In Applied Computing 1996: Proceedings of the 1996 Symposium on Applied Computing, New York:
The Association for Computing Machinery, 1996.

HYBRID GENETIC ALGORITHMS FOR POLYPEPTIDE ENERGY MINIMIZATION

Laurence D. Merkle, Robert L. Gaulke, Gary B. Lamont
Department of Electrical and Computer Engineering
Graduate School of Engineering
Air Force Institute of Technology
Wright-Patterson AFB, OH 45433
{lmerkle, rgaulke, lamont}@afit.af.mil

George H. Gates, Jr., Ruth Pachter
Wright Laboratory
3005 P St., Ste. 1
Wright-Patterson AFB, OH 45433-7702
{gatesgh, pachterr}@ml.wpafb.af.mil

Key Words: Genetic Algorithms, Conjugate Gradient, Baldwinian Evolution, Probabilistically Lamarckian Evolution, Protein Folding.

ABSTRACT

Efforts to predict polypeptide structures nearly always assume that the native conformation corresponds to the global minimum free energy state of the system. Given this assumption, a necessary step in solving the problem is the development of efficient global energy minimization techniques. We describe a hybrid genetic algorithm which incorporates efficient gradient-based minimization directly in the fitness evaluation, which is based on a general full-atom potential energy model. The algorithm includes a replacement frequency parameter which specifies the probability with which an individual is replaced by its minimized counterpart. Thus, the algorithm can implement either Baldwinian, Lamarckian, or probabilistically Lamarckian evolution.

We also describe experiments comparing the effectiveness of the genetic algorithm with and without the local minimization operator, with various probabilities of replacement. The experiments apply the techniques to the minimization of the CHARMM potential for [Met]-Enkephalin.

When fitness proportionate selection is used, the Baldwinian, Lamarckian, and probabilistically Lamarckian approaches obtain better energies (and better basins of attraction) than the standard genetic algorithm. This suggests that the low-energy local minima in polypeptide energy landscapes occur sufficiently regularly to benefit from the proposed hybrid approaches. When tournament selection is used, the results are qualitatively similar, except that the hybrid approaches are prone to premature convergence. Increasing replacement frequency reduces the tendency toward premature convergence for the experiments performed here.

1 INTRODUCTION

The prediction of an arbitrary polypeptide's native conformation (i.e. molecular structure) given only its amino acid sequence is beyond current capabilities, but has numerous potential applications [3]. This structure prediction problem is commonly referred to as the *protein folding problem.* Efforts to solve it nearly always assume that the native conformation corresponds to the global minimum free energy state of the system. Given this assumption, a necessary step in solving the problem is the development of efficient global energy minimization techniques. This is a difficult optimization problem because of the non-linear and multi-modal nature of the energy function. The pentapeptide [Met]-Enkephalin, for example, is estimated to have more than 10^{11} locally optimal conformations. Energy minimization is discussed in slightly more detail in Section 2. Also, Vásquez et al. [27] recently reviewed the literature of

polypeptide conformational energy calculations.

One class of optimization algorithms which has been applied to the energy minimization problem is that of genetic algorithms (GAs), which are described elsewhere (e.g. Goldberg [8], Holland [10], or Michalewicz [16]). The energy models to which GAs have been applied vary from lattice representations [4, 26] to simplified continuum proteins [11, 12, 22, 23], fixed backbones [20, 25], polypeptide-specific full-atom models [13, 15], and general full-atom models [6, 20].

In some cases (e.g. [13, 25]), the genetic algorithm performs a search of conformations constructed from a library of frequently occurring locally optimal single residue conformations (*rotamers*). This approach may be viewed as a sequentially hybrid approach, in which efficient local optimization of single residue conformations precedes global optimization via genetic algorithm of the overall polypeptide conformation.

