

UMA REDE IMUNOLÓGICA ARTIFICIAL PARA A PREDIÇÃO DA ESTRUTURA DE PROTEÍNAS

CAROLINA P. ALMEIDA^{†,*} RICHARD A. GONÇALVES^{†,*} MYRIAM R. DELGADO[†]

* UNICENTRO

Guarapuava, Paraná, Brasil

[†] Universidade Tecnológica Federal do Paraná

Curitiba, Paraná, Brasil

Email: carolpa@cpgei.cefetpr.br richardgoncalves@cpgei.cefetpr.br
myriamdelg@utfpr.edu.br

Abstract— This paper presents an Artificial Immune System based on Immune Network's theory to solve the Protein Structure Prediction Problem (PSPP). The PSPP may be described as the definition of the protein native conformation given its primary structure, where the native conformation is the minimum energy state achieved by the folded protein. The proposed approach uses an immune network associated with a local search method based on tabu search. The system is tested in a benchmark database with different protein sequences. The results show that the proposed approach achieves similar or even superior performance when compared with other approaches described in the literature.

Keywords— Artificial Immune Systems, Immune Network, Tabu Search, Protein Structure Prediction.

Resumo— Este artigo apresenta um Sistema Imunológico Artificial baseado na teoria da rede imune para resolver o problema de Predição da Estrutura de Proteínas (PPEP). O PPEP pode ser definido como a determinação da conformação nativa de uma proteína dada sua estrutura primária, sendo a conformação nativa o estado de energia mínima alcançado pela proteína dobrada. A abordagem proposta utiliza uma rede imune associada a um método de busca local baseado em busca tabu. O sistema é testado em uma base de dados conhecida contendo diferentes seqüências de proteínas. Os resultados mostram que o sistema obtém desempenho semelhante ou até mesmo superior ao de outras abordagens existentes na literatura.

Keywords— Sistema Imunológico Artificial, Rede Imune, Busca Tabu, Predição da Estrutura de Proteínas.

1 Introdução

As proteínas são cadeias polipeptídicas de resíduos de aminoácidos. A seqüência linear de aminoácidos que formam uma proteína é chamada de estrutura primária da proteína. Quando deixada em condições ambientais apropriadas, essa seqüência dobra-se alcançando um estado único de energia mínima. A estrutura tri-dimensional de uma proteína (também chamada estrutura terciária) é determinada por este estado, o qual é denominado conformação nativa da proteína (Alberts et al., 2002). O problema de Predição da Estrutura de Proteínas pode ser definido como a determinação da conformação nativa de uma proteína dada sua estrutura primária. Neste trabalho o problema será baseado num modelo reduzido denominado modelo Hidrofóbico Polar Tridimensional (3D HP).

O uso de métodos heurísticos para resolver o Problema da Predição da Estrutura de Proteínas no modelo HP é bastante justificável, pois não existem métodos determinísticos capazes de resolver esse problema. Neste sentido, os Sistemas Imunológicos Artificiais (SIA) aparecem como uma alternativa promissora.

Os sistemas imunológicos artificiais, como outras técnicas inspiradas na natureza, tentam extrair dos sistemas biológicos idéias para desenvolver ferramentas que resolvam problemas com-

putacionais (de Castro and Timmis, 2002). A teoria da rede imune define o sistema imunológico como uma rede regulada de células e moléculas que se reconhecem mesmo na ausência de agentes externos (antígenos).

Outra metaheurística interessante é a busca tabu que visa guiar o processo de busca para a geração de soluções ótimas ou muito próximas do ótimo, através do histórico deste processo. Esta técnica é em geral aplicada no contexto de otimização combinatória e foi usada com sucesso em diversos problemas como telecomunicações, roteamento, escalonamento, entre outros (Glover, 1995). Neste sentido, o uso da busca tabu permite incrementar o desempenho do modelo proposto uma vez que o PPEP em modelos reduzidos, como o 3D HP, pode ser visto como um problema de otimização combinatorial.

A abordagem proposta neste artigo utiliza um SIA, mais especificamente uma rede imunológica artificial, associada a um método de busca local baseado em busca tabu.

