Garza-Fabre M, Rodriguez-Tello E, Toscano-Pulido G. Comparative analysis of different evaluation functions for protein structure prediction under the HP mode. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 28(5): 868–889 Sept. 2013. DOI 10.1007/s11390-013-1384-7

# Comparative Analysis of Different Evaluation Functions for Protein Structure Prediction Under the HP Model

Mario Garza-Fabre, Eduardo Rodriguez-Tello, and Gregorio Toscano-Pulido

*Information Technology Laboratory, CINVESTAV-Tamaulipas, Km. 5.5 Carretera Ciudad Victoria-Soto La Marina 87130, Ciudad Victoria, Tamaulipas, México*

E-mail: {mgarza, ertello, gtoscano}@tamps.cinvestav.mx

Received January 11, 2013; revised March 27, 2013.

**Abstract** The HP model for protein structure prediction abstracts the fact that hydrophobicity is a dominant force in the protein folding process. This challenging combinatorial optimization problem has been widely addressed through metaheuristics. The evaluation function is a key component for the success of metaheuristics; the poor discrimination of the conventional evaluation function of the HP model has motivated the proposal of alternative formulations for this component. This comparative analysis inquires into the effectiveness of seven different evaluation functions for the HP model. The degree of discrimination provided by each of the studied functions, their capability to preserve a rank ordering among potential solutions which is consistent with the original objective of the HP model, as well as their effect on the performance of local search methods are analyzed. The obtained results indicate that studying alternative evaluation schemes for the HP model represents a highly valuable direction which merits more attention.

**Keywords** evaluation function, protein structure prediction, metaheuristics, combinatorial optimization, bioinformatics

## 1 Introduction

Proteins play a very important role in performing most essential biological and chemical functions in a cell associated with life. They are necessary for carrying out structural, enzymatic, transport, and regulatory functions. It is widely accepted that protein functions are strictly determined by their three-dimensional (3D) conformation. Therefore, to fully understand the biological roles of a protein, it is imperative to first determine its structure. However, given the limitations of the experimental methods, computational approaches to determining the structure of proteins have become increasingly necessary for the understanding of such important biological macromolecules.

The protein structure prediction (PSP) problem is concerned with finding the native conformation of proteins. Such a structure is assumed to be encoded in the amino acid sequence forming the protein and corresponds to the thermodynamically most stable state[1]. Nevertheless, exploring the huge conformational space to find the native structure of a protein represents a very computationally-intensive task, which makes studies at atomic resolution prohibitive even for relatively small proteins. Thus, simplified protein models have been proposed in the literature as valuable tools for studying the most general and essential principles governing the protein folding process[2-6].

One of these simplified formulations of the PSP is the Hydrophobic-Polar (HP) model[7-8]. This model captures the fact that hydrophobicity is one of the main driving forces determining the functional conformation of proteins. Despite its apparent simplicity, the prediction of protein structures based on the HP model represents a hard combinatorial optimization problem. This problem has been proved to be NP-complete[9-10], which justifies the diversity of metaheuristic approaches that have been adopted to address it (see Subsection 2.1).

The success (or failure) of these metaheuristic algorithms depends heavily on a set of key components that must be carefully designed. The evaluation function is a prominent example of these components. It is responsible for assessing the quality of a candidate solution with respect to the optimization objective in order to orient the search towards promising regions of the solutions space. A good evaluation function is expected to be able to distinguish each potential solu-

Regular Paper

This research was partially supported by the National Council of Science and Technology of México (CONACyT) under Grant Nos. 105060 and 99276.

tion from the others, and thus to effectively guide the search method to make the most appropriate choice at each of its iterations. On the contrary, an evaluation function providing a poor discrimination may produce large plateaus in the landscape[11-13], on which local search strategies could fail to detect a promising search direction[14]. Hence, the evaluation issue is expected to seriously compromise the efficiency of metaheuristic algorithms.

The conventional evaluation function of the HP model features a very poor discrimination ability. As a consequence, no preferences can be set among potential conformations, leading the search process to be driven almost at random. For this reason, there exist alternative evaluation functions for the HP model that have been proposed to improve the performance of search algorithms[14-19]. In most of the cases, however, the proposal of these alternative evaluation approaches was not supported, or it was only partially supported, by solid experimental evidence.

This paper extends a preliminary work reported in [20], which was the first intent to formally analyze and compare different alternative evaluation schemes for the HP model. It assessed the discrimination potential of four alternative evaluation functions for the HP model with respect to the conventional one. Comparisons were carried out employing the studied evaluation functions within a basic memetic algorithm over a reduced subset of HP benchmark sequences for the two-dimensional (2D) square lattice. This preliminary study presented some weaknesses, including: the lack of understanding concerning the impact of the analyzed functions on the metaheuristics efficiency when solving sequences for the 3D cubic lattice, the absence of a statistical validation of the experimental results, and the question of whether the conclusions drawn with the basic memetic algorithm could be generalized to other metaheuristics.

The present work extends significantly the study reported in [20]. The main extensions can be summarized as follows: 1) A total of seven different formulations of the evaluation function for the HP model are considered. 2) An in-depth investigation of the discrimination potential for each of the studied functions is performed. 3) A new property to evaluate the alternative evaluation functions capacity to preserve the conventional rank ordering among potential protein conformations is introduced. This property is called *HP-compatibility*, and measures the consistency of an alternative evaluation scheme with respect to the original objective of the HP model. An extensive analysis regarding the HP-compatibility of the studied evaluation functions is carried out. 4) An assessment of the practical usefulness of these evaluation approaches within two different metaheuristic algorithms, best improvement local search and iterated local search, is presented. 5) All the experiments consider a full test-suite composed of 30 well-known benchmark sequences for the HP model (including 2D and 3D lattices). And 6) a rigorous statistical significance analysis of the experimental results is conducted.

The remainder of this article is organized into five other sections. Section 2 formally introduces the protein structure prediction problem and the HP model, analyzes some characteristics of the conventional evaluation function and highlights its potential drawbacks. The six considered alternative evaluation functions for the HP model are described in Section 3. Section 4 details the adopted test cases and the performance assessment methodology. Section 5 is devoted to presenting our experimental results related with a careful examination of two important properties of the studied evaluation functions, the degree of discrimination and the HP-compatibility. The effectiveness of these approaches to guiding the search process is also evaluated within two different metaheuristics. Finally, Section 6 provides our conclusions as well as some possible directions for future research.

## 2 Protein Structure Prediction

Anfinsen's theory of protein folding states that the 3D structure of a protein is determined by the physicochemical properties of its amino acid sequence, and that such a native conformation corresponds to the one that minimizes the overall free energy, i.e., the thermodynamically most stable state of the molecule. This is the so-called *thermodynamic hypothesis*[1]. Anfinsen's theory laid the foundation of one of the most active and challenging areas in bioinformatics: protein structure prediction.

The protein structure prediction (PSP) problem can be defined as the problem of finding the functional conformation for a protein given as the only input data its amino acid sequence. In PSP, one considers a fixed energy model $E : \mathcal{C} \rightarrow \mathbb{R}$, where $\mathcal{C}$ is the set of all possible conformations of the protein, and the native conformation is assumed to be the one with the lowest energy value according to the adopted energy model. That is, the conformation $c^* \in \mathcal{C}$ such that $E(c^*) = \min\{E(c) \mid c \in \mathcal{C}\}$[①].

Thus, we could simply enumerate and evaluate all possible conformations to identify the one with minimal energy. Nevertheless, proteins are very flexible and, consequently, the space of potential conformations

① Hereafter the terms energy function and evaluation function are used indistinctly.

is huge. This makes studies at atomic resolution prohibitive to some extent even for relatively small proteins. In this context, simplified models have emerged as important tools for theoretical studies of protein structure, dynamics and thermodynamics. These models provide a valuable insight to advance the understanding of the most general and essential principles governing the protein folding process[2-6]. This study focuses on one of such simplified protein models: the so-called HP model[7-8], which is described next.

### 2.1 Hydrophobic-Polar Model

Amino acids can be classified on the basis of their affinity for water. Hydrophilic or polar amino acids ($P$) are usually found at the outer surface of proteins. By interacting with the aqueous environment, these amino acids contribute to the solubility of the molecule. In contrast, hydrophobic or nonpolar amino acids ($H$) tend to pack on the inside of proteins, where they interact with one another to form a water-insoluble core. This phenomenon is usually referred to as hydrophobic collapse. The hydrophobicity of the amino acids represents, therefore, one of the major driving forces responsible for the final 3D conformation of proteins.

Following these observations Dill[7] proposed the Hydrophobic-Polar (HP) model, where proteins are abstracted as chains of $H$- and $P$-type beads. Protein sequences, which are originally defined over a 20-letter alphabet, are thus of the form $S = (s_1, s_2, \ldots, s_L)$, where $s_i \in \{H, P\}$ denotes the $i$-th amino acid and $L$ the length of the sequence. The number of $H$ and $P$ amino acids in $S$ are here referred to as $L_H$ and $L_P$, respectively. A feasible protein conformation is modeled as a self-avoiding walk on a given lattice, that is, as an embedding of the protein chain on the lattice such that the following two properties are satisfied: 1) self-avoidance, two different amino acids cannot be mapped to the same lattice position; and 2) connectivity, consecutive amino acids in $S$ are to be also adjacent in the lattice. In this paper, we focus our attention on both, the 2D square lattice and the 3D cubic lattice[8].

With the aim of emulating the so-called hydrophobic collapse, in the HP model the goal is to maximize the interaction among $H$ amino acids in the lattice. Such interactions are to be referred to as topological contacts. Two $H$ amino acids $s_i$ and $s_j$ are said to form a topological contact if they are nonconsecutive in $S$ (i.e., $|j - i| \geqslant 2$) but adjacent in the lattice. The objective is thus to find a feasible protein conformation where the number of $H$-$H$ topological contacts (*HHtc*) is maximized. Adhering to the notation of the field, an energy function, to be minimized, is defined as the negative of *HHtc*; maximizing *HHtc* is equivalent to minimize such an energy function.

Formally, PSP under the HP model is defined as the problem of finding $c^* \in \mathcal{C}_\mathcal{F}$ such that $E_{D85}(c^*) = \min\{E_{D85}(c) \mid c \in \mathcal{C}_\mathcal{F}\}$, where $\mathcal{C}_\mathcal{F}$ is the set of all feasible protein conformations and $\mathcal{C}_\mathcal{F} \subsetneq \mathcal{C}$. The energy function is denoted by $E_{D85} : \mathcal{C} \to \mathbb{R}$ and maps protein conformations to energy values. $E_{D85}(c)$, the energy of a conformation $c \in \mathcal{C}$, is defined as follows②:

$$E_{D85}(c) = \sum_{s_i, s_j} e(s_i, s_j), \tag{1}$$

where

$$e(s_i, s_j) = \begin{cases} -1, & \text{if } s_i \text{ and } s_j \text{ are both } H \text{ and} \\ & \text{form a topological contact,} \\ 0, & \text{otherwise.} \end{cases}$$

As an example, the optimal conformation for a protein sequence of length $L = 20$ on the 2D square lattice is presented in Fig.1. This example corresponds to sequence 2d4, one of the HP benchmark sequences adopted for this study, see Section 4.

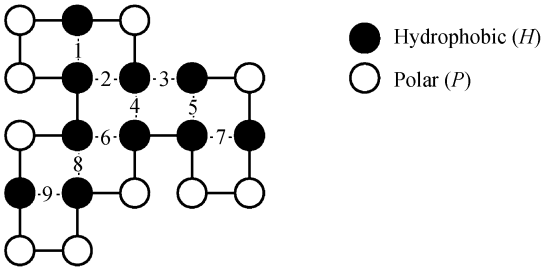


Fig.1. Optimal conformation for sequence 2d4 of length $L = 20$ on the 2D square lattice. Black and white balls denote $H$ and $P$ residues, respectively. $H$-$H$ topological contacts have been numbered. The energy of this conformation is $E_{D85}(c) = -9$, since $HHtc = 9$.

In spite of its apparent conceptual simplicity, the task of finding the optimal structure of a protein in the HP model represents a hard combinatorial optimization problem which has been proved to be NP-complete[9-10]. Such a complexity has motivated the use of a variety of metaheuristics to address this problem, including genetic algorithms[14-16,21-24], memetic and hybrid algorithms[17,19,25-29], tabu search[30-31], ant colony optimization[32-35], immune-based algorithms[36-39], particle swarm optimization[40-41], differential evolution[42-45] and estimation of distribution algorithms[46-47].

②The acronym D85 is used to distinguish this conventional function from the other evaluation approaches considered in this study.

#### 2.1.1 Protein Structure Representation

In the literature, most of the reported metaheuristic algorithms for the HP model are based on an internal coordinates representation. Using internal coordinates, a protein conformation is encoded as a sequence of moves specifying the lattice position for each amino acid with respect to the preceding one; the position of the first amino acid is fixed. Two alternative encoding schemes can be adopted: the absolute moves encoding[48] and the relative moves encoding[49].

In this study, the absolute moves encoding is implemented. Given a global reference system defined by the lattice, the absolute encoding represents 3D conformations as sequences in $\{F, B, L, R, U, D\}^{L-1}$, to denote the forward, backward, left, right, up and down moves from one amino acid to the next; only moves $\{F, B, L, R\}$ are allowed in the 2D case. An example of the absolute moves encoding is provided in Fig.2.

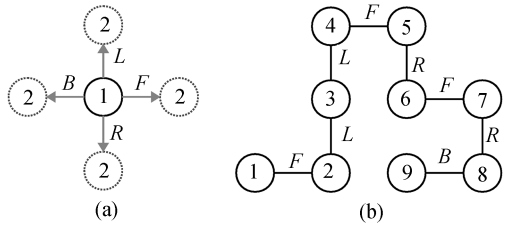


Fig.2. (a) Encoding scheme. (b) Example conformation encoded as FLLFRFRB.