Similarly, McGarrah and Judson [15] use a build-up approach including step-wise local minimization to construct their initial population. Their hybrid algorithm also periodically performs local minimization, and uses the resulting energies as the fitnesses of the corresponding individuals. The individuals are never altered following the local minimization. This is in contrast to one of the algorithms studied earlier by Judson et al. [11] in which individuals are always replaced by their locally optimized structures. Unger and Moult [26] propose a hybrid, similar to the latter, in which each individual undergoes 20 steps of simulated annealing before selection is performed.

Here we describe a hybrid genetic algorithm which incorporates efficient gradient based minimization directly in the fitness evaluation, which is based on a general full-atom potential energy model (Section 2). The algorithm includes a *replacement frequency* parameter p_r which specifies the probability with which an individual is replaced by its minimized counterpart. Thus, the algorithm can implement either Baldwinian ($p_r = 0$) or Lamarckian ($p_r = 1$) evolution [28], or more generally probabilistically Lamarckian ($0 \leq p_r \leq 1$) evolution. We also describe experiments comparing the effectiveness of the genetic algorithm with and without the local minimization operator, and with various probabilities of replacement for the algorithm with the local minimization operator (Section 3). Conclusions are presented in Section 4, and Section 5 discusses directions for future research.

2 METHODOLOGY

In this section we discuss the objective function associated with our polypeptide energy minimization application (Section 2.1) as well as the encoding scheme (Section 2.2). We then discuss the local minimization technique (Section 2.3) which uses the analytical gradient. We outline the derivation of the gradient in Section 2.4. Finally, we discuss the hybridization of the local minimization technique with the genetic algorithm (Section 2.5).

2.1 Objective Function

Our objective function, which we seek to minimize, is based on the CHARMM [2] energy function

$$\begin{aligned} E = & \sum_{(i,j)\in\mathcal{B}} K_{r_{ij}}(r_{ij} - r_{eq})^2 + \\ & \sum_{(i,j,k)\in\mathcal{A}} K_{\Theta_{ijk}}(\Theta_{ijk} - \Theta_{eq})^2 + \\ & \sum_{(i,j,k,l)\in\mathcal{D}} K_{\Phi_{ijkl}}[1 + \cos(n_{ijkl}\Phi_{ijkl} - \gamma_{ijkl})] + \\ & \sum_{(i,j)\in\mathcal{N}} \left[\left(\frac{A_{ij}}{r_{ij}}\right)^{12} - \left(\frac{B_{ij}}{r_{ij}}\right)^{6} + \frac{q_i q_j}{4\pi\varepsilon r_{ij}} \right] + \\ & \frac{1}{2} \sum_{(i,j)\in\mathcal{N}'} \left[\left(\frac{A_{ij}}{r_{ij}}\right)^{12} - \left(\frac{B_{ij}}{r_{ij}}\right)^{6} + \frac{q_i q_j}{4\pi\varepsilon r_{ij}} \right] \end{aligned} \tag{1}$$

where the five terms (which we denote $E_{\mathcal{B}}$, $E_{\mathcal{A}}$, $E_{\mathcal{D}}$, $E_{\mathcal{N}}$, $E_{\mathcal{N}'}$) represent the energy due to bond stretching, bond angle deformation, dihedral angle deformation, non-bonded interactions, and 1-4 interactions, respectively. Specifically,

- $\mathcal{B}$ is the set of bonded atom pairs,
- $\mathcal{A}$ is the set of atom triples defining bond angles,
- $\mathcal{D}$ is the set of atom 4-tuples defining dihedral angles,
- $\mathcal{N}$ is the set of non-bonded atom pairs,
- $\mathcal{N}'$ is the set of 1-4 interaction pairs,
- r_{ij} is the distance between atoms i and j,
- Θ_{ijk} is the angle formed by atoms i, j, and k,
- Φ_{ijkl} is the dihedral angle formed by atoms i, j, k, and l,
- q_i is the partial atomic charges of atom i,
- the $K_{r_{ij}}$'s, r_{eq}'s, $K_{\Theta_{ijk}}$'s, Θ_{eq}'s, $K_{\Phi_{ijkl}}$'s, γ_{ijkl}'s, A_{ij}'s, B_{ij}'s, and ε are empirically determined constants (taken from the QUANTA parameter files).

