comput.-Aided Mol. Des.*, **10**, 293 (1996).
8. N. Pattabiraman, M. Levitt, T. E. Ferrin, and R. Langridge, *J. Comput. Chem.*, **6**, 432 (1985).
9. P. J. Goodford, *J. Med. Chem.*, **28**, 849 (1985).
10. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.*, **21**, 1087 (1953).
11. S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, *Science*, **220**, 671 (1983).
12. E. A. Lunney, S. E. Hagen, J. M. Domagala, C. Humblet, J. Kosinski, B. D. Tait, J. S. Warmus, M. Wilson, D. Ferguson, D. Hupe, P. J. Tummino, E. T. Baldwin, T. N. Bhat, B. Liu, and J. W. Erickson, *J. Med. Chem.*, **37**, 2664 (1994).
13. J. V. N. Vara Prasad, K. S. Para, D. F. Ortwine, J. B. Dunbar Jr., D. Ferguson, P. J. Tummino, D. Hupe, B. D. Tait, J. M. Domagala, C. Humblet, T. N. Bhat, B. Liu, D. M. A. Guerin, E. T. Baldwin, J. W. Erickson, and T. K. Sawyer, *J. Am. Chem. Soc.*, **116**, 6989 (1994).
14. A. R. Friedman, V. A. Roberts, and J. A. Tainer, *Prot. Struct. Func. Genet.*, **20**, 15 (1994).
15. B. L. Stoddard and D. E. Koshland, *Nature*, **358**, 774 (1992).
16. D. S. Goodsell, G. M. Morris, and A. J. Olson, *J. Mol. Recog.*, **9**, 1 (1996).

17. C. M. Oshiro, I. D. Kuntz, and J. S. Dixon, *J. Comput.-Aided Mol. Design*, **9**, 113 (1995).
18. D. R. Westhead, D. E. Clark, D. Frenkel, J. Li, C. W. Murray, B. Robson, and B. Waszkowycz, *J. Comput.-Aided Mol. Design*, **9**, 139 (1995).
19. P. Willet, *TIBTECH*, **13**, 516 (1995).
20. D. E. Clark and D. R. Westhead, *J. Comput.-Aided Mol. Design*, **10**, 337 (1996).
21. C. D. Rosin, R. S. Halliday, W. E. Hart, and R. K. Belew, In *Proceedings of the Seventh International Conference on Genetic Algorithms* (*ICGA97*), T. Baeck, Ed., Morgan Kauffman, San Francisco, CA, 1997.
22. D. K. Gehlhaar, G. M. Verkhivker, P. A. Rejto, C. J. Sherman, D. B. Fogel, L. J. Fogel, and S. T. Freer, *Chem. Biol.*, **2**, 317 (1995).
23. J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI, 1975.
24. S. S. Četverikov, *J. Exper. Biol.*, **2**, 3 (1926).
25. Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, New York, 1996.
26. W. E. Hart, *Adaptive Global Optimization with Local Search*, Ph.D. Thesis, Computer Science and Engineering Department, University of California, San Diego, 1994. See also: ''ftp://ftp.cs.sandia.gov/pub/papers/wehart/thesis.ps.gz.''
27. W. E. Hart, T. E. Kammeyer, and R. K. Belew, In *Foundations of Genetic Algorithms III*, D. Whitley and M. Vose, Eds., Morgan Kauffman, San Francisco, CA, 1994.
28. R. K. Belew and M. Mitchell, *Adaptive Individuals in Evolving Populations: Models and Algorithms. Santa Fe Institute Studies in the Science of Complexity, XXVI*, Addison-Wesley, Reading, MA, 1996.
29. F. J. Solis and R. J.-B. Wets, *Math. Oper. Res.*, **6**, 19 (1981).
30. P.-G. Maillot, In *Graphics Gems*, A. S. Glassner, Ed., Academic Press, London, 1990, p. 498.
31. A. Watt and M. Watt, In *Advanced Animation and Rendering Techniques—Theory and Practice*, ACM Press, New York.
32. P. L'Ecuyer and S. Cote, *ACM Trans. Math. Software*, **17**, 98 (1991).
33. J. B. Lamarck, *Zoological Philosophy*, Macmillan, London, 1914.
34. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta Jr., and P. Weiner, *J. Am. Chem. Soc.*, **106**, 765 (1984).
35. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.*, **117**, 5179 (1995).
36. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.*, **4**, 187 (1983).
37. A. T. Hagler, E. Huler, and S. Lifson, *J. Am. Chem. Soc.*, **96**, 5319 (1977).
38. G. Némethy, M. S. Pottle, and H. A. Scheraga, *J. Phys. Chem.*, **87**, 1883 (1983).
39. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. diNola, and J. R. Haak, *J. Chem. Phys.*, **81**, 3684 (1984).
40. J. H. van der Waals, *Lehrbuch der Thermodynamik, Part 1*, Mass and Van Suchtelen, Leipzig, 1908.
41. I. K. McDonald and J. M. Thornton, *J. Mol. Biol.*, **238**, 777 (1994).
42. M. K. Gilson and B. Honig, *Nature*, **330**, 84 (1987).
43. D. Bashford and M. Karplus, *Biochemistry*, **29**, 10219 (1990).
44. D. Bashford and K. Gerwert, *J. Mol. Biol.*, **224**, 473 (1992).
45. B. Honig and A. Nicholls, *Science*, **268**, 1144 (1995).
46. P. A. Bash, U. C. Singh, F. K. Brown, R. Langridge, and P. A. Kollman, *Science*, **235**, 574 (1987).
47. L. P. Hammett, *J. Am. Chem. Soc.*, **59**, 96 (1937).
48. T. Fujita, J. Iwasa, and C. Hansch, *J. Am. Chem. Soc.*, **86**, 5175 (1964).
49. C. Hansch, A. R. Steward, J. Iwasa, and E. W. Deutsch, *Mol. Pharmacol.*, **1**, 205 (1965).
50. C. D. Selassie, Z. X. Fang, R. L. Li, C. Hansch, G. Debnath, T. E. Klein, R. Langridge, and B. T. Kaufman, *J. Med. Chem.*, **32**, 1895 (1989).
51. D. H. Williams, J. P. L. Cox, A. J. Doig, M. Gardner, U. Gerhard, P. T. Kaye, A. R. Lal, I. A. Nicholls, C. J. Salter, and R. C. Mitchell, *J. Am. Chem. Soc.*, **113**, 7020 (1991).
52. L. Wesson and D. Eisenberg, *Prot. Sci.*, **1**, 227 (1992).
53. H.-J. Böhm, *J. Comput.-Aided Mol. Design*, **6**, 593 (1992).
54. H.-J. Böhm, *J. Comput.-Aided Mol. Design*, **8**, 243 (1994).
55. A. N. Jain, *J. Comput.-Aided Mol. Design*, **10**, 427 (1996).
56. E. L. Mehler and T. Solmajer, *Prot. Eng.*, **4**, 903 (1991).
57. M. F. Sanner, A. J. Olson, and J.-C. Spehner, *Biopolymers*, **38**, 305 (1996).
58. P. F. W. Stouten, C. Frömmel, H. Nakamura, and C. Sander, *Mol. Simul.*, **10**, 97 (1993).
59. S. Hill, In *Graphics Gems IV*, P. S. Heckbert, Ed., Academic Press, London, 1994, p. 521.
60. P. W. Atkins, *Physical Chemistry*, Oxford University Press, Oxford, 1982, p. 263.
61. H. R. Horton, L. A. Moran, R. S. Ochs, J. D. Rawn, and K. G. Scrimgeour, *Principles of Biochemistry*, Prentice-Hall, London, 1993.
62. S-PLUS, Statistical Sciences, Inc., Seattle, WA.
63. F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, *J. Mol. Biol.*, **112**, 535 (1977).
64. E. E. Abola, F. C. Bernstein, S. H. Byant, T. F. Koetzle, and J. Weng, In *Crystallographic Databases-Information Content, Software Systems, Scientific Applications*, F. H. Allen, G. Bergerhoff, and R. Sievers, Eds., Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester, 1987, p. 107.
65. SYBYL, Tripos Associates, Inc., St. Louis, MO.
66. M. Marsili and J. Gasteiger, *Chim. Acta*, **52**, 601 (1980).
67. J. Gasteiger and M. Marsili, *Tetrahedron*, **36**, 3210 (1980).
68. D. N. A. Boobbyer, P. J. Goodford, P. M. McWhinnie, and R. C. Wade, *J. Med. Chem.*, **32**, 1083 (1989).
69. M. Marquart, J. Walter, J. Deisenhofer, W. Bode, and R. Huber, *Acta Crystallogr.* (*Sect. B*), **39**, 480 (1983).
70. T. L. Poulos, B. C. Finzel, and A. J. Howard, *J. Mol. Biol.*, **195**, 687 (1987).
71. E. A. Padlan, G. H. Cohen, and D. R. Davies, *Ann. Immunol.* (*Paris*) (*Sect. C*), **136**, 271 (1985).