### 2.2 Analyzing the Conventional Evaluation Function

It is well-known that metaheuristics rely on an effective evaluation scheme in order to guide the search process towards promising regions in the solutions space. However, as mentioned before the conventional evaluation function of the HP model, originally defined in (1), induces a very poor discrimination among potential conformations. That is, there could be many different conformations for a given protein sequence with the same energy value, see Fig.3.

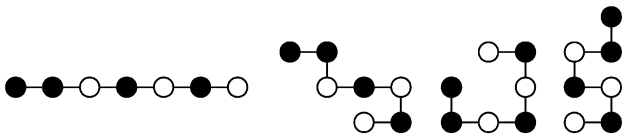

Fig.3. Four different structures for sequence HHPHPHP on the 2D square lattice. All of them with the same energy value, $E_{D85}(c) = 0$.

More precisely, given a protein sequence $S$, with length $L$ and optimal energy value $E^*_{D85}$, there can be at most $|E^*_{D85}| + 1$ available energy levels to classify a search space of size[③] $|\mathcal{C}| = 4^{L-1}$. As an example, consider sequence 2d1, the smallest of the test cases adopted for this study (see Section 4). In this case, $L = 18$ and $E^*_{D85} = -4$, so that there are only five different energy levels which can be used to discriminate among a total of $4^{17} = 17\,179\,869\,184$ potential conformations. Nevertheless, some equally ranked conformations could present better chances than others to be further improved.

The low discrimination provided by the conventional energy function of the HP model is translated into the existence of large plateaus in the search landscape. In such plateaus, metaheuristics (mainly trajectory/local search-based methods) could fail to detect a promising direction, leading the search process to be oriented almost at random.

In the literature, different alternative energy functions for the HP model have been proposed[14-19]. The aim of these alternative formulations of the energy function is to provide a more fine-grained discrimination, as a means of guiding metaheuristics in a more effective manner during the process of finding potential solutions to the original problem. In Section 3, the main details of these alternative energy functions are analyzed.

## 3 Alternative Energy Functions for the HP Model

This section describes several alternative formulations of the HP model's energy (evaluation) function which have been proposed in the literature. A three-letter acronym has been assigned to each of the studied evaluation functions. The acronyms adopted are as follows: K99[14], C04[15], L06[16,50], B08[17], C08[18,51] and I09[19,26,52]. Below, each one of these alternative energy functions is defined.

### 3.1 Alternative Energy Function K99

In the conventional energy function of the HP model, only $H$-$H$ topological contacts contribute to the quality assessment of conformations. Given two conformations with the same number of $H$-$H$ topological contacts, it is possible, however, that one of them has better characteristics (more compact) than the other.

Based on this observation, Krasnogor *et al.*[14] proposed the following distance-dependent energy function:

$$E_{K99}(c) = \sum_{s_i, s_j} e(s_i, s_j), \tag{2}$$

where $e(s_i, s_j) = -1$ if $s_i$ and $s_j$ are both $H$

③The given size of the search space assumes the use of the absolute moves representation of the protein conformations on the 2D square lattice, see Subsection 2.1.1.

and they form a topological contact; $e(s_i, s_j) = -1/(d(s_i, s_j)^k L_H)$ if $s_i$ and $s_j$ are both $H$ but the lattice distance between them is $d(s_i, s_j) > 1$; and $e(s_i, s_j) = 0$, otherwise. Krasnogor *et al.*[14] suggested to use the values $k = 4$ for the square lattice and $k = 5$ for the cubic and triangular lattices, respectively.

According to Krasnogor *et al.*[14], this alternative formulation of the evaluation function preserves the conventional rank ordering of the conformations, at the same time it enables a finer level of distinction among conformations with the same number of $H$-$H$ topological contacts. The behavior of this evaluation function was investigated using a genetic algorithm over only five relatively short protein sequences (less than 50 amino acids). Experiments were performed for the 2D square and triangular lattices, as well as for the 3D cubic lattice. No detailed results are provided; the authors pointed out that no significant improvements in performance were obtained by using this modified energy function. However, they suggested that the advantages of using this function can become more evident for larger protein sequences and when this approach is implemented within local search strategies. The relevance of using this proposal needs to be further investigated.

### 3.2 Alternative Energy Function C04

Given that the aim of the HP model is only to maximize interactions between $H$ amino acids, the positioning of $P$ amino acids is not directly optimized. This may result in unnatural structures for sequences with long $P$ segments and, particularly, when such $P$ segments are located at the ends of the chain[15]. An example of this scenario is presented in Fig.4.

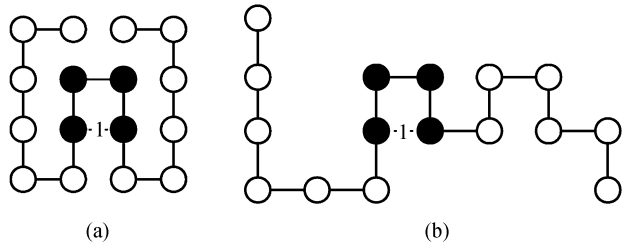


Fig.4. Two conformations with the same number of $H$-$H$ topological contacts ($HHtc = 1$). However, the structure of (a) is more natural-like (globular) than the one of (b).

Custódio *et al.*[15] proposed a modified energy function based on the assumption that it may be preferable for an $H$ amino acid to have a $P$ neighbor rather than to be in contact with the aqueous solvent. In the proposed function, the energy of a conformation is computed as the weighted sum of the number of $H$-$H$ contacts ($HHc$), $H$-$P$ contacts ($HPc$) and $H$-Solvent contacts ($HSc$). A free lattice location (not assigned to any amino acid) is said to be occupied by the solvent. Formally, the energy of a conformation $c$ is given by:

$$E_{C04}(c) = \omega_1 HHc + \omega_2 HPc + \omega_3 HSc, \tag{3}$$

where $\omega_1$, $\omega_2$ and $\omega_3$ denote the relative importance of $HHc$, $HPc$ and $HSc$, respectively. Although not specified by the authors, these weighting coefficients were set to $\omega_1 = 0$, $\omega_2 = 10$ and $\omega_3 = 40$ for the reported experiments④. Thus, given these weights, the minimization of (3) penalizes $H$-$P$ and $H$-solvent contacts, $H$-$P$ contacts being favored over $H$-solvent contacts, while $H$-$H$ interactions are not penalized ($H$-$H$ contacts have no contribution to the energy value using these weights).

Custódio *et al.*[15] evaluated the suitability of this proposal by using a genetic algorithm. A total of 10 instances for the 3D cubic lattice were considered. Seven of the sequences have 27 amino acids and the remaining 3 sequences are of length $L = 64$. The proposed function allows to improve the performance of the implemented algorithm for some of the adopted test cases. The reported results also suggest that this function presents a greater tendency to form more natural-like conformations.

### 3.3 Alternative Energy Function L06

In [16, 50], an alternative energy function for the HP model which is based on the concept of radius of gyration was proposed. The radius of gyration is a measure of the compactness of conformations; the more compact the conformation, the smaller the value for this measure. The proposed function is defined in (4):

$$E_{L06}(c) = HnLB \times RadH \times RadP. \tag{4}$$

The $HnLB$ term comprises the number of $H$-$H$ topological contacts in the conformation ($HHtc$) and a penalty factor which takes into account the violation of the self-avoiding constraint. Formally:

$$HnLB = HHtc - (NC \times PW), \tag{5}$$

where $NC$ is the number of collisions (i.e., lattice nodes assigned to more than one amino acid) in the conformation and $PW$ is the penalty weight. The value of $PW$ depends on the chain length, $L$, and it can be computed as $PW = (0.033 \times L) + 1.33$[50]⑤.

Before defining the $RadH$ and $RadP$ terms, let us define $R_H$ as the radius of gyration for $H$ amino acids:

④This information was obtained through personal communication with the authors.

⑤In this study, only feasible protein structures are considered; the penalty factor in (5) is simply omitted.

$$R_H = \sqrt{\frac{\sum_{s|s=H} (x_s - X_H)^2 + (y_s - Y_H)^2}{L_H}}, \quad (6)$$

where $x_s$ and $y_s$ are the lattice coordinates of amino acid $s$. $X_H$ and $Y_H$ denote the arithmetic mean of the coordinates for all $H$ amino acids. Analogously, we can compute $R_P$, the radius of gyration for $P$ amino acids, by considering $P$ rather than $H$ amino acids in (6)[⑥].

Once $R_H$ has been defined, the *RadH* term measures how compact the hydrophobic core of the conformation is. *RadH* is given by:

$$RadH = MaxR_H - R_H, \quad (7)$$

where $MaxR_H$ denotes the radius of gyration for $H$ amino acids in a totally unfolded conformation, i.e., the maximum possible $R_H$ value.

Finally, the *RadP* term aims to push $P$ amino acids away from the hydrophobic core. Given the previously defined $R_H$ and $R_P$ measures, the *RadP* term is computed as:

$$RadP = \begin{cases} 1, & \text{if } (R_P - R_H) \geqslant 0 \\ \dfrac{1}{1 - (R_P - R_H)}, & \text{otherwise.} \end{cases} \quad (8)$$

The *RadP* term will always lie in the range $[0, 1]$. A value of $(R_P - R_H) > 0$ means that $P$ amino acids are more exposed than $H$ amino acids. This is a convenient scenario, so the *RadP* term has no contribution to the final energy value ($RadP = 1$). Otherwise, $(R_P - R_H) < 0$ suggests that $H$ amino acids are more spread than the $P$ ones, so *RadP* is used to penalize the energy value of the conformation. Note that (4) is to be maximized[⑦].

Lopes and Scapin[16,50] argued that the above described function provides an adequate discrimination among conformations with the same number of *H*-*H* topological contacts. This function was implemented within a genetic algorithm in order to solve several instances on the 2D square lattice. However, no results are provided on the advantages of using this alternative function with regard to the conventional energy formulation of the HP model.

### 3.4 Alternative Energy Function B08

Berenboym and Avigal[17] proposed an alternative energy function called the *global energy*. In this function, each pair of nonconsecutive $H$ amino acids contributes to the energy value even if they are not topological neighbors. The global energy for a given conformation $c$ is defined as:

$$E_{B08}(c) = \sum_{s_i, s_j} e(s_i, s_j), \quad (9)$$

where $e(s_i, s_j) = \frac{-1}{(x_{s_i} - x_{s_j})^2 + (y_{s_i} - y_{s_j})^2}$ if $s_i$ and $s_j$ are both $H$ and they are nonconsecutive in $S$ ($|j - i| \geqslant 2$); otherwise, $e(s_i, s_j) = 0$[⑧].

In [17], the effects of using a local search operator within a genetic algorithm were analyzed for both the conventional and the proposed energy functions. However, an explicit comparison to demonstrate the advantages of using a particular energy function was not reported. This issue needs to be further explored.

### 3.5 Alternative Energy Function C08

In [18, 51], an alternative energy formulation which measures the deviation that each pair of $H$ amino acids presents with respect to the unit distance (i.e., topological contact distance) was introduced.

Let $d(s_i, s_j)^2 = (x_{s_i} - x_{s_j})^2 + (y_{s_i} - y_{s_j})^2 + (z_{s_i} - z_{s_j})^2$ be the lattice distance between amino acids $s_i$ and $s_j$, and let $dv(s_i, s_j) = d(s_i, s_j)^2 - 1$ denote its deviation from the unit distance. The energy value of a conformation $c$ is given by:

$$E_{C08}(c) = \sum_{s_i, s_j | s_i = s_j = H} dv(s_i, s_j)^k, \quad (10)$$

where $k \geqslant 1$ is a parameter of the function, whose larger values give more weight to unit distances. We adopt $k = 2$ for this study, since this value provides the best behavior according to the results reported in [18]. $E_{C08}(c^*) = 0$ would refer to the ideal (potentially unrealistic) scenario where all pairs of $H$ amino acids are at a unit distance in conformation $c^*$. In [18, 51], no experimental results supporting the benefits of using the proposed energy function were reported.

### 3.6 Alternative Energy Function I09

In [19, 26, 52], the authors reported a memetic algorithm with a modified energy function which incorporates two additional measures: $H$-compliance ($H_C$) and $P$-compliance ($P_C$).

$H$-compliance measures the proximity of $H$ amino acids to the center of a hypothetical rectangle (or

⑥Note that (6) has been defined for the 2D lattice, but this equation can be easily generalized to the 3D case.

⑦The negative of (4) can be used as an energy-minimization formulation of the problem which adheres to the notation commonly used in this field.

⑧This definition assumes a 2D lattice, but it can be extended to the 3D case.

cuboid in 3D space) enclosing all $H$ amino acids, which is denoted by the reference point $(x_r, y_r)$. Formally, this measure is given by:

$$H_C = \frac{\sum_{s|s=H} (x_r - x_s)^2 + (y_r - y_s)^2}{L_H}, \quad (11)$$

where $x_s$ and $y_s$ denote the lattice coordinates of the $s$ amino acid.

$P$-compliance computes how close $P$ amino acids are to the boundaries of a hypothetical rectangle enclosing all $P$ amino acids. Such a cuboid is defined by $x_{\min}$, $x_{\max}$, $y_{\min}$ and $y_{\max}$. The $P$-compliance measure is formally given by:

$$P_C = \frac{\sum_{s|s=P} \min \left\{ \begin{array}{l} |x_{\min} - x_s|, |x_{\max} - x_s|, \\ |y_{\min} - y_s|, |y_{\max} - y_s| \end{array} \right\}}{L_P}. \quad (12)$$

Finally, the energy of a given conformation $c$ is defined as:

$$E_{I09}(c) = \alpha E_{D85} + H_C + P_C, \quad (13)$$

where $E_{D85}$ is the conventional energy function of the HP model (as defined in (1), see Subsection 2.1) and $\alpha$ is large enough to ensure this will be the dominant term in (13).

In [19], the authors demonstrated the advantages of using the proposed energy function using an 85-length HP protein sequence on the 2D square lattice. However, the influence of using this function should be further explored for a larger set of test cases.