Este artigo é estruturado da forma descrita a seguir. Após esta introdução, a Seção 2 discute sobre o problema da predição de estrutura de proteínas, detalhando o modelo 3D HP adotado. Os conceitos básicos de SIA são apresentados na Seção 3, com atenção especial para as redes imunológicas artificiais. A Seção 4 traz a descrição da abordagem que está sendo proposta. Na seção

5 são discutidos os experimentos realizados, assim como os resultados obtidos através da comparação do modelo proposto com outros existente na literatura. Finalmente, a Seção 6 apresenta as conclusões e perspectivas de trabalhos futuros.

2 Predição de Estrutura de Proteínas

Descobrir a estrutura tri-dimensional de uma proteína é fundamental para encontrar sua função biológica em um organismo. Esse conhecimento é essencial para o desenvolvimento de novos medicamentos para alguns tipos de doenças, para o tratamento ou prevenção de doenças causadas pelo mau dobramento de proteínas - Fibrose Cística, Mal de Alzheimer e o mau da "vaca louca", por exemplo, (Cohen and Kelly, 2003) - e para o desenvolvimento de polímeros biológicos com propriedades materiais específicas (Shmygelska and Hoos, 2005).

A estrutura de proteínas pode ser determinada com exatidão por técnicas como Ressonância Nuclear Magnética (NMRI) e Cristalografia, porém essas técnicas são muito custosas em termos de equipamentos, tempo e esforço computacional. Além disso, elas exigem isolamento, purificação e cristalização da proteína alvo (Alberts et al., 2002), dificultando ainda mais o processo. Por isso, métodos computacionais para a solução deste problema são muito atrativos.

A Predição da Estrutura de Proteínas é um processo complexo que envolve conceitos biológicos, químicos e físicos. Então, os métodos computacionais desenvolvidos para solucionar o Problema de Predição da Estrutura de Proteínas (PPEP) são geralmente baseados em modelos reduzidos. Apesar desses modelos reduzidos abstraírem apenas as características mais relevantes de todo o processo, o PPEP ainda é uma tarefa desafiadora. Como discutido por Berger and Leighton (1998), a Predição da Estrutura de Proteínas associada ao modelo 3D HP é comprovadamente um problema NP-difícil.

Vários métodos computacionais têm sido aplicados na solução do PPEP no modelo 3D HP, tais como Otimização por Colônias de Formigas (OCF) (Shmygelska and Hoos, 2005) (Chu et al., 2005), Algoritmos Evolucionários (AEs) (Cotta, 2003), Algoritmos de Monte-Carlo (AMC) (Liang and Wong, 2001) (Hsu et al., 2003) e Sistemas Imunológicos Artificiais (SIAs) (Cutello et al., 2005) (Cutello et al., 2007) (Almeida et al., 2007).

2.1 Modelo Hidrofóbico-Polar

Dentre esses modelos, o modelo Hidrofóbico-Polar (HP) (Lau and Dill, 1989) é o mais estudado e aplicado. O HP é um modelo de energia

livre baseado em dois importantes fatos (Lau and Dill, 1989) (Richards, 1977):

- A interação hidrofóbica é a força guia do dobramento de proteínas e a hidrofobicidade dos aminoácidos é a principal força para o desenvolvimento da conformação nativa de pequenas proteínas globulares;
- As estruturas nativas de muitas proteínas são compactas e têm núcleos compactos com grande concentração de resíduos hidrofóbicos, assim como uma mínima área de superfície não-polar exposta no solvente.

No modelo HP os vinte (20) aminoácidos que compõem as proteínas são divididos em duas categorias: resíduos Hidrofóbicos/Não-polares (H) e resíduos Hidrofílicos/Polares (P). Então, a estrutura primária de uma proteína pode ser representada como uma string sobre o alfabeto $\{H,P\}^+$.

Como dois resíduos não podem ocupar o mesmo lugar no espaço, conformações factíveis de uma seqüência no modelo HP são restritas a um caminho sem colisões em um lattice. Para o modelo 2D HP, um lattice quadrado bi-dimensional é tipicamente utilizado, enquanto o modelo 3D HP é baseado em um lattice cúbico tri-dimensional. Um exemplo de representação no modelo 3D HP de uma proteína com 50 resíduos é mostrado na Figura 1.

Uma conformação factível no modelo HP é associada a um nível de energia livre. Nesse modelo, a energia livre de uma conformação é proporcional ao número de contatos topológicos entre resíduos hidrofóbicos que não são vizinhos na seqüência dada. Mais especificamente, a energia livre de uma certa conformação com η contatos hidrofóbicos não-locais é $-\eta$.

