



Article

# Reoptimized UNRES Potential for Protein Model Quality Assessment

Eshel Faraggi<sup>1,2,3</sup> , Pawel Krupa<sup>3,4</sup> , Magdalena A. Mozolewska<sup>3,5</sup>, Adam Liwo<sup>6,7</sup> and Andrzej Kloczkowski<sup>3,8,9,10,\*</sup><sup>1</sup> Research and Information Systems, LLC, Indianapolis, IN 46240, USA; efaraggi@iupui.edu<sup>2</sup> Department of Physics, Indiana University Purdue University Indianapolis, Indianapolis, IN 46202, USA<sup>3</sup> Battelle Center for Mathematical Medicine, Nationwide Children's Hospital, Columbus, OH 43215, USA; pkrupa@ifpan.edu.pl (P.K.); magdalena.mozolewska@gmail.com (M.A.M.)<sup>4</sup> Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, PL-02-668 Warsaw, Poland<sup>5</sup> Institute of Computer Science, Polish Academy of Sciences, ul. Jana Kazimierza 5, 01-248 Warszawa, Poland<sup>6</sup> Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland; adam@sun1.chem.univ.gda.pl<sup>7</sup> Center for In Silico Protein Structure and School of Computational Sciences, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, Seoul 130-722, Korea<sup>8</sup> Department of Pediatrics, The Ohio State University, Columbus, OH 43215, USA<sup>9</sup> Kavli Institute for Theoretical Physics China, Chinese Academy of Sciences, Beijing 100190, China<sup>10</sup> Future Value Creation Research Center, Graduate School of Informatics, Nagoya University, Nagoya 464-8601, Japan

\* Correspondence: andrzej.kloczkowski@nationwidechildrens.org; Tel.: +1-614-355-5622

Received: 15 October 2018; Accepted: 27 November 2018; Published: 3 December 2018



**Abstract:** Ranking protein structure models is an elusive problem in bioinformatics. These models are evaluated on both the degree of similarity to the native structure and the folding pathway. Here, we simulated the use of the coarse-grained UNited RESidue (UNRES) force field as a tool to choose the best protein structure models for a given protein sequence among a pool of candidate models, using server data from the CASP11 experiment. Because the original UNRES was optimized for Molecular Dynamics simulations, we reoptimized UNRES using a deep feed-forward neural network, and we show that introducing additional descriptive features can produce better results. Overall, we found that the reoptimized UNRES performs better in selecting the best structures and tracking protein unwinding from its native state. We also found a relatively poor correlation between UNRES values and the model's Template Modeling Score (TMS). This is remedied by reoptimization. We discuss some cases where our reoptimization procedure is useful.

**Keywords:** protein energy; optimization; model ranking; UNRES; Seder; GENN; OUNRES

## 1. Introduction

The problem of evaluating protein energy and scoring protein conformations has been an important aspect of protein research [1–24]. The energy and scoring functions serve to both guide protein simulation studies and to rank putative protein models. There are two main approaches used, categorized as *physical* and *knowledge-based* [25]. In the physical approach, an energy function is built based on a physical model of the atomic interactions and is then optimized based on experimental results. In the knowledge-based approach, the model itself relies on experimental results, typically by matching experimental distributions.

Selecting the best models among putative models is an important application of a protein energy or scoring function [2–10,12–16,18–24]. In such cases, models tailored to a specific sequence are

produced, and the task is to rank them according to a specified criterion, usually a measure of the deviation from the native structure corresponding to the given sequence. In this respect, there is a strong overall match between the energy values and the spatial deviation from native scores, such as the Template Modeling Score (TMS) [26]. However, this correspondence is not exact. While energies such as the coarse-grained UNited RESidue (UNRES) force field account for charge distributions, the TMS and similar measures do not. One can imagine transient states arising in cases where the charge distribution has a different transition time than the timescale associated with a structural change. However, protein native and decoy structures are steady states, i.e., they are allowed enough time to relax and escape any unfavorable transient states. Therefore, unstable configurations with lower TMSs but large energies are expected to be excluded. Taking ensemble averages would tend to decrease this effect further.