## 4 Experimental Setup

A total of 30 well-known benchmark sequences for the HP model have been considered for the experimentation of this research project. Out of them, 15 are for the 2D square lattice and the other 15 are for the 3D cubic lattice. Tables 1 and 2 present the full HP sequences, their length ($L$) and the optimal or best-known energy value ($E^*_{D85}$) reported in the literature[25,38,52-55].

Although alternative evaluation functions for the HP model are considered in this study, it is important to remark that the goal of the optimization process remains to maximize the number of $H$-$H$ topological contacts (*HHtc*), which is the singular objective in the HP model (see Subsection 2.1). Therefore, all the obtained experimental results are evaluated in terms of the conventional energy function of the HP model[⑨].

**Table 1.** HP Instances for the 2D Square Lattice

| Sequence | | $L$ | $E^*_{D85}$ |
|---|---|---|---|
| 2d1 | $H_2P_5H_2P_3HP_3HP$ | 18 | −4 |
| 2d2 | $HPHPH_3P_3H_4P_2H_2$ | 18 | −8 |
| 2d3 | $PHP_2HPH_3PH_2PH_5$ | 18 | −9 |
| 2d4 | $HPHP_2H_2PHP_2HPH_2P_2HPH$ | 20 | −9 |
| 2d5 | $H_3P_2HPHPHP_2HPHPHP_2H$ | 20 | −10 |
| 2d6 | $H_2P_2HP_2HP_2HP_2HP_2HP_2HP_2H_2$ | 24 | −9 |
| 2d7 | $P_2HP_2H_2P_4H_2P_4H_2P_4H_2$ | 25 | −8 |
| 2d8 | $P_3H_2P_2H_2P_5H_7P_2H_2P_4H_2P_2HP_2$ | 36 | −14 |
| 2d9 | $P_2HP_2H_2P_2H_2P_5H_{10}P_6H_2P_2H_2P_2H\ P_2H_5$ | 48 | −23 |
| 2d10 | $H_2(PH)_4H_3P(HP_3)_3(P_3H)_3PH_4(PH)_4H$ | 50 | −21 |
| 2d11 | $P_2H_3PH_8P_3H_{10}PHP_3H_{12}P_4H_6PH_2PHP$ | 60 | −36 |
| 2d12 | $H_{12}PHPH(P_2H_2P_2H_2P_2H)_3PHPH_{12}$ | 64 | −42 |
| 2d13 | $H_4P_4H_{12}P_6(H_{12}P_3)_3HP_2H_2P_2H_2P_2HPH$ | 85 | −53 |
| 2d14 | $P_6HPH_2P_5H_3PH_5PH_2P_4H_2P_2H_2PH_5PH_{10}$ $PH_2PH_7P_{11}H_7P_2HPH_3P_6HPH_2$ | 100 | −48 |
| 2d15 | $P_3H_2P_2H_4P_2H_3PH_2PH_2PH_4P_8H_6P_2H_6P_9$ $HPH_2PH_{11}P_2H_3PH_2PHP_2HPH_3P_6H_3$ | 100 | −50 |

**Table 2.** HP Instances for the 3D Cubic Lattice

| Sequence | | $L$ | $E^*_{D85}$ |
|---|---|---|---|
| 3d1 | $HPHP_2H_2PHP_2HPH_2P_2HPH$ | 20 | −11 |
| 3d2 | $H_2P_2HP_2HP_2HP_2HP_2HP_2HP_2H_2$ | 24 | −13 |
| 3d3 | $P_2HP_2H_2P_4H_2P_4H_2P_4H_2$ | 25 | −9 |
| 3d4 | $P_3H_2P_2H_2P_5H_7P_2H_2P_4H_2P_2HP_2$ | 36 | −18 |
| 3d5 | $P_2H_3PH_3P_3HPH_2PH_2P_2HPH_4PHP_2H_5PH$ $PH_2P_2H_2P$ | 46 | −35 |
| 3d6 | $P_2HP_2H_2P_2H_2P_5H_{10}P_6H_2P_2H_2P_2HP_2H_5$ | 48 | −31 |
| 3d7 | $H_2(PH)_4H_3P(HP_3)_3(P_3H)_3PH_4(PH)_4H$ | 50 | −34 |
| 3d8 | $PH(PH_3)_2P(PH_2PH)_2H(HP)_3(H_2P_2H)_2$ $PHP_4(H(P_2H)_2)_2$ | 58 | −44 |
| 3d9 | $P_2H_3PH_8P_3H_{10}PHP_3H_{12}P_4H_6PH_2PHP$ | 60 | −55 |
| 3d10 | $H_{12}PHPH(P_2H_2P_2H_2P_2H)_3PHPH_{12}$ | 64 | −59 |
| 3d11 | $P(HPH_2PH_2PHP_2H_3P_3)_3(HPH)_3P_2H_3P$ | 67 | −56 |
| 3d12 | $P(HPH)_3P_2H_2(P_2H)_6H(P_2H_3)_4P_2(HPH)_3$ $P_2HP(PHP_2H_2P_2HP)_2$ | 88 | −72 |
| 3d13 | $P_2H_2P_5H_2P_2H_2PHP_2HP_7HP_3H_2PH_2P_6H$ $P_2HPHP_2HP_5H_3P_4H_2PH_2P_5H_2P_4H_4PH$ $P_8H_5P_2HP_2$ | 103 | −58 |
| 3d14 | $P_3H_3PHP_4HP_5H_2P_4H_2P_2H_2(P_4H)_2P_2H$ $P_2H_2P_3H_2PHPH_3P_4H_3P_6H_2P_2HP_2HPH$ $P_2HP_7HP_2H_3P_4HP_3H_5P_4H_2(PH)_4$ | 124 | −75 |
| 3d15 | $HP_5HP_4HPH_2PH_2P_4HPH_3P_4HPHPH_4P_{11}$ $HP_2HP_3HPH_2P_3H_2P_2HP_2HPHPHP_8HP_3$ $H_6P_3H_2P_2H_3P_3H_2PH_5P_9HP_4HPHP_4$ | 136 | −83 |

Additionally, the overall average performance ($OAP$) measure is adopted in order to assess the overall behavior of the studied approaches. $OAP$ is defined as the average ratio of the obtained mean values to the optimum ($E^*_{D85}$). Formally:

$$OAP = \frac{100\%}{|\mathcal{T}|}\Big(\sum_{t\in\mathcal{T}} \frac{mean(t)}{E^*_{D85}(t)}\Big), \quad (14)$$

⑨The same criterion used in the literature to evaluate the performance of the algorithms employed for solving the PSP problem under the HP model.

where $mean(t)$ denotes the arithmetic mean of the energy values obtained when solving a particular test instance $t$, computed over multiple executions of the experiment, and $\mathcal{T}$ is the set of all test cases. Thus, $OAP$ expresses the performance of the evaluated approaches in a 0% to 100% scale, being $OAP(t) = 100\%$ the preferred value for this measure. $OAP(t) = 100\%$ suggests the ideal situation where the optimal conformation for each instance is reached during all the performed executions.

Finally, in the experiments presented in this paper, the statistical significance analysis is conducted as follows. First, D'Agostino-Pearson omnibus $K^2$ test is used to evaluate the normality of data distributions. For normally distributed data, either ANOVA or the Welch's $t$ parametric tests are used depending on whether the variances across the samples are homogeneous (homoskedasticity) or not. This is investigated using the Bartlett's test. For non-normal data, the nonparametric Kruskal-Wallis test is adopted. A significance level of $\alpha = 0.05$ has been considered.

The algorithms used for the experiments of this study are coded in C language and compiled with *gcc* using the optimization flag -*O*3. All of them are run sequentially into a CPU Xeon X5650 at 2.66 GHz, 2 GB of RAM with Linux operating system.

## 5 Results

In this section, seven different formulations of the energy function for the HP model are evaluated and compared: the conventional energy function of the HP model, D85[7–8], and the six alternative formulations described in Section 3. Important properties of the studied energy functions are first examined in Subsections 5.1 and 5.2. Then, the effectiveness of these approaches to guide the search process is evaluated in Subsections 5.3 and 5.4. For all the experiments reported in this chapter, protein conformations are encoded using an internal coordinates representation based on absolute moves. Moreover, only solutions encoding feasible protein conformations have been considered (see Subsection 2.1).

### 5.1 Degree of Discrimination

The discrimination potential is an important property of the evaluation scheme which impacts directly on the behavior of metaheuristics. That is, if it is not possible to set preferences among candidate solutions, then the progress in the search could become practically dominated by random decisions.

In this subsection, the degree of discrimination provided by the studied energy functions is investigated. This is done by analyzing the distribution of ranks that these approaches induce on a set of protein conformations. A ranking expresses the relationship among a set of items according to a given property. In the context of this study, protein conformations are to be ranked and the property to set such a relationship corresponds to the energy value. Given a set of protein conformations, the first ranking position is assigned to the conformation with the best energy value, the next ranking position to the one with the second best energy value, and so on. If two or more conformations present the same energy, then they will share the same rank.

The relative entropy ($RE$) measure proposed in [56] is adopted. Given a set of $n$ ranked conformations (there are at most $n$ ranks, and at least 1), the relative entropy of the distribution of ranks $D$ is defined as:

$$RE(D) = \frac{\sum_{r} \frac{D(r)}{n} \log\left(\frac{D(r)}{n}\right)}{\log(1/n)}, \tag{15}$$

where $D(r)$ denotes the number of conformations with rank $r$. $RE(D)$ tends to be 1 as approaching to the ideal situation where each conformation has a different rank (i.e., the maximum discrimination). On the other hand, when all the conformations share the same ranking position (i.e., the poorest discrimination), $RE(D)$ takes a value of zero.

In this experiment, 1 000 different feasible conformations are generated at random. For each of the studied energy functions, these solutions are evaluated and ranked to finally compute the $RE$ measure. A total of 100 repetitions of this experiment are performed for all the adopted test instances. The overall statistics of this experiment are presented in Fig.5. Instance-specific results are provided in Figs. 6 and 7, where bars represent the $RE$ values obtained by the different analyzed functions.

From Fig.5, it can be seen that some of the studied functions discriminate more strongly than others. The obtained results are quite similar for both the 2D and the 3D lattices. In all the test cases, the conventional energy function of the HP model, D85, achieves the lowest $RE$ values. This confirms the poor discrimination capabilities of this function, which has been the main factor motivating the exploration of alternative approaches. Among the alternative functions, C04 presents the worst performance in terms of discrimination. Function L06 reaches high $RE$ values most of the time. However, this function presents a moderate discrimination for the shortest test sequences (see Figs. 6 and 7). Regarding I09, it is possible to note that the $RE$ values obtained by this function are almost always above 0.9, which indicates a strong discrimination. Finally, it is important to remark the high degree of discri-

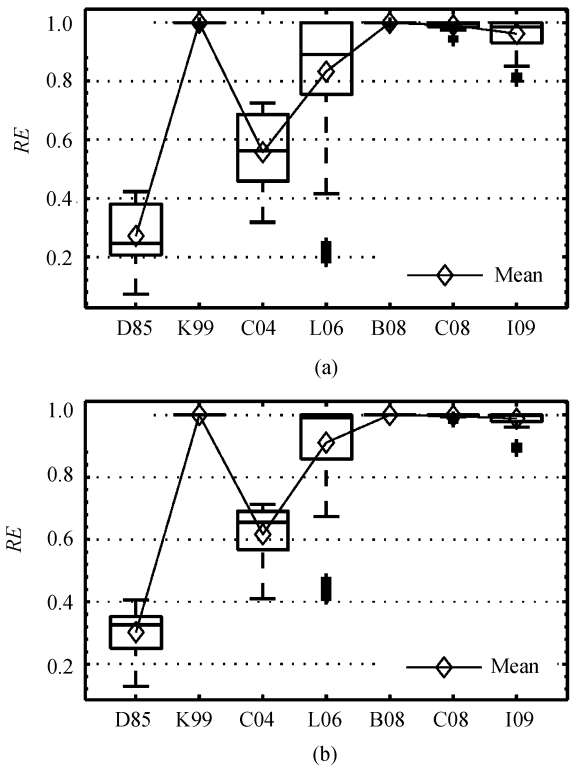


Fig.5. Relative entropy ($RE$) of the distribution of ranks obtained using the different energy functions analyzed. Overall statistics for all (a) 2D and (b) 3D test cases.

mination provided by functions K99, B08 and C08. Functions K99 and B08 are the most discriminative functions according to the obtained results, followed by C08 which suffers slight decreases on some of the instances.

The above results can be better understood by analyzing the histograms with the distribution of ranks achieved by each of the studied energy functions. Fig.8 presents such histograms for a single repetition of this experiment regarding sequence 2d4 on the 2D square lattice (similar results are obtained for other test instances).

From Fig.8, it is possible to note how poor the distribution of ranks achieved by function D85 is. Only seven different ranking positions are induced to classify the 1 000 generated conformations. It can be seen that there are almost 400 conformations sharing the sixth rank. As stated in Subsection 2.2, using function D85 there can be only $|E^*_{D85}| + 1$ different energy levels.

Therefore, no matter the amount of generated conformations, the maximum number of ranks which can be assigned through function D85 is 10, since $E^*_{D85} = -9$ for this benchmark sequence (2d4). The second worst scenario is presented by function C04, where only 40 different ranking positions are produced, out of which one is assigned to more than 100 conformations.

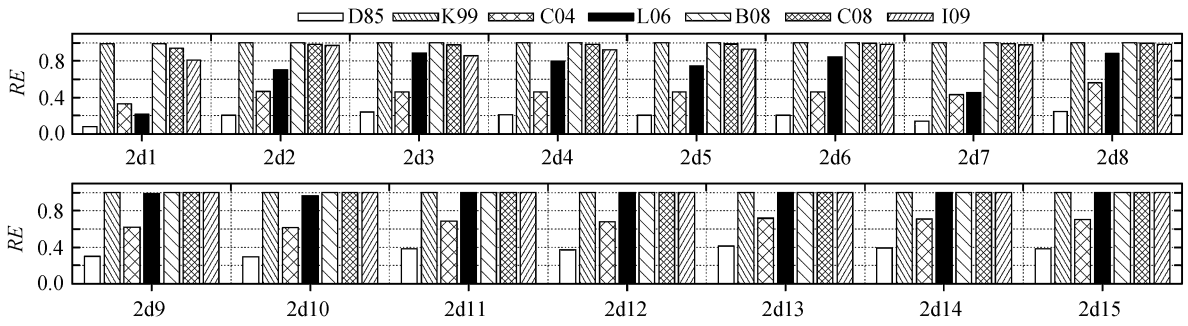


Fig.6. Relative entropy ($RE$) of the distribution of ranks obtained by using the different energy functions analyzed. Average of 100 independent executions. 2D test cases.