The primary determinants of a protein's 3-D structure, and thus the energetics of the system, are its independent dihedral angles [27]. Our genetic algorithm operates on individuals which encode these dihedral angles [6]. In Equation 1, E is expressed as a function

of both the internal coordinates (bond lengths r_{ij} for $(i,j) \in \mathcal{B}$, bond angles Θ_{ijk}, and dihedral angles Φ_{ijkl}) and the interatomic distances r_{ij} for $(i,j) \in \mathcal{N} \cup \mathcal{N}'$. Thus, in order to calculate E (and hence the fitness) for the conformation encoded by an individual, it is necessary to calculate its Cartesian coordinates from its internal coordinates. We use the transformation method proposed by Thompson [24]. This method requires at most one 4×4 matrix multiplication per atom per conformation.

2.2 Encoding Scheme

Each individual is a fixed length binary string encoding the independent dihedral angles of a polypeptide conformation. The decoding function used is the affine mapping $D : \{0,1\}^{10} \rightarrow [-\pi, \pi]$ of 10 bit subsequences to dihedral angles such that

$$D(a_1, a_2, \ldots, a_{10}) = -\pi + 2\pi \sum_{j=1}^{10} a_j 2^{-j}. \qquad (2)$$

This encoding yields a precision of approximately one third of one degree.

The particular biomolecule investigated here is the pentapeptide [Met]-enkephalin. This molecule is chosen because it has been used as a test problem for many other energy minimization investigations (e.g. [13, 17]), and its minimum energy conformation is known (with respect to the ECEPP/2 energy model). Twenty-four dihedral angles determine [Met]-enkephalin's structure, hence the string length is 240.

2.3 Local Minimization

The objective function defined by Equation 1 is such that all of its second partial derivatives exist and are continuous almost everywhere.[1] We consider three local minimization techniques which exploit to varying degrees this smoothness property and the ready availability of software [19, pp. 422, 426].

1. **First derivative method.** Apply an appropriate first derivative method (e.g. conjugate gradient or quasi-Newton) directly to the local minimization of E.

2. **Critical point method.** Apply a first derivative method to the minimization of $\|\nabla E\|$ (or $\|\nabla E\|^2$). This is equivalent to solving for a point for which $\nabla E = 0$, i.e. a critical point of E.

3. **Exact second derivative method.** Apply either Newton's method or conjugate gradient with the exact Hessian to the minimization of E.

[1] That is, for each derivative the set of discontinuities is countable. It is in fact finite.

The first derivative method is the easiest of the three to implement, because it does not require second derivatives and it uses readily available local minimization software. It is also guaranteed to find a local minimum (as opposed to a critical point). This method converges quickly for our energy function. For the application to [Met]-Enkephalin, convergence to within 0.1 kcal/mol typically occurs in five or fewer steps for individuals in the initial population and only one or two steps for subsequent individuals.

The critical point method also uses readily available local minimization software, but it does require both the first and second derivatives. It also may find either a local minimum, a local maximum, or a saddle point. It is an interesting question whether the possibility of the latter events is disadvantageous. Their occurrence indicates that the GA individual being evaluated is in some sense closer to a maximum (or saddle point) than to a local minimum. This information might be useful in directing the GA search. We have not investigated this method experimentally.

Software implementing the exact second derivative method is not as readily available, and the method requires both the first and second derivatives. It is guaranteed to find a local minimum. We also have not experimented with this method.