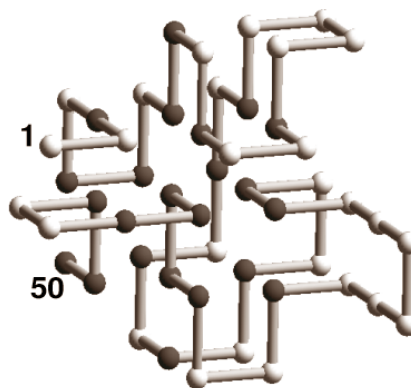


Figura 1: Representação de uma Proteína no modelo 3D HP. As esferas claras representam resíduos polares e as esferas escuras representam resíduos hidrofóbicos.

O problema de Predição da Estrutura de Proteínas no modelo HP pode ser formalmente

definido como: dada uma seqüência $s = s_1 s_2 \dots s_n$ no alfabeto $\{H,P\}$ encontre uma conformação de energia mínima de s , isto é, encontre $c^* \in C(s)$ tal que $E(c^*) = \min\{E(c) | c \in C\}$, onde C é o conjunto de todas as conformações factíveis para s (Shmygelska and Hoos, 2005).

3 Sistema Imunológico Artificial

O sistema imunológico (SI) natural é um mecanismo complexo responsável pela defesa do organismo contra o ataque de microorganismos externos, os antígenos, que podem causar doenças. Quando ocorre o ataque de um antígeno a um organismo, as células do SI respondem a esse ataque através de um conjunto (população) de anticorpos. O SI é altamente distribuído, adaptativo, auto-organizável por natureza, mantém uma memória de encontros (antígenos/anticorpos) passados e tem a habilidade de aprender continuamente a partir de novos encontros. Os Sistemas Imunológicos Artificiais herdaram essas características tornando-se uma poderosa ferramenta computacional.

Sistemas Imunológicos Artificiais (SIAs) podem ser definidos como sistemas computacionais inspirados pela teoria imunológica e funções, princípios e modelos imunes, os quais são aplicados na solução de problemas (de Castro and Von Zuben, 2000). As metáforas biológicas mais usadas nos SIAs são: o princípio da seleção clonal, a seleção negativa, a seleção positiva e a teoria das redes imunes (de Castro and Von Zuben, 2000) (Kephart et al., 1998).

Como o SIA proposto nesse trabalho é baseado na teoria da rede imune, ela será melhor explicada na subseção seguinte.

3.1 Teoria da Rede Imune

A teoria da Rede Imunológica, ou Rede Idiotípica, foi proposta por Jerne (1974) para explicar as capacidades de memória e aprendizado dos SI. A principal hipótese desta teoria diz que a memória imune é mantida mesmo sem a presença de antígenos devido às interações entre as células do sistema imunológico.

Um antígeno possui apenas regiões que podem ser reconhecidas por um anticorpo, essas regiões são chamadas de epitopo. Já um anticorpo possui uma região de reconhecimento de antígenos (paratopo) e também uma região que pode ser reconhecida por outros anticorpos (idiotopo). Assim, a produção de um anticorpo Ab_1 estimula (ou suprime) a produção de outro anticorpo Ab_2 que estimula (ou suprime) a produção de outro anticorpo Ab_3 e assim por diante (Parisi, 1990). Se um anticorpo reconhece uma célula sua produção é estimulada pela rede. Caso um anticorpo seja reconhecido por outro anticorpo sua produção é

inibida (supressão). Desta forma, a rede possui um mecanismo auto-regulatório de produção de anticorpos quer um antígeno esteja presente no sistema ou não.

As redes imunológicas possuem três características principais (Bersini and Varela, 1990):

- **Estrutura:** descrição dos componentes celulares e moleculares da rede e suas interconexões, desconsiderando as interações entre estes componentes;
- **Dinâmica:** interação constante entre os diversos componentes do sistema;
- **Meta-dinâmica:** o sistema imunológico pode interagir com qualquer elemento, para isso a rede precisa estar sempre produzindo novos anticorpos. Como o repertório imunológico tem tamanho constante, para que novos anticorpos sejam gerados outros precisam morrer o que faz com que o sistema imunológico esteja em constante renovação. Essa renovação garante diversidade para que o SI possa reagir a novos ataques.