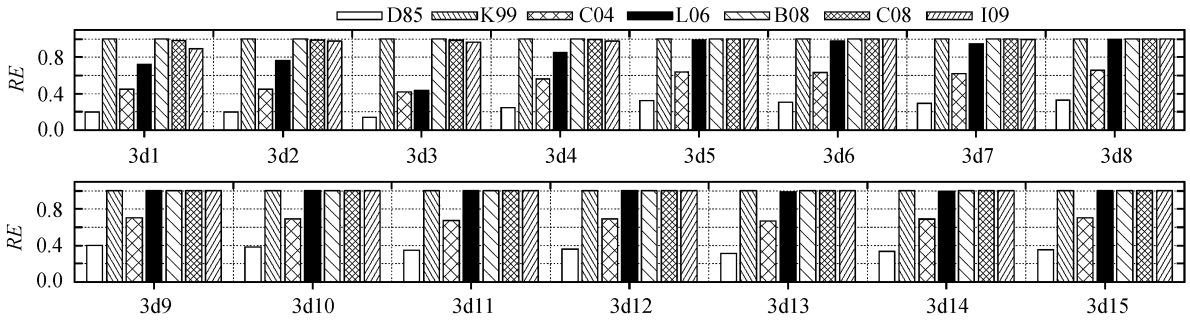


Fig.7. Relative entropy ($RE$) of the distribution of ranks obtained by using the different energy functions analyzed. Average of 100 independent executions. 3D test cases.

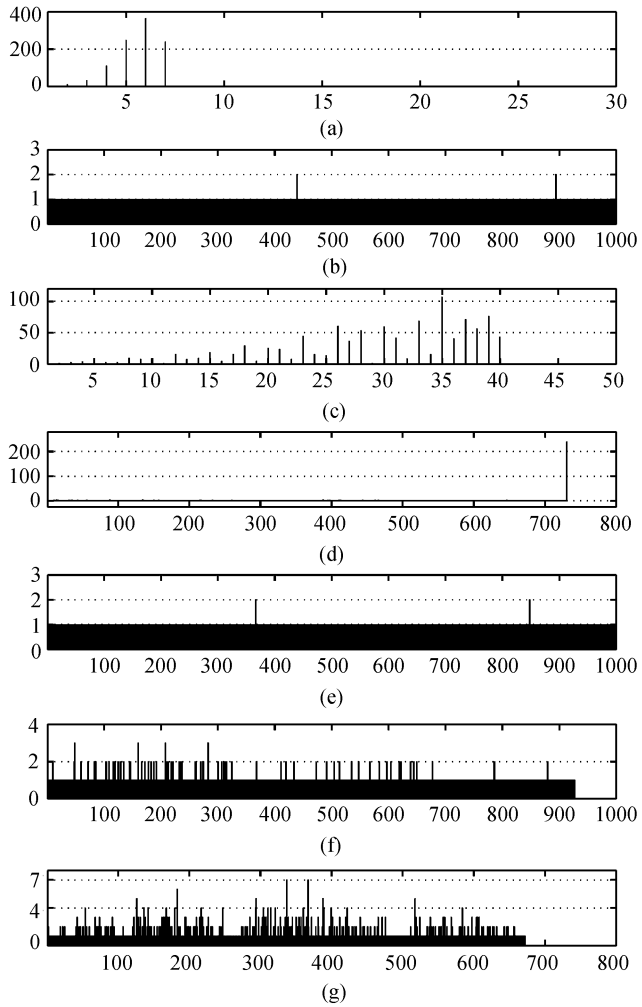


Fig.8. Density of the distribution of ranks achieved by the studied energy functions. Results for a single repetition over sequence 2d4 (2D square lattice). (a) Function D85. (b) Function K99. (c) Function C04. (d) Function L06. (e) Function B08. (f) Function C08. (g) Function I09.

Functions L06 and I09 enable a more fine-grained discrimination, since about 730 and 680 different ranking positions are occupied to classify the totality of conformations, respectively. In the case of function I09, a maximum of seven conformations are assigned to the same rank. On the other hand, the histogram for function L06 presents a high peak indicating that there are about 250 equally ranked conformations. Function L06 is defined as the product of three terms, out of which one corresponds to the number of *H-H* topological contacts, *HHtc* (see Subsection 3.3). Thus, all conformations for which $HHtc = 0$ will share the same energy value, $E_{L06} = 0$. This can be seen as a drawback; function L06 will not be able to discriminate among these conformations even if some of them present better characteristics than the others.

Finally, the histograms for K99, B08 and C08 confirm the high degree of discrimination that these functions provide. Function C08 allows roughly 930 different ranking positions to be assigned. K99 and B08 exhibit the strongest discrimination among all the studied energy functions. The corresponding histograms for these functions reveal that almost every conformation is mapped to a different ranking position. Only a few ranks are assigned to at most two conformations.

### 5.2 HP-Compatibility

Alternative energy functions for the HP model are used in order to perform a more effective exploration through the space of potential protein conformations. Nevertheless, they should remain consistent with the original objective of the HP model of the PSP problem, which consists in minimizing the conventional energy function D85 (by maximizing the number of *H-H* topological contacts, *HHtc*). Therefore, an important issue to be investigated is whether or not these alternative energy formulations are consistent with such an original objective.

The alternative energy functions should not contradict the conventional function D85 at the time of discriminating among potential conformations. Otherwise, the search process could be oriented towards solutions which differ from the original optima in the HP model (false optima can potentially be introduced). In this study, functions that meet this requirement (not contradicting function D85) are said to feature the HP-compatibility property or, in other words, they are HP-compatible. Thus, HP-compatibility can be defined as the capability of an alternative energy function to preserve the conventional rank ordering among potential protein conformations. More formally[10]:

**Definition 1.** *An alternative energy function $E : \mathcal{C}_{\mathcal{F}} \rightarrow \mathbb{R}$ is said to be HP-compatible if and only if $E(c_1) < E(c_2) \Rightarrow E_{D85}(c_1) \leqslant E_{D85}(c_2)$ for every pair of conformations $c_1, c_2 \in \mathcal{C}_{\mathcal{F}}$. Otherwise, if there exists at least a pair of conformations $c_1, c_2 \in \mathcal{C}_{\mathcal{F}}$ such that $E_{D85}(c_1) < E_{D85}(c_2)$ but $E(c_1) > E(c_2)$, then function $E$ is not HP-compatible.*

Note, however, that the case where $E_{D85}(c_1) = E_{D85}(c_2)$ but $E(c_1) \neq E(c_2)$ is not considered as a contradiction. This is a convenient scenario, since the aim of using the alternative function $E$ is to enable a more fine-grained discrimination.

In this subsection, the HP-compatibility property is explored for all the alternative energy functions consi-

[10] By convention, this definition assumes that lower energy values correspond to higher quality conformations.

dered in this study. An experiment is conducted where 1 000 different feasible structures are generated at random and all pairwise comparisons among them are performed. The percentage of such comparisons where the alternative energy function agrees with (does not contradict) the conventional one is computed. The resulting value is to be referred to as relative compatibility ($RC$). Although a value of $RC = 100\%$ does not guarantee the HP-compatibility property for a given function, $RC < 100\%$ is enough to disprove it. To some extent, the $RC$ value allows us to inquire into the severity of the cases where the HP-compatibility property is not satisfied. For all the selected test instances, 100 repetitions of this experiment are performed. The average $RC$ obtained for each of the instances is depicted in Figs. 10 and 11, while Fig.9 provides the overall statistics produced in this experiment.

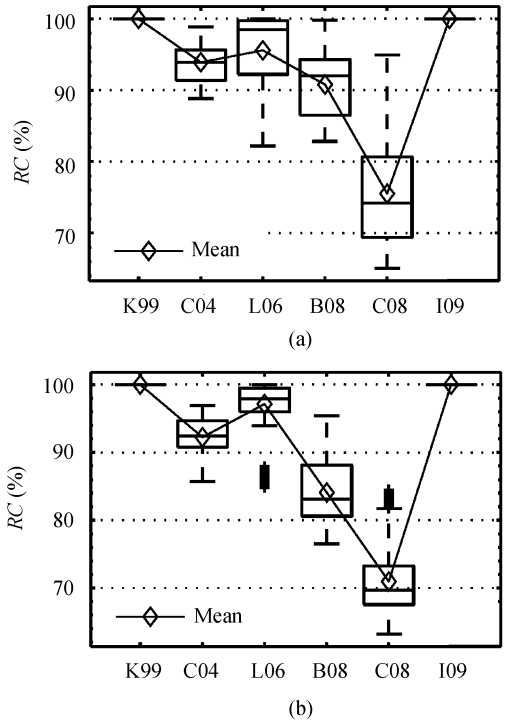


Fig.9. Relative compatibility ($RC$) obtained by each of the alternative energy functions analyzed. Overall statistics for all (a) 2D and (b) 3D test cases.

From Figs. 9 to 11, it is possible to note that functions K99 and I09 show 100% of agreement with the conventional HP energy function for all the instances of this experiment. These results suggest that functions K99 and I09 are HP-compatible. On the other hand, the obtained results reveal that functions C04, L06, B08 and C08 do not present the HP-compatibility property, which becomes more evident with the increasing of problem size.

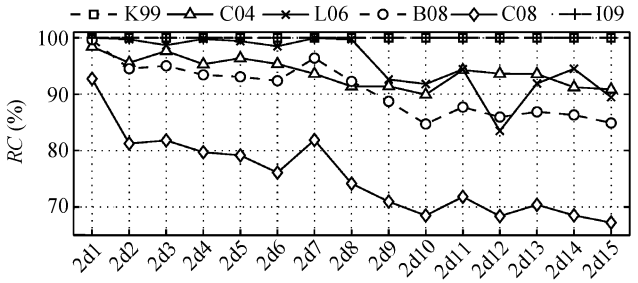


Fig.10. Relative compatibility ($RC$) obtained by each of the alternative energy functions analyzed. Average results for all the 2D test cases.

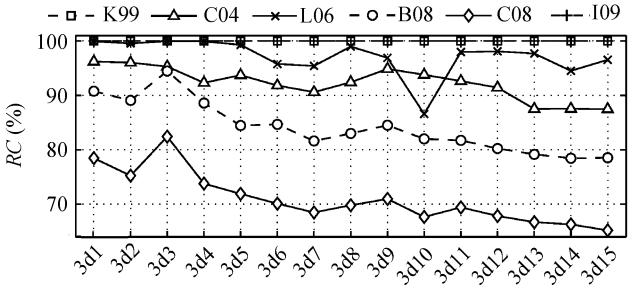


Fig.11. Relative compatibility ($RC$) obtained by each of the alternative energy functions analyzed. Average results for all the 3D test cases.

Function L06 scores very competitive results for the shortest 2D and 3D test sequences. However, its performance declines for the largest test cases, especially when facing sequences 2d12 and 3d10. The average $RC$ values obtained by L06 are almost always above 95%. The performance of function C04 gradually decreases as the problem size increases. The $RC$ values achieved by this approach range from 90% to 95% most of the time. Function B08 presents the second worst overall behavior in this experiment. In the 2D instances, the performance of B08 is above $RC = 90\%$ for the shortest sequences but at around 85% for the largest ones. Regarding the 3D instances, B08 obtains $RC$ values below 85% in most of the cases.

Finally, we want to highlight the poor performance exhibited by function C08. This approach achieves the lowest $RC$ values for all the adopted test cases. The average $RC$ obtained by function C08 is roughly 75% for 2D benchmarks, while it is about 70% for the 3D cases. Fig.12 presents an example scenario where function C08 contradicts the conventional function D85.

In this example, a couple of 2D conformations $c_1$ and $c_2$ for sequence 2d4 are compared with respect to each other by using functions D85 and C08. As a result, the conventional energy function D85 prefers conformation $c_1$ (with $HHtc = 7$) to $c_2$ (with $HHtc = 0$), while function C08 induces the opposite order of preferences between them.

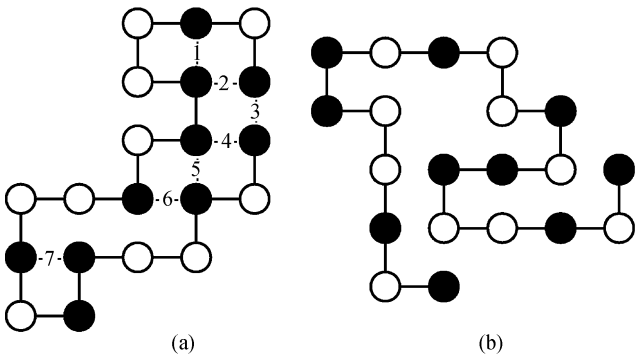


Fig.12. Two conformations $c_1$ and $c_2$ for sequence 2d4 on the 2D square lattice. This is an example where function C08 contradicts function D85: $E_{D85}(c_1) = -7 < E_{D85}(c_2) = 0$ but $E_{C08}(c_1) = 5548 > E_{C08}(c_2) = 5308$. (a) $c_1$. (b) $c_2$.

The low $RC$ values obtained by some functions, particularly C08, suggest serious implications. The lower the $RC$ value, the more likely that the global optimum induced by the alternative function differs from the global optimum of the original problem. Therefore, alternative functions which are not HP-compatible cannot be expected to steer the search process in an effective manner.