Combinations and variations of these methods are possible, such as beginning with conjugate gradient and then transitioning to quasi-Newton. Also, the termination criteria may depend on either the number of iterations or a convergence tolerance, which may be measured in dihedral angle space, energy, or both. In the remainder of the paper, we consider only the first derivative method based on conjugate gradient, terminating after a single step. This method is less computationally expensive than one in which a full minimization is performed for each individual, but retains the benefits of the hybrid approach.

We use a readily available implementation of conjugate gradient [19], except that we modify the bracketing procedure used in the line minimizations. The standard bracketing procedure (`mymnbrak.c`) assumes that the domain of each of the independent variables is the set of all real numbers, whereas our independent variables assume values only in the interval $[-\pi, \pi]$. Consequently, the intervals produced by the standard procedure typically are not limited to the basin of attraction in which the encoded conformation lies. Our method heuristically corrects this problem by choosing an interval over which no dihedral angle varies by more than $\frac{\pi}{6}$. Neglecting non-bonded interactions, this guarantees that the bracketed interval is contained in the conformation's basin of attraction, and that it contains the local minimum along the direction of minimization.

2.4 Analytical Gradient

Because we vary only the dihedral angles, we have

$$\nabla E_{\mathcal{B}} = \nabla E_{\mathcal{A}} = 0. \tag{3}$$

Thus, the components of the gradient are given by the sums of the partial derivatives of $E_{\mathcal{D}}$, $E_{\mathcal{N}}$, and $E_{\mathcal{N}'}$ with respect to the dihedral angles. In our application the constants $\gamma_{ijkl} \in \{0, \pi\}$ for all i, j, k,and l. Thus, the partial derivatives of $E_{\mathcal{D}}$ may be calculated straightforwardly as

$$\frac{\partial E_{\mathcal{D}}}{\partial \Phi_{ijkl}} = \pm n_{ijkl} K_{\Phi_{ijkl}} \sin(n_{ijkl}\Phi_{ijkl}) \tag{4}$$

for $(i,j,k,l) \in \mathcal{D}$ and zero otherwise, where the positive sign is taken for $\gamma_{ijkl} = \pi$ and the negative for $\gamma_{ijkl} = 0$. The partial derivatives of $E_{\mathcal{N}}$ and $E_{\mathcal{N}'}$ are obtained by twice applying the chain rule [7]. For each $(a,b,c,d) \in \mathcal{D}$, we have

$$\frac{\partial E_{\mathcal{N}}}{\partial \Phi_{abcd}} = \sum_{(i,j)\in\mathcal{N}} \left[-12 A_{ij}^{12} r_{ij}^{-13} + 6 B_{ij}^{6} r_{ij}^{-7} - \frac{q_i q_j}{4\pi\varepsilon r_{ij}^2} \right] \cdot \frac{\partial r_{ij}}{\partial \Phi_{abcd}} \tag{5}$$

and

$$\begin{aligned} \frac{\partial r_{ij}}{\partial \Phi_{abcd}} &= \frac{\partial r_{ij}}{\partial x_i}\frac{\partial x_i}{\partial \Phi_{abcd}} + \frac{\partial r_{ij}}{\partial y_i}\frac{\partial y_i}{\partial \Phi_{abcd}} + \\ &\quad \frac{\partial r_{ij}}{\partial z_i}\frac{\partial z_i}{\partial \Phi_{abcd}} + \frac{\partial r_{ij}}{\partial x_j}\frac{\partial x_j}{\partial \Phi_{abcd}} + \\ &\quad \frac{\partial r_{ij}}{\partial y_j}\frac{\partial y_j}{\partial \Phi_{abcd}} + \frac{\partial r_{ij}}{\partial z_j}\frac{\partial z_j}{\partial \Phi_{abcd}} \end{aligned} \tag{6}$$

where (x_i, y_i, z_i) and (x_j, y_j, z_j) are the Cartesian coordinates of atoms i and j respectively. For each $(i,j) \in \mathcal{N}$ we have