Na seção seguinte é descrito como o algoritmo proposto implementa essas características.

4 Modelo de Rede Imune Proposto

Na rede imune proposta um antígeno é a representação primária da proteína para a qual se deseja encontrar a estrutura tri-dimensional e os anticorpos representam conformações factíveis para tal proteína. A forma geral do algoritmo proposto pode ser resumida pelo Pseudocódigo 1.

```

Rede Imune(TamPROT, TamPOP, FatorDUP,
IdadeMAX, TaxaHIP, LimiarEST, LimiarSUP,
TamMIN, PercSEL)
  gerações ← 0;
  POP = Inicialização();
  Avalia(POP);
  Enquanto (!Condição de Parada()) faça
    Enquanto !(Estabilidade da Rede())
      POPc ← Clonagem(POP, dup);
      POPh ← Hipermuta(POPc, TaxaHIP);
      POPm ← Hipermacro(POPc);
      Avalia(POPh);
      Avalia(POPm);
      POP ← Seleção(POP, POPh, POPm);
      gerações ← gerações+1;
    Fim Enquanto
  POP ← Supressão(POP, LimiarSUP);
  Se (TamPOPatual < TamMIN) então
    Gerar novos anticorpos para completar POP;
  Fim Se
  Se !(K-M melhoram em K-I) então
    POP ← MaturaçãoFracca(POP);
  Fim Se
Fim Enquanto
  POP ← MaturaçãoIntesiva(POP);

```

Pseudocódigo 1: Algoritmo da Rede Imune

A população inicial (Geração 0) é gerada aleatoriamente de tal forma que cada anticorpo

na população de soluções (POP) representa uma solução factível para o PPEP (isto é, representa um caminho sem colisões da seqüência da proteína em um lattice). Esse caminho geralmente é feito utilizando-se coordenadas internas. As coordenadas internas dependem da topologia particular do lattice considerado. Neste trabalho foi considerada a representação em um lattice cúbico onde cada localização tem no máximo seis vizinhos (Cotta, 2003). Além disso, foi adotado o esquema absoluto para representar os movimentos internos, pois trabalhos anteriores compararam ambas as representações (absoluta e relativa), sendo que a representação absoluta foi a que obteve os melhores resultados (Cotta, 2003) (Cutello et al., 2007).

Depois de criado, cada anticorpo deve ser avaliado. A função de *fitness* mede o número de ligações hidrofóbicas não-locais (LHNL) da proteína (ver Seção 2.1). Então, o problema de encontrar a energia mínima de uma conformação é transformado em um problema equivalente de maximização do número de LHNL. Isto é feito pela função *Avalia()*, a qual recebe a população de anticorpos como parâmetro e retorna o *fitness* de cada anticorpo. *Condição_Parada()* é a função que retorna verdadeiro sempre que o processo evolucionário deve parar e falso caso contrário. Neste trabalho o critério de parada é definido como um número máximo de avaliações da função de *fitness*.

A rede imunológica entra então na etapa de maturação de afinidade, a qual é a etapa onde ocorre a maior parte da dinâmica da rede. Nesta etapa são aplicados os operadores de clonagem, hipermutação, hipermacromutação, avaliação e seleção, conforme detalhado a seguir. Cabe salientar que todos os operadores aplicados geram apenas anticorpos factíveis, ou seja, sem colisões. Esta etapa é repetida até que a rede alcance sua estabilidade, a qual é medida pela função *Estabilidade da Rede()*. A estabilidade da rede é medida pela diferença entre a afinidade média da população atual e anterior, quando este valor é inferior ao limiar de estabilidade ($Limiar_{EST}$) a rede é considerada estabilizada.

O operador de clonagem (*Clonagem()*) produz algumas cópias (clones) de cada anticorpo. Esse operador gera uma população intermediária de clones (POP_c) com tamanho igual a $Tam_{POP} * dup$, onde Tam_{POP} é o número de anticorpos da população inicial e dup é o parâmetro que define o número de cópias de cada anticorpo.