### 5.3 Search Performance Using a Basic Local Search Algorithm

A best improvement local search (BILS) algorithm was implemented in order to evaluate the effectiveness of the studied energy functions at guiding the search process (see Algorithm 1). BILS starts with a randomly generated conformation, denoted by $c$. Iteratively, $c$ is replaced by the best among all the improving conformations defined in the neighborhood of $c$, $\mathcal{N}(c)$. The search process stops when given the current conformation $c$ and the adopted neighborhood structure it is not possible to achieve an improvement, i.e., $c$ is locally optimal.

**Algorithm 1.** Best Improvement Local Search (BILS) Algorithm

1: Choose $c \in \mathcal{C}$ uniformly at random
2: **repeat**
3: $c \leftarrow Best\_Improvement(\mathcal{N}(c))$
4: **until** no improvement is possible

As stated at the beginning of Section 5, only solutions encoding feasible conformations are considered in this study. Hence, the initial solutions for the BILS algorithm were generated using the backtracking procedure proposed in [22]. The implemented neighborhood structure $\mathcal{N}(c)$ is defined by all conformations which can be reached through single 1-variable perturbations of $c$, i.e., $\mathcal{N}(c) = \{c' \in \mathcal{C}_{\mathcal{F}} \mid h(c, c') = 1\}$, where $h(c, c')$ denotes the Hamming distance between $c$ and $c'$. Given a protein sequence of length $L$, the size of such a neighborhood is $|\mathcal{N}(c)| = 3(L-1)$ in the 2D square lattice and $|\mathcal{N}(c)| = 5(L-1)$ for the 3D case.

The motivation for using such a simple BILS algorithm is as follows. On the one hand, BILS seems to be a suitable algorithm for evaluating the impact of varying the evaluation scheme. Once the neighborhood structure has been defined, the behavior and performance of the algorithm will be mainly determined by the discrimination capabilities of the different energy functions. "A local search is effective if it is able to find good local minima"[57]. BILS stops at a local optimum, and the effectiveness of the discrimination will depend on the characteristics of such a local optimum. Moreover, due to the low degree of discrimination provided by some of the functions, the search process can be expected to stop early (after a reduced number of iterations). On the other hand, no additional parameters of the algorithm have to be adjusted, which avoids affecting (neither negatively nor positively) the behavior induced by the studied energy functions through parameter settings.

The behavior of the BILS algorithm is evaluated when using each of the studied energy functions. A total of 100 independent executions are performed. Fig.13 presents the results obtained for all the 2D instances, while Fig.14 details the results for the 3D case. Plots in these figures show the average number of $H$-$H$ topological contacts ($HHtc$) achieved by the algorithm as the search progressed (at each iteration), for each considered test case.

From Figs. 13 and 14, it is possible to derive some general conclusions. As expected, the conventional energy function D85 presents a limited performance for this experiment. For all the test instances (except for sequence 3d9), the algorithm reaches the lowest number of iterations due to the poor discrimination that function D85 provides (see Subsection 5.1). In most cases, however, the poorest performance of the algorithm is obtained when using function C08. Although functions B08 and C04 behave better than function D85 in most of the 2D instances, these functions report a poorer search performance than D85 for some of the 3D test cases. Function L06 obtains very competitive results most of the time. L06 allows the algorithm to score the highest $HHtc$ values for some of the test cases (e.g., 2d3, 2d5, 2d10, 3d2), while showing a slight inferior performance for some other instances (e.g., 2d1, 2d7, 3d10). Finally, it is possible to highlight the promising behavior that functions I09 and K99 consistently exhibit for all the considered test cases.

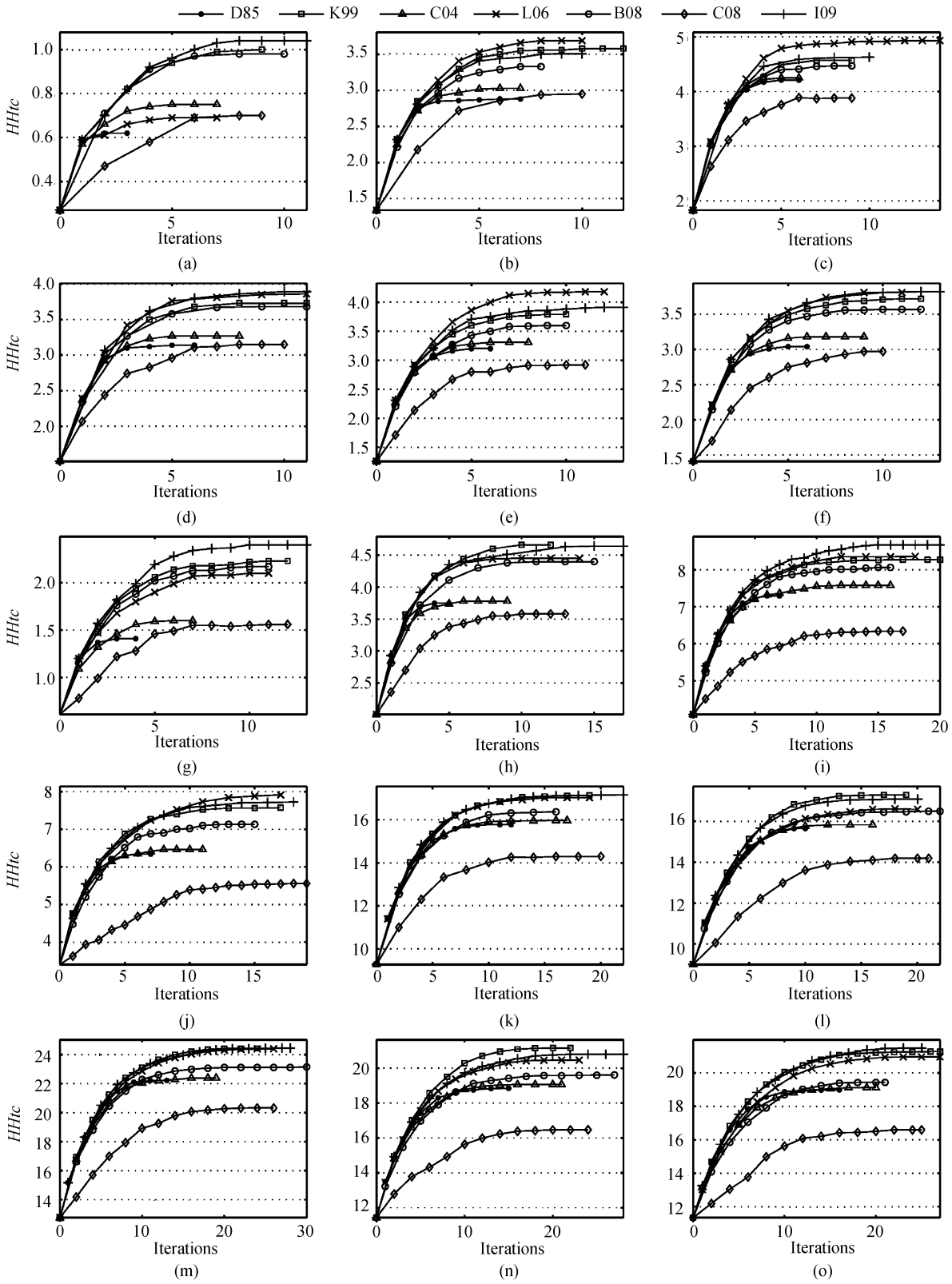


Fig.13. Results of the BILS on the 2D instances. Number of $H$-$H$ topological contacts ($HHtc$) obtained at each iteration. Average of 100 independent executions. (a) 2d1. (b) 2d2. (c) 2d3. (d) 2d4. (e) 2d5. (f) 2d6. (g) 2d7. (h) 2d8. (i) 2d9. (j) 2d10. (k) 2d11. (l) 2d12. (m) 2d13. (n) 2d14. (o) 2d15.

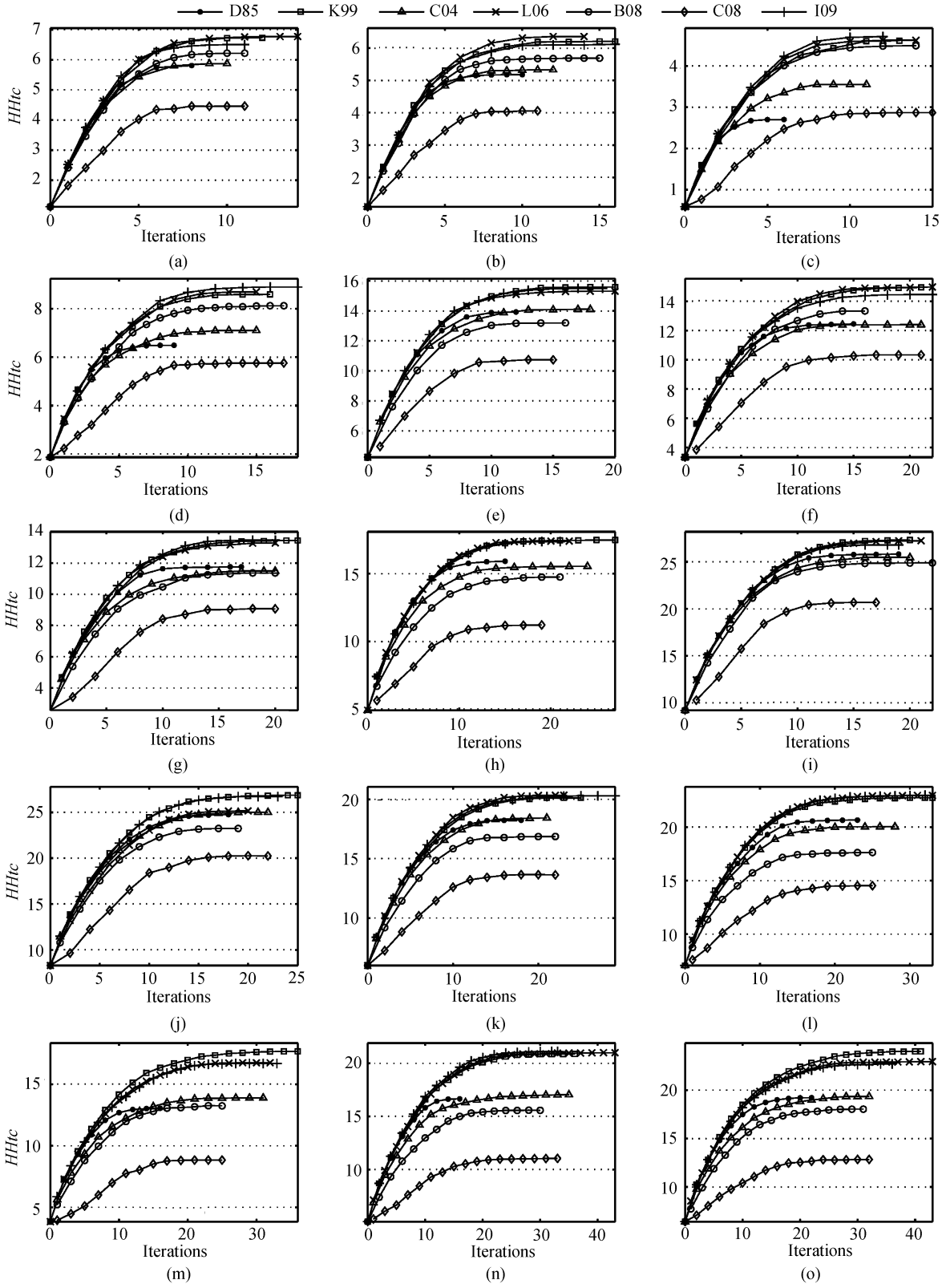


Fig.14. Results of the BILS on the 3D instances. Number of $H$-$H$ topological contacts (*HHtc*) obtained at each iteration. Average of 100 independent executions. (a) 3d1. (b) 3d2. (c) 3d3. (d) 3d4. (e) 3d5. (f) 3d6. (g) 3d7. (h) 3d8. (i) 3d9. (j) 3d10. (k) 3d11. (l) 3d12. (m) 3d13. (n) 3d14. (o) 3d15.

A more detailed comparison and the results of the statistical significance analysis are provided in Tables 3 and 4. For the different analyzed energy functions and all the adopted test cases, these tables detail the best obtained energy value (Best), the number of BILS executions where this solution quality is reached (freq), and the arithmetic mean (Mean). The obtained $OAP$ values are presented at the bottom of the tables. Each time that a significant performance difference exists with respect to the conventional function D85, the mean energy of the corresponding alternative function is either marked + or marked − depending on whether such a difference favors the alternative function or not. In addition, the lowest average energy for each of the instances appears shaded in the tables.

Tables 3 and 4 confirm the superiority that functions K99, I09 and L06 have shown in this experiment. In the vast majority of the instances, it can be seen from the tables that functions K99, I09 and L06 significantly improve the performance of the BILS algorithm with respect to the conventional function D85. There are no significant differences between functions D85 and C04 except for sequences 3d3 and 3d4, in both cases favoring C04. Function B08 scores significantly better results than function D85 in 9 out of the 15 2D instances. Note, however, that this function is significantly outperformed by function D85 in five of the largest 3D test cases. Finally, it can also be confirmed the poor performance presented by function C08. Function C08 performs significantly worse than function D85 for the largest 2D instances and for all but one of the 3D cases.

### 5.4 Search Performance Using the Iterated Local Search Algorithm

In Subsection 5.3, a basic local search algorithm was employed as a first step in analyzing the effectiveness of the studied energy functions at guiding the search process. Through local search it is possible to converge towards local optima. However, the performance of these algorithms is usually unsatisfactory in terms of finding global optimum solutions[57-58]. Therefore, it is required to implement additional strategies to foster exploration and to allow the search process escaping from local optima. One possible strategy consists in iteratively applying local search each time starting from a different initial solution, such as it is done in the iterated local search (ILS) algorithm[59-61].

In this subsection, a basic ILS algorithm is used for inquiring into the suitability of the studied energy functions (outlined in Algorithm 2). The ILS algorithm starts with a feasible conformation generated at random[11], denoted as $c$. Then, a local search strategy (embedded heuristic) is applied to $c$ until a local optimum $c^*$ is found. A perturbation $c'$ of the current local optimum $c^*$ is obtained and used as a starting point of another round of local search. After each local search the new local optimum solution found $c'^*$ may be accepted as the new incumbent solution $c^*$, based on a given acceptance criterion. This iterative procedure is repeated until a given stop condition is met.