$$r_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2. \tag{7}$$

Thus,

$$\begin{aligned} \frac{\partial r_{ij}}{\partial x_i} &= \frac{x_i - x_j}{r_{ij}}, \quad \text{and} \\ \frac{\partial r_{ij}}{\partial x_j} &= -\frac{x_i - x_j}{r_{ij}} \end{aligned} \tag{8}$$

and similarly for the partial derivatives with respect to y_i, y_j, z_i, and z_j. Combining Equations 5, 6, 7, and 8 yields the required derivatives of $E_{\mathcal{N}}$ as functions of the atoms' Cartesian coordinates and the partial derivatives thereof with respect to the dihedral angles. We use Thompson's method [24] to calculate the partial derivatives of the Cartesian coordinates with respect to the dihedral angles. The derivatives of $E_{\mathcal{N}'}$ are obtained similarly except for the leading factor of $\frac{1}{2}$ and the set over which the summation is taken.

2.5 Hybrid Genetic Algorithm

In the context of constrained optimization problems, Orvosh and Davis [18] propose replacing infeasible individuals by their repaired counterparts with probability $p_r = 0.05$. Pseudocode for this algorithm is shown in Figure 1. This algorithm may be viewed as probabilistically Lamarckian. Alternatively, one may view the local minimization operator as a repair operator in the sense that it maps individuals to the "feasible region," where the nonlinear equality constraint to be satisfied is $\nabla E = 0$.

```
initialize();
for (gen=0 ; gen < max_gen; gen++){
   for (i=0 ; i < pop_size ; i++) {
      temp = pop[i];
      local_min(temp);
      pop[i].fitness = temp.fitness;
      if (Rand() < p_r)
         pop[i] = temp;
   }
   select();
   recombine();
   mutate();
}
```

Figure 1: Probabilistically Lamarckian genetic algorithm pseudocode

3 RESULTS

In this section we present the results of experiments in which we empirically compare the minimum energies found by the standard genetic algorithm (denoted SGA), the SGA followed by one step of conjugate gradient minimization (denoted SGA+1CG), and probabilistically Lamarckian genetic algorithms using various replacement probabilities $p_r \in \{0, 0.05, 0.10, 1.00\}$ (denoted Baldwinian, $p_r = 0.05$, $p_r = 0.10$, and Lamarckian, respectively). The experiments are performed using a modification of the 1990 version of GENESIS running on SPARC workstations. The input parameters are as given in the typical input file shown in Figure 2. The minimum energies obtained in each generation, averaged over 5 runs per algorithm, are shown in Figure 3, except those for SGA+1CG. The results for the latter are identical to those for the SGA except in the final generation. The probabilistically Lamarckian genetic algorithms quickly find substantially lower energy conformations than the SGA, most notably the purely Lamarckian algorithm. The means and standard deviations of the final generation minimum energies are

```
      Experiments = 1
     Total Trials = 50000
  Population Size = 50
Structure Length = 240
   Crossover Rate = 0.65
    Mutation Rate = 0.003
   Generation Gap = 1.0
   Scaling Window = 1
  Report Interval = 1
Structures Saved = 1
Max Gens w/o Eval = 10
    Dump Interval = 0
      Dumps Saved = 0
          Options = ce
      Random Seed = 987654321
```

Figure 2: Typical GENESIS run time parameter file

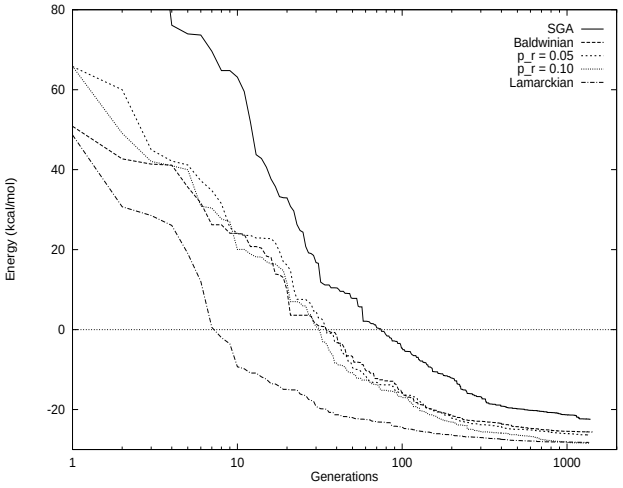

Figure 3: Average minimum energy vs. generation using fitness proportionate selection

shown in Table 1.