Durante o processo evolucionário dois tipos de operadores de mutação são aplicados: hipermutação inversamente proporcional ao *fitness* e Hipermacromutação. A função de Hipermutação (*Hipermuta()*) recebe dois parâmetros - a população de clones e a taxa de hipermutação ($Taxa_{HIP}$) - e retorna uma população intermediária (POP_h). A função de Hipermacromutação (*Hipermacro()*) recebe apenas a população de clones e retorna uma

população intermediária (POP_m). No operador de Hipermutação, M_{max} (o número máximo de mutações permitido em um anticorpo) é inversamente proporcional ao valor da afinidade de cada anticorpo e é determinado pela Eq. 1.

$$M_{max}(A(x)) = \begin{cases} (1 + \frac{E^*}{A(x)}) * \alpha, & \text{if } A(x) > 0 \\ (1 + E^*) * \alpha + \alpha, & \text{if } A(x) = 0 \end{cases} \quad (1)$$

onde $A(x)$ é o valor do *fitness* do anticorpo x , E^* é o melhor valor de *fitness* conhecido e $\alpha = Taxa_{HIP} * Tam_{PROT}$. O tamanho da proteína é dado por Tam_{PROT} .

A Hipermacromutação aplicada é a mesma utilizada em (Cutello et al., 2007) em um algoritmo baseado na seleção clonal. Ela tenta mutar cada anticorpo, sempre gerando indivíduos factíveis. O número máximo de mutações, que é independente da afinidade do anticorpo sendo hipermacromutado, pode ser definido como $M_{max} = j - i + 1$, onde i e j são dois inteiros aleatórios tal que $i < j \leq Tam_{PROT}$. O operador de Hipermacromutação seleciona aleatoriamente a posição inicial (i) e final (j) da perturbação, a direção da perturbação também é aleatória tanto da posição i para a posição j (esquerda para a direita) quanto da posição j para a posição i (esquerda para a direita).

Após a aplicação destes operadores os anticorpos gerados são submetidos a avaliação, que é a mesma que foi aplicada à população inicial. Então ocorre a seleção dos anticorpos que farão parte da próxima população. Esta seleção é feita dentre os anticorpos da população atual (POP), os anticorpos hipermutados (POP_{hiper}) e os anticorpos hipermacromutados (POP_{macro}). Os Tam_{POP} melhores anticorpos são selecionados pela função de seleção (*Seleção()*). Um anticorpo a é considerado melhor que um anticorpo b se ambos são factíveis (não possuem colisões) e a afinidade de a é maior que a de b , ou a é factível e b não ou se ambos são infactíveis mas a tem menos colisões que b .

Após a rede atingir sua estabilidade é aplicado o operador de supressão, que é um mecanismo de interação anticorpo-anticorpo. Este mecanismo visa eliminar da população anticorpos redundantes e é responsável por implementar a meta-dinâmica da rede. Para medir a semelhança entre dois anticorpos foi utilizada a distância de Hamming normalizada. Se dois anticorpos possuem uma semelhança superior a $Limiar_{SUP}$, o pior deles deve ser eliminado da população. Se após a supressão o tamanho mínimo da população não for atingido (Tam_{MIN}), novos anticorpos são gerados de forma aleatória para completar a população. A supressão faz com que a diversidade dos anticorpos seja sempre boa.

A cada K_I iterações é testado se houve melhora na afinidade dos K_M melhores anticorpos.

Se não ocorreu nenhuma melhora, os K_M melhores anticorpos são submetidos a um processo de maturação fraca. A maturação fraca realiza uma busca local baseada nos operadores de hipermutação e hipermacromutação. A principal diferença entre este processo e os operadores nos quais ele se baseia é que a maturação fraca interrompe as mutações assim que um anticorpo possui uma afinidade superior ao anticorpo original.

Inspirados pelos resultados obtidos em (Blazewicz et al., 2005) onde uma busca tabu foi aplicada na solução do Problema da Predição da Estrutura de Proteínas no modelo 3D HP, optou-se por introduzir uma etapa de maturação intensiva da afinidade após o critério de parada ser atendido. Nesta etapa os anticorpos são submetidos a uma busca tabu (Glover, 1995), com o intuito de melhorar os resultados obtidos durante os passos anteriores.

5 Experimentos e Resultados

Nesta seção, um *benchmark* popular para a Predição da Estrutura de Proteínas no modelo 3D HP é utilizado para testar o método baseado em SIA proposto neste trabalho. As descrições das instâncias de seqüências de proteínas consideradas podem ser encontradas em (Cutello et al., 2005).