**Algorithm 2.** Iterated Local Search (ILS) Algorithm

```
1: choose c ∈ C uniformly at random
2: c* ← LocalSearch(c)
3: repeat
4:     c' ← Perturbation(c*)
5:     c'* ← LocalSearch(c')
6:     c* ← AcceptanceCriterion(c*, c'*)
7: until ⟨stop condition⟩
```

In order to implement the ILS algorithm, three basic components have to be defined: the local search strategy, the perturbation strength, and the acceptance criterion. In this study, these components are defined as follows:

• *Local Search.* The best improvement local search (BILS) algorithm described in Subsection 5.3 is adopted as the embedded heuristic.

• *Perturbation Strength.* Six different values for the perturbation strength are considered: $\{2, 3, 4, 6, 8, 10\}$. The perturbation strength refers to the number of encoding positions of the conformation which are to be affected by the perturbation.

• *Acceptance Criterion.* Three different acceptance criteria are explored:

– IMP: the new local optimum $c'^*$ is accepted if it has a better energy value than the incumbent solution $c^*$.

– IEQ: the new local optimum $c'^*$ is accepted if it is at least as good as the incumbent solution $c^*$.

– ALL: the new local optimum $c'^*$ is always accepted.

The three different acceptance criteria, together with the six considered values for the perturbation strength, lead to a total of 18 parameter configurations of the ILS algorithm. All these parameter configurations are evaluated in order to identify the most appropriate conditions for the compared approaches. In all the cases, the algorithm is allowed to run until a maximum number of $5 \times 10^5$ solution evaluations is reached, and 50 independent executions are performed. Figs. 15 and 16 present (2D and 3D instances, respectively) the overall average performance ($OAP$) measure obtained by each of the studied energy functions for the different parameter settings of the ILS algorithm. Higher $OAP$ values are preferred, see Section 4.

[11] It is generated using the backtracking algorithm proposed in [22].

**Table 3.** Performance of the BILS Algorithm When Using the Seven Studied Energy Formulations (2D Test Cases)

| | D85 | | K99 | | C04 | | L06 | | B08 | | C08 | | I09 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Seq. | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean |
| 2d1 | −3(1) | −0.6 | −3(2) | −1.0 + | −3(1) | −0.8 | −3(2) | −0.7 | −3(2) | −1.0 + | −2(12) | −0.7 | −3(3) | −1.0 + |
| 2d2 | −7(1) | −2.9 | −7(3) | −3.6 + | −7(1) | −3.0 | −7(3) | −3.7 + | −7(3) | −3.3 + | −7(1) | −3.0 | −7(2) | −3.5 + |
| 2d3 | −8(1) | −4.2 | −8(1) | −4.6 | −8(1) | −4.2 | −8(1) | −4.9 + | −7(4) | −4.5 | −7(1) | −3.9 | −8(1) | −4.6 + |
| 2d4 | −7(1) | −3.1 | −7(3) | −3.7 + | −7(1) | −3.3 | −7(2) | −3.9 + | −7(3) | −3.7 + | −7(1) | −3.1 | −7(3) | −3.9 + |
| 2d5 | −6(7) | −3.2 | −7(1) | −3.8 + | −6(7) | −3.3 | −7(4) | −4.2 + | −7(1) | −3.6 | −6(6) | −2.9 | −7(4) | −3.9 + |
| 2d6 | −7(2) | −3.0 | −7(1) | −3.7 + | −7(1) | −3.2 | −7(1) | −3.8 + | −7(1) | −3.6 + | −6(2) | −3.0 | −7(2) | −3.8 + |
| 2d7 | −5(1) | −1.4 | −7(1) | −2.2 + | −6(1) | −1.6 | −7(1) | −2.1 + | −7(1) | −2.2 + | −7(1) | −1.6 | −7(1) | −2.4 + |
| 2d8 | −7(2) | −3.8 | −8(1) | −4.7 + | −7(2) | −3.8 | −7(6) | −4.5 + | −7(7) | −4.4 + | −7(4) | −3.6 | −9(1) | −4.6 + |
| 2d9 | −12(3) | −7.3 | −12(4) | −8.3 + | −12(2) | −7.6 | −13(1) | −8.4 + | −12(4) | −8.1 + | −11(7) | −6.3 − | −15(1) | −8.7 + |
| 2d10 | −10(6) | −6.3 | −13(1) | −7.6 + | −11(1) | −6.5 | −12(3) | −7.9 + | −13(1) | −7.1 + | −11(1) | −5.6 − | −13(2) | −7.7 + |
| 2d11 | −22(2) | −15.8 | −24(3) | −17.1 + | −22(1) | −16.0 | −24(3) | −17.0 + | −25(1) | −16.4 | −24(1) | −14.3 − | −25(1) | −17.1 + |
| 2d12 | −22(1) | −15.7 | −24(1) | −17.3 + | −22(1) | −15.8 | −22(2) | −16.6 + | −21(4) | −16.5 + | −22(1) | −14.2 − | −23(1) | −17.1 + |
| 2d13 | −30(2) | −22.2 | −35(1) | −24.4 + | −30(3) | −22.4 | −35(1) | −24.4 + | −31(2) | −23.1 | −33(1) | −20.3 − | −35(1) | −24.4 + |
| 2d14 | −28(1) | −18.8 | −30(1) | −21.1 + | −26(4) | −19.1 | −29(1) | −20.5 + | −28(1) | −19.6 | −25(1) | −16.5 − | −29(1) | −20.8 + |
| 2d15 | −26(2) | −19.0 | −28(3) | −21.3 + | −29(1) | −19.1 | −28(1) | −20.9 + | −26(1) | −19.4 | −22(5) | −16.6 − | −27(3) | −21.5 + |
| *OAP* | 33.70% | | 39.72% | | 34.77% | | 39.64% | | 37.87% | | 31.54% | | 40.29% | |

**Table 4.** Performance of the BILS Algorithm When Using the Seven Studied Energy Formulations (3D Test Cases)

| | D85 | | K99 | | C04 | | L06 | | B08 | | C08 | | I09 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Seq. | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean |
| 3d1 | −10(2) | −5.8 | −11(2) | −6.7 + | −10(1) | −5.9 | −11(1) | −6.8 + | −11(1) | −6.2 | −9(2) | −4.5 − | −10(4) | −6.5 + |
| 3d2 | −9(4) | −5.2 | −11(1) | −6.2 + | −10(1) | −5.3 | −10(3) | −6.4 + | −9(6) | −5.7 + | −7(7) | −4.1 − | −10(1) | −6.1 + |
| 3d3 | −7(2) | −2.7 | −9(1) | −4.6 + | −7(3) | −3.5 + | −8(2) | −4.6 + | −9(1) | −4.5 + | −7(3) | −2.9 | −8(2) | −4.7 + |
| 3d4 | −12(2) | −6.5 | −13(2) | −8.6 + | −13(1) | −7.1 + | −14(1) | −8.7 + | −14(2) | −8.1 + | −13(1) | −5.8 − | −15(1) | −8.9 + |
| 3d5 | −22(1) | −13.9 | −23(1) | −15.6 + | −21(1) | −14.1 | −22(2) | −15.3 + | −22(1) | −13.2 | −17(2) | −10.7 − | −22(1) | −15.5 + |
| 3d6 | −19(4) | −12.4 | −22(1) | −14.9 + | −19(1) | −12.4 | −21(3) | −15.0 + | −19(4) | −13.3 + | −18(1) | −10.3 − | −21(3) | −14.5 + |
| 3d7 | −18(2) | −11.8 | −20(2) | −13.4 + | −18(1) | −11.5 | −22(1) | −13.3 + | −17(2) | −11.4 | −17(1) | −9.1 − | −18(4) | −13.5 + |
| 3d8 | −23(2) | −15.9 | −25(1) | −17.5 + | −22(2) | −15.6 | −24(1) | −17.4 + | −23(1) | −14.8 − | −19(1) | −11.2 − | −24(2) | −17.4 + |
| 3d9 | −36(2) | −25.8 | −38(4) | −27.3 + | −36(1) | −25.5 | −36(3) | −27.2 + | −36(1) | −24.9 | −32(1) | −20.7 − | −38(2) | −26.8 |
| 3d10 | −34(1) | −24.8 | −38(1) | −26.9 + | −36(1) | −25.0 | −35(2) | −25.1 | −33(1) | −23.2 − | −33(1) | −20.2 − | −37(1) | −26.7 + |
| 3d11 | −28(2) | −18.2 | −31(1) | −20.1 + | −26(2) | −18.4 | −29(1) | −20.3 + | −27(1) | −16.9 − | −25(1) | −13.6 − | −31(2) | −20.3 + |
| 3d12 | −29(5) | −20.6 | −31(6) | −22.7 + | −30(1) | −20.0 | −33(1) | −22.9 + | −27(2) | −17.6 − | −22(1) | −14.5 − | −34(2) | −23.0 + |
| 3d13 | −22(1) | −13.0 | −28(1) | −17.6 + | −22(1) | −13.9 | −24(4) | −16.7 + | −22(2) | −13.2 | −17(1) | −8.8 − | −24(3) | −16.7 + |
| 3d14 | −24(2) | −16.6 | −30(1) | −20.9 + | −29(1) | −17.1 | −32(1) | −21.0 + | −26(1) | −15.6 − | −20(1) | −11.0 − | −34(1) | −21.1 + |
| 3d15 | −30(1) | −19.2 | −36(1) | −24.0 + | −30(1) | −19.3 | −35(2) | −22.9 + | −28(2) | −18.0 | −21(2) | −12.8 − | −32(3) | −22.7 + |
| *OAP* | 35.16% | | **41.98%** | | 36.14% | | 41.71% | | 36.51% | | 27.58% | | 41.62% | |

Among the alternative energy functions, Fig.15 and Fig.16 show that K99, L06 and I09 consistently present the best performance for the different parameter configurations of the ILS. In the 2D instances, the performance of function B08 is competitive for most of the ILS configurations. In contrast, this function exhibits a low performance in all cases when facing the 3D instances. Function C08 obtains the lowest $OAP$ values in most of the cases, followed by function C04. Functions C08 and C04 are thus the worst performers of this experiment. Regarding the conventional energy function D85, an interesting behavior can be observed when comparing the results obtained using the different acceptance criteria. While the ranking among the alternative energy functions remains consistent in most of the cases from one acceptance criterion to another, there is a significant increase in the performance of function D85 when using the IEQ acceptance criterion. That is, the IEQ acceptance criterion allows the algorithm to exploit the low discrimination associated with function D85 as a means of escaping from local optima.

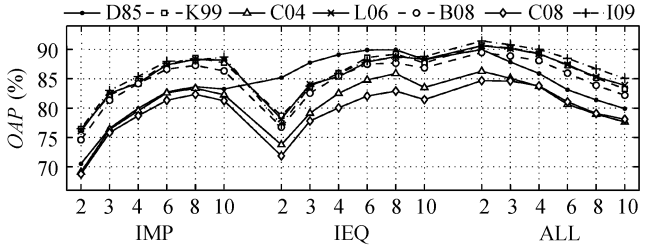


Fig.15. Overall average performance ($OAP$) obtained for all parameter configurations of the ILS algorithm. 2D test cases.

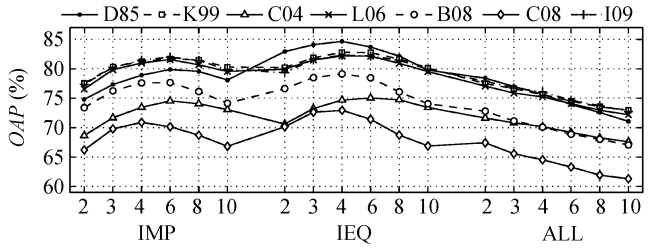


Fig.16. Overall average performance ($OAP$) obtained for all parameter configurations of the ILS algorithm. 3D test cases.

In order to provide a more detailed analysis, the parameters adjustment which allows each of the studied energy functions to reach the highest $OAP$ value has been selected. Table 5 summarizes the selected ILS configurations for the next experiment.

Tables 6 and 7 detail the obtained results for all 2D and 3D test cases, respectively. For each instance, these tables show the best obtained energy value, the number of times that this solution is found and the arithmetic mean achieved using the different energy functions. Also, the $OAP$ measure is presented at the bottom of the tables. In these tables, values marked with + highlight a statistically significant increase in performance achieved by the alternative energy function with regard to the conventional function D85. Conversely, values marked with − indicate that a statistically significant performance decrease is obtained as a consequence of using the alternative formulation. Additionally, the best average performance (the lowest average energy) for each test case has been shaded in these tables.

**Table 5.** Selected Parameter Settings for the ILS Algorithm

| | 2D Benchmarks | | 3D Benchmarks | |
|---|---|---|---|---|
| | Acceptance Criterion | Perturbation Strength | Acceptance Criterion | Perturbation Strength |
| D85 | ALL | 2 | IEQ | 4 |
| K99 | ALL | 2 | IEQ | 4 |
| C04 | ALL | 2 | IEQ | 6 |
| L06 | ALL | 2 | IEQ | 4 |
| B08 | ALL | 2 | IEQ | 4 |
| C08 | ALL | 2 | IEQ | 4 |
| I09 | ALL | 2 | IEQ | 4 |

From Table 6, it is possible to observe that function I09 reaches the lowest average energy on 73.33% of the studied 2D instances (11 out of 15), obtaining the highest $OAP$ value. In five of the instances, the improvements obtained by function I09 are statistically significant with respect to the conventional energy function D85. The second best performer is function K99, which shows the best average performance for seven of the instances and significantly improves the results of function D85 in three other cases. Function L06 achieves significantly better results than function D85 for five of the instances; note, however, that there is a significant difference against function L06 in four of the largest test cases. Slightly similar results are obtained by function B08. Although the conventional function D85 does not present a remarkable performance, the results of this function are still considered competitive. Finally, the poorest performances are obtained by functions C04 and C08, whose results are significantly worse than those of the conventional function D85 in most of the cases.