The SGA+1CG algorithm results in slightly lower final energies than the SGA. The probabilistically Lamarckian algorithms obtain final energies which are significantly lower than those of the SGA+1CG at the 0.005 level of significance, as determined by the Kruskal-Wallis H Test [1]. The final energies obtained by the various probablistically Lamarckian algorithms in these experiments are not different from each other at any interesting level of statistical significance.

Examination of the distributions of energies (not shown) indicates that in each generation, most individuals have energies close to the best individual, but that there are a few individuals with much larger energies. This causes most individuals to have very similar fitnesses, thereby reducing selective pressure. Thus, we

Table 1: Final generation minimum energies (kcal/mol) using fitness proportionate selection

Algorithm	Mean	Std. Dev.
SGA	-22.46	1.10
SGA+1CG	-22.89	1.62
Baldwinian	-25.61	2.22
$p_r = 0.05$	-26.41	1.52
$p_r = 0.10$	-28.37	2.00
Lamarckian	-28.23	1.66

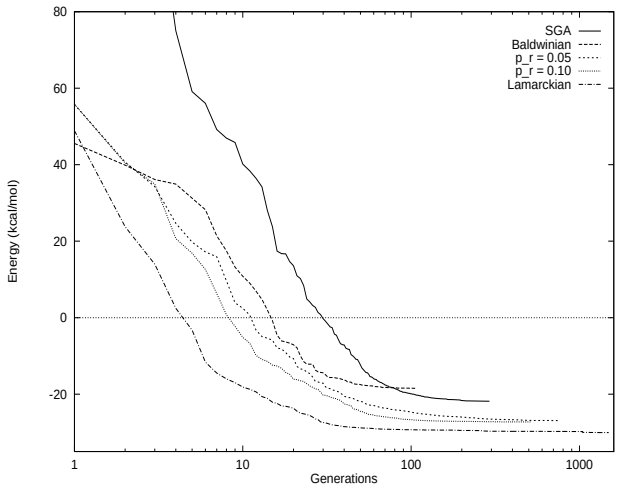

Figure 4: Average minimum energy vs. generation using tournament selection

also perform the experiments using binary tournament selection variants of each of the algorithms (c.f. [15]). Our implementation of tournament selection creates the new population in two identical steps, each of which consists of randomly pairing the individuals in the old population and including the more fit of each pair in the new population. Thus, each individual in the old population participates in exactly two tournaments, and the presence of high energy individuals in the population does not prevent discrimination between low energy individuals. The minimum energies obtained via tournament selection are shown in Figure 4 and in Table 2.

Again, the SGA+1CG algorithm results in slightly

Table 2: Final generation minimum energies (kcal/mol) using tournament selection

Algorithm	Mean	Std. Dev.
SGA	-21.85	2.50
SGA+1CG	-21.91	2.48
Baldwinian	-18.47	2.43
$p_r = 0.05$	-26.90	2.08
$p_r = 0.10$	-27.26	1.07
Lamarckian	-30.05	2.75

Table 3: Number of trials performed prior to obtaining 99% convergence using tournament selection

Algorithm	Mean	Std. Dev.
SGA	5545	2486
Baldwinian	2562	573
$p_r = 0.05$	15031	5815
$p_r = 0.10$	12964	2472

lower final energies than the SGA. In contrast to the results for fitness proportionate selection, the final energies obtained by the probabilistically Lamarckian strategies are as a group not lower than those obtained by the SGA at any interesting level of statistical significance.[2] This is primarily due to premature convergence, which is most notable in the Baldwinian algorithm and absent in the purely Lamarkian algorithm. The latter obtained between 90.3% and 96.0% convergence at 50000 trials. The mean and standard deviation of the number of trials performed by the remaining algorithms prior to obtaining 99% convergence is shown in Table 3.