Os experimentos foram executados de maneira independente 50 vezes. A condição de parada foi a mesma adotada em (Cotta, 2003) Backtracking-EA e (Cutello et al., 2007), ou seja, o máximo de 100.000 execuções da função de avaliação. Os parâmetros da rede imune foram determinados empiricamente através de simulações com diversos valores para cada parâmetro, utilizando a seqüência de tamanho 48. Os melhores resultados foram obtidos com os seguintes parâmetros: $Tam_{POP} = 50$, $dup = 4$, $Taxa_{HIP} = 0.6$, $Limiar_{EST} = 0.002$, $K_M = 30$, $K_I = 50$, $Limiar_{SUP} = 0.9$. Os parâmetros da busca tabu foram baseados nos resultados obtidos em (Blazewicz et al., 2005) e foram: tamanho da lista tabu igual a 20, número máximo de iterações igual a 200, vizinhança de uma mutação e o critério de aspiração melhoria do melhor anticorpo encontrado até o momento. Portanto, todas as simulações foram realizadas com esse conjunto de parâmetros.

A Tabela 1 apresenta a comparação entre os resultados obtidos pela Rede Imune proposta, o Backtracking-EA (Cotta, 2003) e o Aging-SIA (Cutello et al., 2007). Nesta tabela pode-se observar que o algoritmo proposto é capaz de encontrar os melhores valores de energia (e, por conseguinte, as melhores estruturas) para quatro das cinco instâncias (ele apenas não é capaz de encontrar a energia mínima para a terceira instância). Com relação ao valor médio de energia obtido durante as 50 execuções dos algoritmos é possível ver que a

Rede Imune encontra médias melhores para quatro das cinco instâncias, sendo superada apenas pelo Aging-SIA na segunda instância.

Considerando-se os resultados das duas últimas instâncias fica evidente a superioridade da Rede Imune em encontrar melhores resultados para instâncias maiores: nestas instâncias as médias dos valores de energia obtidas pela Rede Imune são superiores aos melhores valores encontrados pelos demais algoritmos.

Testes estatísticos seriam necessários para atestar com precisão se as diferenças encontradas nos resultados são realmente significativas, mas como não foram publicados detalhes sobre a distribuição dos resultados obtidos pelo Backtracking-EA e pelo Aging-SIA tais testes não puderam ser realizados.

6 Conclusões

Neste artigo foi proposto um Sistema Imunológico Artificial baseado na teoria das Redes Imunes hibridizado com Busca Tabu para a resolução do Problema da Predição da Estrutura de Proteínas no modelo Hidrofóbico Polar Tri-Dimensional. A metodologia proposta incorpora vários operadores imunes (clonagem, hipermutação, hipermacromutação, supressão, maturação fraca da afinidade, maturação intensiva da afinidade e estabilidade da rede) se enquadrando adequadamente nas características propostas por (Bersini and Varela, 1990).

Os resultados obtidos pela Rede Imune proposta foram comparados com algoritmos do estado-da-arte (Backtracking-EA (Cotta, 2003) e Aging-SIA (Cutello et al., 2007)). A Rede Imune foi capaz de obter resultados comparáveis ou, na maioria dos casos, superiores aos obtidos pelos demais algoritmos. Os resultados são particularmente interessantes nas duas últimas instâncias, onde a média dos resultados obtidos pela rede foi superior ao melhor valor obtido pelos algoritmos do estado-da-arte. Estes resultados estabelecem a Rede Imune proposta como um método bastante eficiente para a solução do PPEP no modelo 3D HP.

Como trabalhos futuros pretende-se experimentar outras técnicas de busca local na etapa de maturação intensiva da afinidade e testar a Rede Imune em outras representações de proteínas, principalmente em representações livre de lattice.

Referências

- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002). *Molecular Biology of the Cell*, 4 edn, Garland Publishing, New York, USA.

Tabela 1: Comparação entre Rede Imune, Backtracking-EA e Aging-SIA.