A quite different scenario can be observed regarding the 3D test cases. It can be seen from Table 7 that the conventional energy function D85 scores the best average performance for all the considered test cases. The statistical analysis indicates that function D85 significantly outperforms all the alternative energy functions in the vast majority of the cases. Among the alternative functions, the best results are obtained by K99, followed by functions I09 and L06. Finally, the worst overall behavior is presented by functions B08, C04 and particularly C08.

**Table 6.** Detailing the Results of the ILS Algorithm When Using the Seven Studied Energy Formulations (2D Test Cases)

| | D85 | | K99 | | C04 | | L06 | | B08 | | C08 | | I09 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Seq. | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean |
| 2d1 | −4(50) | −4.0 | −4(50) | −4.0 | −4(49) | −4.0 | −4(50) | −4.0 | −4(50) | −4.0 | −4(46) | −3.9 − | −4(50) | −4.0 |
| 2d2 | −8(50) | −8.0 | −8(50) | −8.0 | −8(49) | −8.0 | −8(50) | −8.0 | −8(50) | −8.0 | −8(36) | −7.7 − | −8(50) | −8.0 |
| 2d3 | −9(44) | −8.9 | −9(47) | −8.9 | −9(46) | −8.9 | −9(50) | −9.0 + | −9(49) | −9.0 | −9(48) | −9.0 | −9(47) | −8.9 |
| 2d4 | −9(50) | −9.0 | −9(50) | −9.0 | −9(50) | −9.0 | −9(50) | −9.0 | −9(50) | −9.0 | −9(45) | −8.9 − | −9(50) | −9.0 |
| 2d5 | −10(50) | −10.0 | −10(50) | −10.0 | −10(49) | −10.0 | −10(50) | −10.0 | −10(50) | −10.0 | −10(27) | −9.5 − | −10(50) | −10.0 |
| 2d6 | −9(47) | −8.9 | −9(50) | −9.0 | −9(42) | −8.8 | −9(50) | −9.0 | −9(50) | −9.0 | −9(47) | −8.9 | −9(50) | −9.0 |
| 2d7 | −8(36) | −7.7 | −8(47) | −7.9 + | −8(25) | −7.5 − | −8(50) | −8.0 + | −8(50) | −8.0 + | −8(49) | −8.0 + | −8(50) | −8.0 + |
| 2d8 | −14(2) | −12.3 | −14(2) | −12.4 | −13(1) | −11.2 − | −14(7) | −12.9 + | −14(9) | −12.8 + | −14(5) | −12.4 | −14(15) | −13.0 + |
| 2d9 | −21(3) | −18.8 | −22(1) | −19.6 + | −20(1) | −17.4 − | −21(4) | −19.5 + | −21(5) | −19.1 | −21(1) | −17.7 − | −22(1) | −20.1 + |
| 2d10 | −20(1) | −18.2 | −21(1) | −18.3 | −19(1) | −16.8 − | −21(1) | −18.4 | −19(1) | −17.1 − | −17(4) | −15.5 − | −21(1) | −18.7 + |
| 2d11 | −33(7) | −31.3 | −34(1) | −31.5 | −33(1) | −29.3 − | −33(2) | −30.7 − | −33(3) | −30.9 | −32(1) | −27.7 − | −34(2) | −31.0 |
| 2d12 | −34(1) | −30.5 | −35(2) | −31.2 + | −34(1) | −29.3 − | −35(1) | −32.1 + | −33(1) | −29.3 − | −30(3) | −26.4 − | −35(2) | −32.2 + |
| 2d13 | −46(3) | −42.6 | −47(1) | −43.0 | −45(1) | −40.2 − | −46(1) | −41.9 − | −46(1) | −42.2 | −43(2) | −38.7 − | −46(1) | −42.9 |
| 2d14 | −42(3) | −38.6 | −41(2) | −38.2 | −38(1) | −34.4 − | −40(3) | −37.2 − | −40(3) | −37.2 − | −38(1) | −33.1 − | −41(4) | −38.4 |
| 2d15 | −42(1) | −39.1 | −42(3) | −39.0 | −39(1) | −35.1 − | −41(1) | −38.2 − | −41(2) | −37.0 − | −37(1) | −31.7 − | −42(5) | −39.3 |
| *OAP* | 89.98% | | 90.70% | | 86.28% | | 90.61% | | 89.48% | | 84.68% | | **91.43%** | |

**Table 7.** Detailing the Results of the ILS Algorithm When Using the Seven Studied Energy Formulations (3D Test Cases)

| | D85 | | K99 | | C04 | | L06 | | B08 | | C08 | | I09 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Seq. | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean | Best(freq) | Mean |
| 3d1 | −11(50) | −11.0 | −11(50) | −11.0 | −11(50) | −11.0 | −11(50) | −11.0 | −11(50) | −11.0 | −11(50) | −11.0 | −11(50) | −11.0 |
| 3d2 | −13(50) | −13.0 | −13(49) | −13.0 | −13(37) | −12.7 − | −13(50) | −13.0 | −13(49) | −13.0 | −13(43) | −12.8 − | −13(50) | −13.0 |
| 3d3 | −9(50) | −9.0 | −9(50) | −9.0 | −9(41) | −8.8 − | −9(50) | −9.0 | −9(50) | −9.0 | −9(49) | −9.0 | −9(50) | −9.0 |
| 3d4 | −18(49) | −18.0 | −18(38) | −17.6 − | −18(8) | −16.2 − | −18(38) | −17.6 − | −18(27) | −17.3 − | −18(16) | −16.8 − | −18(37) | −17.6 − |
| 3d5 | −33(4) | −31.1 | −33(4) | −30.4 − | −31(3) | −27.9 − | −34(1) | −30.3 − | −33(1) | −30.0 − | −30(2) | −26.8 − | −33(1) | −30.1 − |
| 3d6 | −31(13) | −29.6 | −31(3) | −28.7 − | −29(5) | −26.4 − | −31(5) | −29.0 − | −31(2) | −28.4 − | −31(2) | −27.7 − | −31(7) | −28.9 − |
| 3d7 | −32(1) | −29.2 | −31(3) | −28.9 | −30(1) | −25.5 − | −32(3) | −28.2 − | −32(1) | −27.3 − | −31(1) | −23.8 − | −30(8) | −28.1 − |
| 3d8 | −40(2) | −36.2 | −40(1) | −35.3 − | −36(1) | −31.2 − | −39(1) | −34.5 − | −40(1) | −33.5 − | −35(1) | −29.9 − | −40(1) | −35.1 − |
| 3d9 | −52(1) | −48.3 | −51(1) | −47.6 | −49(1) | −44.5 − | −50(5) | −47.2 − | −50(3) | −46.3 − | −49(4) | −44.8 − | −51(4) | −47.6 |
| 3d10 | −56(1) | −50.2 | −54(1) | −48.4 − | −49(1) | −43.7 − | −55(1) | −48.6 − | −52(1) | −45.6 − | −50(2) | −41.6 − | −54(1) | −49.2 − |
| 3d11 | −45(1) | −41.4 | −44(2) | −39.7 − | −40(1) | −35.4 − | −44(2) | −39.5 − | −43(1) | −37.1 − | −40(1) | −32.8 − | −45(1) | −39.5 − |
| 3d12 | −57(2) | −50.5 | −55(2) | −48.7 − | −50(1) | −41.4 − | −54(1) | −48.2 − | −50(3) | −43.0 − | −42(1) | −35.7 − | −54(2) | −48.1 − |
| 3d13 | −47(1) | −40.6 | −46(1) | −39.6 | −40(1) | −31.8 − | −45(2) | −38.6 − | −42(1) | −36.4 − | −38(1) | −29.0 − | −44(4) | −39.0 − |
| 3d14 | −55(1) | −49.4 | −54(2) | −48.3 | −49(1) | −38.2 − | −55(1) | −45.8 − | −47(1) | −39.5 − | −47(1) | −33.1 − | −57(1) | −46.4− |
| 3d15 | −61(1) | −53.8 | −59(1) | −50.0 − | −50(1) | −40.1 − | −58(3) | −50.0 − | −55(1) | −44.7 − | −48(1) | −37.0 − | −57(1) | −49.0 − |
| *OAP* | **84.67%** | | 82.80% | | 75.04% | | 82.16% | | 79.12% | | 72.94% | | 82.31% | |

The obtained results confirm that an effective evaluation scheme is essential in order to guide the search process towards high quality conformations. For different parameter configurations of the ILS algorithm, the best results are obtained using alternative energy functions which provide a fine-grained discrimination. Nevertheless, a particular acceptance criterion (IEQ, in this case) increases the performance of the ILS algorithm when using the conventional energy function, D85. Using such an acceptance criterion, the results of function D85 are statistically superior compared with those obtained by the different alternative functions. This suggests that it is possible to take advantage of the low degree of discrimination provided by the conventional energy formulation of the HP model.

## 6 Conclusions and Future Work

The conventional energy function of the HP model is known to provide a very poor discrimination among potential conformations. Nevertheless, an effective evaluation scheme is an essential component of metaheuristics, being the responsible for steering the search process towards promising regions of the solutions space. Therefore, alternative formulations of the energy function have been proposed in the literature to cope with this issue. This paper presented the results of a comparative study where seven different evaluation functions for the HP model were considered.

The first step in this study was concerned with the analysis of the degree of discrimination that each of the considered energy functions provides. Through such an analysis it is possible to confirm the poor discrimination capabilities of the conventional energy function of the HP model, D85, which has been the main motivation for exploring alternative energy formulations. All the alternative functions were found to provide a more fine-grained discrimination. From the obtained results, the most discriminative functions are K99 and B08, followed by C08 and I09.

The HP-compatibility property was defined and investigated for each alternative energy function. This important property refers to the capability of an alternative energy function to preserve a rank ordering among potential conformations which is consistent with the original objective of the HP model. The obtained results suggest that functions K99 and I09 feature this property. Very competitive results were also obtained by function L06. However, this was not the case for functions C04, B08 and particularly C08, which obtained the worst results in the experiment. Alternative energy functions which are not HP-compatible may not be able to guide the search process properly since they can potentially introduce a false optimum.

The effectiveness of the studied energy functions to guide the search process was examined using a best improvement local search (BILS) algorithm. The conventional energy function D85 exhibited a low performance for this experiment. In most of the adopted test cases, however, the worst performance of the algorithm was obtained when using the alternative function C08. Also, functions B08 and C04 showed a poor search performance for most of the instances. In contrast, the alternative functions I09, L06 and K99 consistently presented a very promising behavior.

In order to further explore the suitability of the studied energy functions, a more sophisticated metaheuristic was implemented: iterated local search (ILS). In most of the cases, the results of the ILS were similar to those obtained in the previous experiment using the BILS algorithm. Among the alternative energy functions, K99, I09 and L06 consistently exposed a promising behavior, while functions B08, C04 and particularly C08 presented the worst overall performance in this test. On the other side, the results obtained for the conventional function D85 suggest that, using a proper acceptance criterion, it is possible to exploit the neutrality of the search landscape[13,62] induced by the low discrimination of this function.

From this study, it is possible to derive some general conclusions. First, intensity of discrimination does not necessarily imply effectiveness at guiding the search process. Even when functions K99, B08, C08 and I09 were all identified to provide a strong discrimination, only K99 and I09 presented a promising search behavior. In contrast, functions B08 and C08 showed a poor search performance in most of the cases. Such a poor performance can be explained by the fact that functions B08 and C08 are not HP-compatible. Function C04 is also not HP-compatible; the low discrimination capabilities of C04 gives further explanation to the reduced search performance obtained when using this function. Finally, function L06 obtained very competitive results in terms of both the degree of discrimination and HP-compatibility. As a consequence, function L06 consistently competed at the top of the ranking regarding search performance together with functions K99 and I09. Therefore, the degree of discrimination and the HP-compatibility property were found to be useful in explaining the success or failure of the studied energy functions at guiding the search process.

The conventional energy function D85 presented a limited search performance for the BILS algorithm and most parameter configurations of the ILS. This supports the relevance of exploring alternative evaluation schemes for the HP model. There exists evidence in the literature, however, which suggests that the neutrality

property of a fitness landscape can be exploited for designing more efficient search algorithms[62-68]. The performance of function D85 for some parameter configurations of the ILS provides additional clues in this regard. Therefore, future work may focus on investigating how to benefit from a fine-grained discrimination, at the same time that the inherent neutrality of the HP model can be exploited. Finally, an interesting research direction involves the evaluation of how some characteristics of the fitness landscape (e.g., neutrality, ruggedness[13,62,69]) change when using the different evaluation functions. Such an analysis would certainly be helpful to further support the findings of the study presented in this paper.

**Acknowledgment** The authors would like to thank the anonymous reviewers for their valuable feedback that greatly contributes to improving this paper. Also, the first author acknowledges the support from CONACyT through a scholarship to pursue graduate studies at the Information Technology Laboratory, CINVESTAV-Tamaulipas.

## References

[1] Anfinsen C. Principles that govern the folding of protein chains. *Science*, 1973, 181(4096): 223-230.

[2] Chandru V, DattaSharma A, Kumar V. The algorithmics of folding proteins on lattices. *Discrete Applied Mathematics*, 2003, 127(1): 145-161.

[3] Kolinski A, Skolnick J. Reduced models of proteins and their applications. *Polymer*, 2004, 45(2): 511-524.

[4] Hart W, Newman A. Protein structure prediction with lattice models. In *Handbook of Computational Molecular Biology*, Chapman & Hall/CRC, 2005.

[5] Clementi C. Coarse-grained models of protein folding: Toy models or predictive tools? *Current Opinion in Structural Biology*, 2008, 18(1): 10-15.

[6] Pierri C, De Grassi A, Turi A. Lattices for ab initio protein structure prediction. *Proteins: Structure, Function, and Bioinformatics*, 2008, 73(2): 351-361.

[7] Dill K A. Theory for the folding and stability of globular proteins. *Biochemistry*, 1985, 24(6): 1501-1509.