The increased selective pressure of tournament selection causes the Baldwinian algorithm to abandon higher energy basins of attraction before it produces individuals representing the associated local minima, resulting in a loss of apparently critical information. This occurs to a lesser degree in the SGA and the probabilistically Lamarckian algorithms for $p_r \in \{0.05, 0.10\}$. The Baldwinian algorithm is the only one for which the final energies obtained using fitness proportionate selection and tournament selection are significantly different, the latter being higher at the 0.01 level of significance.

4 CONCLUSIONS

While Lamarckian genetic algorithms obtain good solutions for some applications (e.g. [11]), it has also been shown that Baldwinian algorithms are superior for other applications [28], while probabilistically Lamarckian approaches are superior for others [18]. All of the probabilistically Lamarckian algorithms used in this investigation obtained better energies than the SGA for the minimization of the CHARMM potential for [Met]-Enkephalin, with the exception of the Baldwinian algorithm using tournament selection.

The effectiveness of the probabilistically Lamarckian algorithms suggests that the low-energy local minima in the energy landscape of [Met]-Enkephalin occur somewhat regularly within the conformation space. If this is the case for [Met]-Enkephalin, it seems likely that it will hold for larger polypeptides as well, and hence that probabilistically Lamarckian algorithms are appropriate techniques for protein structure prediction. Replacement frequencies must be appropriate to the level of selective pressure, in order to ensure the presence of enough locally optimal individuals to prevent premature convergence.

[2] The final energies obtained using $p_r \in \{0.05, 0.10, 1.00\}$ are lower than those obtained using SGA+1CG at the 0.005 level of significance.

5 FUTURE DIRECTIONS

Comparison of the effectiveness of the algorithms used in this investigation to that of other algorithms requires that they be applied to the same energy model. Previous work involving [Met]-Enkephalin has been based on the ECEPP/2 model [13, 17]. The minimum energy structure with respect to that model has an energy of -28.96 kcal/mol when evaluated in CHARMM, which is higher than that of four of the structures obtained in this investigation. We are currently comparing our algorithm to the Monte Carlo with Minimization algorithm [17], using both the CHARMM and ECEPP/2 energy models.

Application of these algorithms to significantly larger molecules requires computational resources which are only available through the use of highly scalable architectures. We have previously used island model genetic algorithms successfully for protein structure prediction [6], and we are now studying parallel designs of the hybrid algorithms presented here.

References

[1] Allen, Arnold O. *Probability, Statistics, and Queueing Theory: With Computer Science Applications.* Computer Science and Scientific Computing, San Diego, California: Academic Press, Inc., 1990.

[2] Brooks, Bernard R. and others. "CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamic Calculations," *Jounal of Computational Chemistry*, *4*(2):187–217 (1983).

[3] Chan, Hue Sun and Ken A. Dill. "The Protein Folding Problem," *Physics Today*, 24–32 (February 1993).

[4] Dandekar, Thomas and Patrick Argos. "Potential of Genetic Algorithms in Protein Folding and Protein Engineering Simulations," *Protein Engineering*, *5*(7):637–645 (1992).

[5] Forrest, Stephanie, editor. *Proceedings of the Fifth International Conference on Genetic Algorithms*, San Mateo CA: Morgan Kaufmann Publishers, Inc., July 1993.

[6] Gates, Jr., George H., et al. "Parallel Simple and Fast Messy GAs for Protein Structure Prediction." *Proceedings of the Intel Supercomputer Users' Group 1995 Annual North America Users Conference*. Beaverton, Oregon: Intel Supercomputer Systems Division, 1995.