72. J. Novotny, R. E. Bruccoleri, and F. A. Saul, *Biochemistry*, **28**, 4735 (1989).

73. P. C. Weber, D. H. Ohlendorf, J. J. Wendolski, and F. R. Salemme, *Science*, **243**, 85 (1989).

74. V. T. Moy, E.-L. Florin, and H. E. Gaub, *Science*, **266**, 257 (1994).

75. P. Y. S. Lam, P. K. Jadhav, C. J. Eyerman, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Rayner, Y. Wong, C.-H. Chang, P. C. Weber, D. A. Jackson, T. R. Sharpe, and S. Erickson-Viitanen, *Science*, **263**, 380 (1994).

76. W. I. Weis, A. T. Brünger, J. J. Skehel, and D. C. Wiley, *J. Mol. Biol.*, **212**, 737 (1990).

77. C. K. Matthews and K. E. van Holde, *Biochemistry*, Benjamin/Cummings, Redwood City, CA, 1990.

78. C. A. Reynolds, W. G. Richards, and P. J. Goodford, *Anti-Cancer Drug Des.*, **1**, 291 (1987).

79. J. T. Bolin, D. J. Filman, D. A. Matthews, R. C. Hamlin, and J. Kraut, *J. Biol. Chem.*, **257**, 13650 (1982).

80. M. L. Raymer, P. C. Sanschagrin, W. F. Punch, S. Venkataraman, E. D. Goodman, and L. A. Kuhn, *J. Mol. Biol.*, **265**, 445 (1997).

81. M. K. Gilson, J. A. Given, and M. S. Head, *Chem. Biol.*, **4**, 87 (1997).