[8] Lau K, Dill K A. A lattice statistical mechanics model of the conformational and sequence spaces of proteins. *Macromolecules*, 1989, 22(10): 3986-3997.

[9] Berger B, Leighton T. Protein folding in the hydrophobic-hydrophilic (HP) model is NP-complete. In *Proc. International Conference on Research in Computational Molecular Biology*, March 1998, pp.30-39.

[10] Crescenzi P, Goldman D, Papadimitriou C *et al.* On the complexity of protein folding. In *Proc. the 30th ACM Symposium on Theory of Computing*, May 1998, pp.597-603.

[11] Stadler P F. Correlation in landscapes of combinatorial optimization problems. *Europhysics Letters*, 1992, 20(6): 479-482.

[12] Hoos H, Stützle T. Stochastic Local Search: Foundations And Applications. Morgan Kaufmann, 2004.

[13] Pitzer E, Affenzeller M. A comprehensive survey on fitness landscape analysis. In *Studies in Computational Intelligence 378*, Fodor J, Klempous R, Araújo C (eds.), Springer, 2012, pp.161-191.

[14] Krasnogor N, Hart W, Smith J *et al.* Protein structure prediction with evolutionary algorithms. In *Proc. Conf. Genetic and Evolutionary Computation*, July 1999, pp.1569-1601.

[15] Custódio F, Barbosa H, Dardenne L. Investigation of the three-dimensional lattice HP protein folding model using a genetic algorithm. *Genetics and Molecular Biology*, 2004, 27(4): 611-615.

[16] Lopes H, Scapin M. An Enhanced genetic algorithm for protein structure prediction using the 2D hydrophobic-polar model. In *Proc. the 7th Int. Conf. Artificial Evolution*, October 2005, pp.238-246.

[17] Berenboym I, Avigal M. Genetic algorithms with local search optimization for protein structure prediction problem. In *Proc. the 10th Conference Genetic and Evolutionary Computation*, July 2008, pp.1097-1098.

[18] Cebrián M, Dotú I, Van Hentenryck P, Clote P. Protein structure prediction on the face centered cubic lattice by local search. In *Proc. the 23rd Conference on Artificial Intelligence*, July 2008, Vol.1, pp.241-246.

[19] Islam M, Chetty M. Novel memetic algorithm for protein structure prediction. In *Lecture Notes in Computer Science 5866*, Nicholson A, Li X (eds.), Springer Berlin/Heidelberg, 2009, pp.412-421.

[20] Garza-Fabre M, Rodriguez-Tello E, Toscano-Pulido G. Comparing alternative energy functions for the HP model of protein structure prediction. In *Proc. IEEE Congress on Evolutionary Computation*, June 2011, pp.2307-2314.

[21] Khimasia M, Coveney P. Protein structure prediction as a hard optimization problem: The genetic algorithm approach. *Molecular Simulation*, 1997, 19(4): 205-226.

[22] Cotta C. Protein structure prediction using evolutionary algorithms hybridized with backtracking. In *Lecture Notes in Computer Science 2687*, Mira J, Alvarez J (eds.), Springer Berlin/Heidelberg, 2003, pp.321-328.

[23] Bui T, Sundarraj G. An efficient genetic algorithm for predicting protein tertiary structures in the 2D HP model. In *Proc. Conf. Genetic and Evolutionary Computation*, June 2005, pp.385-392.

[24] Hoque M, Chetty M, Lewis A, Sattar A. Twin removal in genetic algorithms for protein structure prediction using low-resolution model. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2011, 8(1): 234-245.

[25] Krasnogor N, Blackburne B, Burke E, Hirst J. Multimeme algorithms for protein structure prediction. In *Proc. the 7th International Conference on Parallel Problem Solving from Nature*, September 2002, pp.769-778.

[26] Islam M, Chetty M. Clustered memetic algorithm for protein structure prediction. In *Proc. IEEE Congress on Evolutionary Computation*, July 2010, pp.1-8.

[27] Zhang X, Wang T, Luo H, Yang J, Deng Y, Tang J, Yang M. 3D protein structure prediction with genetic tabu search algorithm. *BMC Systems Biology*, 2010, 4(Suppl. 1): S6.

[28] Chira C. A hybrid evolutionary approach to protein structure prediction with lattice models. In *Proc. IEEE Congress on Evolutionary Computation*, June 2011, pp.2300-2306.

[29] Chira C, Horvath D, Dumitrescu D. Hill-climbing search and diversification within an evolutionary approach to protein structure prediction. *BioData Mining*, 2011, 4: Article No.8.

[30] Blazewicz J, Lukasiak P, Milostan M. Application of tabu search strategy for finding low energy structure of protein. *Artificial Intelligence in Medicine*, 2005, 35(1/2): 135-145.

[31] Pardalos P, Liu X, Xue G. Protein conformation of a lattice model using tabu search. *Journal of Global Optimization*, 1997, 11(1): 55-68.

[32] Shmygelska A, Hoos H. An ant colony optimisation algorithm for the 2D and 3D hydrophobic polar protein folding problem. *BMC Bioinformatics*, 2005, 6: Article No.30.

[33] Chu D, Till M, Zomaya A. Parallel ant colony optimization for 3D protein structure prediction using the HP lattice model. In *Proc. the 19th IEEE International Parallel and Distributed Processing Symposium*, April 2005, Vol.7, p.193b.

[34] Hu X M, Zhang J, Li Y. Flexible protein folding by ant colony optimization. In *Studies in Computational Intelligence 151*, Smolinski T G, Milanova M G, Hassgnien A (eds.), Springer, 2008, pp.317-336.

[35] Guo H, Lv Q, Wu J, Huang X, Qian P. Solving 2D HP protein folding problem by parallel ant colonies. In *Proc. the 2nd International Conference on Biomedical Engineering and Informatics*, October 2009, pp.1-5.

[36] De Almeida C, Gonçalves R, Delgado M. A hybrid immune-based system for the protein folding problem. In *Proc. the 7th Evolutionary Computation in Combinatorial Optimization*, April 2007, pp.13-24.

[37] Cutello V, Morelli G, Nicosia G, Pavone M. Immune algorithms with aging operators for the string folding problem and the protein folding problem. In *Proc. the 5th European Conf. Evolutionary Computation in Combinatorial Optimization*, April 2005, pp.80-90.

[38] Cutello V, Nicosia G, Pavone M *et al.* An immune algorithm for protein structure prediction on lattice models. *IEEE Trans. Evolutionary Computation*, 2007, 11(1): 101-117.

[39] Cutello V, Morelli G, Nicosia G, Pavone M, Scollo G. On discrete models and immunological algorithms for protein structure prediction. *Natural Computing*, 2011, 10(1): 91-102.

[40] Kanj F, Mansour N, Khachfe H, Abu-Khzam F. Protein structure prediction in the 3D HP model. In *Proc. IEEE/ACS International Conference on Computer Systems and Applications*, May 2009, pp.732-736.

[41] Băutu A, Luchian H. Protein structure prediction in lattice models with particle swarm optimization. In *Proc. the 7th Int. Conf. Swarm Intelligence*, Sept. 2010, pp.512-519.

[42] Bitello R, Lopes H. A differential evolution approach for protein folding. In *Proc. IEEE Symp. Computational Intelligence, Bioinformatics and Computational Biology*, Sept. 2006, pp.1-5.

[43] Lopes H, Bitello R. A differential evolution approach for protein folding using a lattice model. *Journal of Computer Science and Technology*, 2007, 22(6): 904-908.

[44] Santos J, Diéguez M. Differential evolution for protein structure prediction using the HP model. In *Lecture Notes in Computer Science 6686*, Ferrández J M, Sánchez J R Á, Paz F (eds.), 2011, pp.323-333.

[45] Jana N, Sil J. Protein structure prediction in 2D HP lattice model using differential evolutionary algorithm. In *Proc. the Int. Conf. Information Systems Design and Intelligent Applications*, January 2012, pp.281-290.

[46] Santana R, Larranaga P, Lozano J A. Protein folding in simplified models with estimation of distribution algorithms. *IEEE Trans. Evolutionary Computation*, 2008, 12(4): 418-438.

[47] Chen B, Li L, Hu J. A novel EDAs based method for hp model protein folding. In *Proc. IEEE Congress on Evolutionary Computation*, May 2009, pp.309-315.

[48] Unger R, Moult J. Genetic algorithms for protein folding simulations. *Journal of Molecular Biology*, 1993, 231(1): 75-81.

[49] Patton A, Punch III W, Goodman E. A standard GA approach to native protein conformation prediction. In *Proc. the 6th Int. Conf. Genetic Algorithms*, July 1995, pp.574-581.

[50] Lopes H, Scapin M. A hybrid genetic algorithm for the protein folding problem using the 2D-HP lattice model. In *Studies in Computational Intelligence 92*, Yang A, Shan Y, Bui L (eds.), Springer, 2008, pp.121-140.

[51] Dotú I, Cebrián M, Van Hentenryck P, Clote P. On lattice protein structure prediction revisited. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2011, 8(6): 1620-1632.

[52] Islam M, Chetty M, Murshed M. Novel local improvement techniques in clustered memetic algorithm for protein structure prediction. In *Proc. IEEE Congress on Evolutionary Computation*, June 2011, pp.1003-1011.

[53] Zhang J, Kou S C, Liu J S. Biopolymer structure simulation and optimization via fragment regrowth Monte Carlo. *The Journal of Chemical Physics*, 2007, 126(22): 225101.

[54] Thachuk C, Shmygelska A, Hoos H. A replica exchange Monte Carlo algorithm for protein folding in the HP model. *BMC Bioinformatics*, 2007, 8: 342.

[55] Wüst T, Li Y W, Landau D P. Unraveling the beautiful complexity of simple lattice model polymers and proteins using Wang-Landau sampling. *Journal of Statistical Physics*, 2011, 144(3): 638-651.

[56] Corne D, Knowles J. Techniques for highly multiobjective optimisation: Some nondominated points are better than others. In *Proc. the 9th Conf. Genetic and Evolutionary Computation*, Volume 1, July 2007, pp.773-780.

[57] Blum C, Roli A. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys*, 2003, 35(3): 268-308.

[58] Talbi E. Metaheuristics: From Design to Implementation. Wiley Publishing, 2009.

[59] Martin O, Otto S, Felten E. Large-step Markov chains for the traveling salesman problem. *Complex Systems*, 1991, 5(3): 299-326.

[60] Lourenço H, Martin O, Stützle T. Iterated local search. In *International Series in Operations Research & Management Science 57*, Glover F, Kochenberger G (eds.), Kluwer Academic Publishers, 2002, pp.321-353.

[61] Lourenço H, Martin O, Stützle T. Iterated local search: Framework and applications. In *International Series in Operations Research & Management Science 146*, Gendreau M, Potvin J Y (eds.), Springer, 2010, pp.363-397.

[62] Vanneschi L, Tomassini M, Collard P, Vérel S, Pirola Y, Mauri G. A comprehensive view of fitness landscapes with neutrality and fitness clouds. In *Proc. the 10th European Conf. Genetic Programming*, April 2007, pp.241-250.

[63] Collard P, Clergue M, Defoin-Platel M. Synthetic neutrality for artificial evolution. In *Proc. the 4th European Conf. Artificial Evolution*, November 1999, pp.254-265.

[64] Yu T, Miller J. Finding needles in haystacks is not hard with neutrality. In *Proc. the 5th European Conf. Genetic Programming*, April 2002, pp.13-25.

[65] Vérel S, Collard P, Clergue M. Scuba search: When selection meets innovation. In *Proc. IEEE Congress on Evolutionary Computation*, June 2004, pp.924-931.

[66] Marmion M, Dhaenens C, Jourdan L, Liefooghe A, Vérel S. On the neutrality of flowshop scheduling fitness landscapes. In *Lecture Notes in Computer Science 6683*, Coello C A C (ed.), Springer, 2011, pp.238-252.

[67] Marmion M, Dhaenens C, Jourdan L, Liefooghe A, Vérel S. NILS: A neutrality-based iterated local search and its application to flowshop scheduling. In *Proc. Evolutionary Computation in Combinatorial Optimization*, April 2011, pp.191-202.

[68] Marmion M, Dhaenens C, Jourdan L, Liefooghe A, Vérel S. The road to VEGAS: Guiding the search over neutral networks. In *Proc. Genetic and Evolutionary Computation Conference*, July 2011, pp.1979-1986.

[69] Watson J. An introduction to fitness landscape analysis and cost models for local search. In *International Series in Operations Research & Management Science 146*, Gendreau M, Potvin J Y (eds.), Springer, 2010, pp.599-623.

**Mario Garza-Fabre** received the M.Sc. degree in computer science from the Information Technology Laboratory at the Center for Research and Advanced Studies of the National Polytechnic Institute, CINVESTAV-Tamaulipas, México, in 2009. He is currently working towards the Ph.D. degree at the same institution. His current research interests include evolutionary computation and metaheuristics, multi- and many-objective optimization, bioinformatics and protein structure prediction.

**Eduardo Rodriguez-Tello** is an associate professor/researcher at the Information Technology Laboratory at CINVESTAV-Tamaulipas, México, since 2008. He received the M.Sc. degree in computer science from the Tecnológico de Monterrey, México, and the Ph.D. degree in informatics from the University of Angers, France, in 1999 and 2007, respectively. Dr. Rodriguez-Tello's research focuses on the design and implementation of effective metaheuristic algorithms for solving large-scale combinatorial optimization problems in various application areas. His current research interests in bioinformatics include phylogenetic tree reconstruction, gene regulatory network model inference, and protein structure prediction.

**Gregorio Toscano-Pulido** received the Ph.D. degree from the Center for Research and Advanced Studies of the National Polytechnic Institute (CINVESTAV), México city, México, in 2005. He is currently a full-time assistant professor with the Information Technology Laboratory, CINVESTAV-Tamulipas, México. He maintains an active interest in the design of multi-objective evolutionary algorithms. His further research interest includes many-objective optimization, dynamic multi-objective optimization, and bioinformatics.