[7] Gaulke, Robert L. *The Application of Hybridized Parallel Genetic Algorithms to the Protein Folding Problem*. MS thesis, Air Force Institute of Technology, 1995. In Preparation.

[8] Goldberg, David E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading MA: Addison–Wesley Publishing Company, 1989.

[9] Grefenstette, John J., editor. *Genetic Algorithms and Their Applications: Proceedings of the Second International Conference on Genetic Algorithms*, Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers, July 1987.

[10] Holland, John H. *Adaptation in Natural and Artificial Systems* (First mit press Edition). Cambridge, MA: MIT Press, 1992.

[11] Judson, R. S., et al. "Do Intelligent Configuration Search Techniques Outperform Random Search for Large Molecules?," *International Journal of Quantum Chemistry*, *44*:277–290 (1992).

[12] Judson, Richard S. "Teaching Polymers to fold," *The Journal of Physical Chemistry*, *96*(25):10102 (1992).

[13] LeGrand, Scott M. and Kenneth M. Merz Jr. "The Application of the Genetic Algorithm to the Minimization of Potential Energy Functions," *Journal of Global Optimization*, *3*:49–66 (1991).

[14] Männer, Reinhard and Bernard Manderick, editors. *Parallel Problem Solving from Nature, 2*, Amsterdam: North-Holland, September 1992.

[15] McGarrah, D.B. and R.S. Judson. "Analysis of the Genetic Algorithm Method of Molecular Conformation Determination," *Journal of Computational Chemistry*, *14*(11):1385–1395 (1993).

[16] Michalewicz, Z. *Genetic Algorithms + Data Structures = Evolution Programs* (Second Edition). Springer, 1994.

[17] Nayeem, Akbar, et al. "A Comparative Study of the Simulated-Annealing and Monte Carlo-with-Minimization Approaches to the Minimum-Energy Structures of Polypeptides: [Met]-Enkephalin," *Journal of Computational Chemistry*, *12*(5):594–605 (1991).

[18] Orvosh, David and Lawrence Davis. "Shall We Repair? Genetic Algorithms, Combinatorial Optimization, and Feasibility Constraints." In Forrest [5], 650.

[19] Press, William H., et al. *Numerical Recipes in C: The Art of Scientific Computing* (Second Edition). New York: Cambridge University Press, 1992.

[20] Schulze-Kremer, Steffen. "Genetic Algorithms for Protein Tertiary Structure Prediction." In Stender [21], 129–149.

[21] Stender, Joachim, editor. *Parallel Genetic Algorithms: Theory and Applications*. Amsterdam: IOS Press, 1993.

[22] Sun, S., et al. *Biophysics Journal*, *62*:104 (1992).

[23] Sun, Shaojian. "Reduced representation model of protein structure prediction: Statistical potential and genetic algorithms," *Protein Science*, *2*:762–785 (1993).

[24] Thompson, H. Bradford. "Calculation of Cartesian Coordinates and Their Derivatives from Internal Molecular Coordinates," *The Journal of Chemical Physics*, *47*(9):3407–3410 (November 1967).

[25] Tuffery, P., et al. "A New Approach to the Rapid Determination of Protein Side Chain Conformations," *Journal of Biomolecular Structure & Dynamics*, *8*(6):1267 (1991).

[26] Unger, Ron and John Moult. "Genetic Algorithms for Protein Folding Simulations," *Journal of Molecular Biology*, *231*:75–81 (1993).

[27] Vásquez, Maximiliano, et al. "Conformational Energy Calculations on Polypeptides and Proteins," *Chemical Reviews*, *94*:2183–2239 (1994).

[28] Whitley, Darrell, et al. "Lamarckian Evolution, The Baldwin Effect and Function Optimization." In Männer and Manderick [14], 6–15.