N.	E^*	Rede Imune			Backtracking-EA			Aging-SIA		
		Melhor	Média	σ	Melhor	Média	σ	Melhor	Média	σ
1	-9	-9	-9	0	-9	-7.98	0.71	-9	-9	0
2	-18	-18	-16.02	1.13	-18	-14.38	1.26	-18	-16.76	1.02
3	-29	-28	-25.94	1.07	-25	-20.80	1.61	-29	-25.16	0.45
4	-26	-26	-23.75	1.68	-23	-20.20	1.50	-23	-22.60	0.40
5	-49	-49	-43.98	2.22	-39	-34.18	2.31	-41	-39.28	0.24

- Almeida, C. P., Gonçalves, R. A. and Delgado, M. R. (2007). A hybrid immune-based system for the protein folding problem, in C. Cotta and J. V. Hemer (eds), *Seventh European Conference on Evolutionary Computation in Combinatorial Optimisation - EvoCOP 2007*, Lecture Notes in Computer Science, EvoCOP, Springer, Valencia, Spain.
- Berger, B. and Leighton, T. (1998). Protein folding in the hidrophobic-hidrophilic model is np complete, *Journal of Computational Biology* **5**: 27–40.
- Bersini, H. and Varela, F. J. (1990). Hint for adaptive problem solving gleaned from immune networks, *Parallel Problem Solving from Nature*, PPSN, Dortmund, FRG, pp. 343–354.
- Blazewicz, J., Lukasiak, P. and Milostan, M. (2005). Application of tabu search strategy for finding low energy structure of protein, *Artificial Intelligence in Medicine* **35**: 135–145.
- Chu, D., Till, M. and Zomaya, A. Y. (2005). Parallel ant colony optimization for 3d protein structure prediction using the hp lattice model, *19th International Parallel and Distributed Processing Symposium*, IPDPS, Nice, France.
- Cohen, F. E. and Kelly, F. E. (2003). Therapeutic approaches to protein-misfolding diseases, *Nature* **426**: 905–909.
- Cotta, C. (2003). Protein structure prediction using evolutionary algorithms hybridized with backtracking, *7th International Work-Conference on Artificial and Natural Neural Networks* **2687**: 321–328.
- Cutello, V., Morelli, G., Nicosia, G. and Pavone, M. (2005). Immune algorithms with aging operators for the string folding problem and the protein folding problem, *Evolutionary Computation in Combinatorial Optimization*, EvoCOP, Lausanne, Switzerland, pp. 80–90.
- Cutello, V., Nicosia, G., Pavone, M. and Timmis, J. (2007). An immune algorithm for protein structure prediction on lattice models, *IEEE Transactions on Evolutionary Computation* **11**(1): 101–117.
- de Castro, L. N. and Von Zuben, F. J. (2000). The clonal selection algorithm with engineering applications, *Workshop on Artificial Immune Systems and Their Applications*, GECCO, Las Vegas, Nevada, USA, pp. 36–37.
- de Castro, L. and Timmis, J. (2002). *Artificial Immune Systems: A New Computational Approach*.
- Glover, F. (1995). Tabu search fundamentals and uses, University of Colorado at Boulder.
- Hsu, H. P., Mehra, V., Nadler, W. and Grassberger, P. (2003). Growth algorithm for lattice heteropolymers at low temperatures, *Journal of Chemical Physics* **118**: 444–451.
- Jerne, N. K. (1974). Towards a network theory of the immune system, *Immunology* **125C**: 373–389.
- Kephart, J., Sorkin, B., Swimmer, M. and White, S. (1998). *Artificial Immune Systems and Their Applications*, Springer-Verlag, Berlin, German, chapter Blueprint for a Computer Immune System, pp. 242–260.
- Lau, K. F. and Dill, K. A. (1989). Lattice statistical mechanics model of conformation and sequence space of proteins, *Macromolecules* **22**: 3986–3997.
- Liang, F. and Wong, W. H. (2001). Evolutionary monte carlo for protein folding simulation, *Journal of Chemical Physics* **115**: 3374–3380.
- Parisi, G. (1990). A simple model for the immune network, *Immunology* **87**: 429–433.
- Richards, F. M. (1977). Areas, volumes, packing and protein structures, *Annual Review of Biophysics and Bioengineering* **6**: 151–176.
- Shmygelska, A. and Hoos, H. H. (2005). An ant colony optimisation algorithm for the 2d and 3d hidrofobic polar protein folding problem, *BMC Bioinformatics* **6**: 1–22.