The current UNRES was optimized to carry out free simulations and not to score decoys [27–29]. An attempt at threading was made with a very early version of UNRES [9]; however, even this application involved decoy energy minimization. Given the success of UNRES in free simulations, we considered it worth trying this force field in decoy scoring. It should be noted that free simulations imply that the computed structures are relaxed; if not at all configurations, then at least the end ones. Decoy structures are fixed, and, therefore, clashes from side-chain–side-chain interactions can appear, in general. Consequently, to better design UNRES for decoy scoring, the long-range repulsive components of the potentials need to be better regulated. Thus, a new optimization of UNRES could improve UNRES for decoy scoring. To optimize UNRES for this purpose, we applied a methodology based on neural networks, developed in our earlier work [18,30].

## 2. Methods

### 2.1. UNited RESidue (UNRES)

UNRES [31–34] is a coarse-grained model for proteins in which each amino-acid residue is reduced to two interaction sites: the united peptide group (p), located halfway between two consecutive C<sup>α</sup> atoms (which are not interaction sites and are only used to define the geometry of the chain), and the united side chain (SC) (Figure 1). Due to the model's reduced number of interaction sites and the exclusion of averaging out the degrees of freedom (the secondary degrees of freedom), the UNRES force field provides a speed of at least 3 order of magnitude higher compared with the all-atom simulations [35].

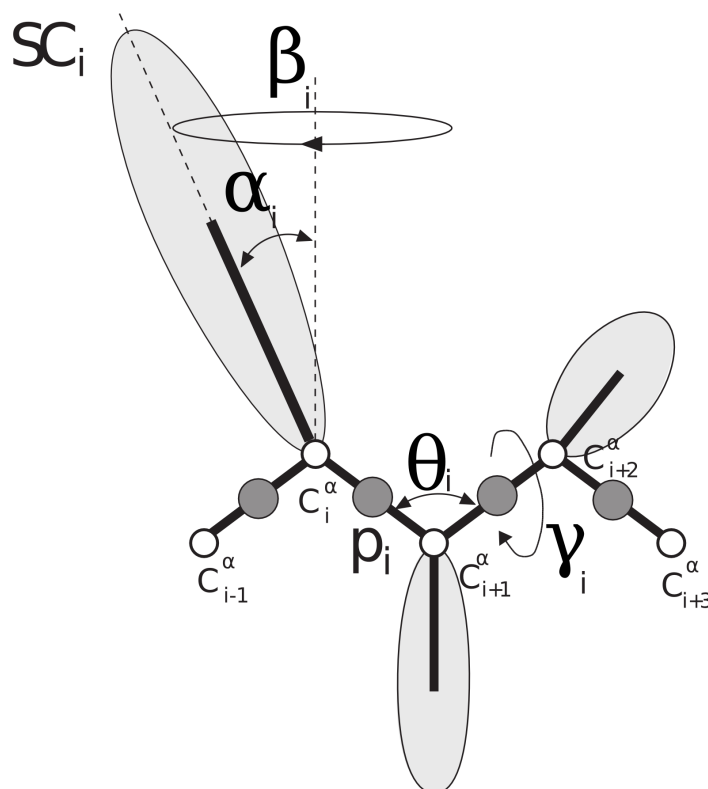
The effective energy function in the UNRES model is defined as the restricted free energy (RFE) or the potential of mean force (PMF), and it is given by Equation (1). A detailed description of UNRES is provided elsewhere [36].

$$\begin{aligned}
 U = & w_{SC} \sum_{i < j} U_{SC_i SC_j} + w_{SC_p} \sum_{i \neq j} U_{SC_i p_j} + w_{pp}^{VDW} \sum_{i < j-1} U_{p_i p_j}^{VDW} + w_{pp}^{el} f_2(T) \sum_{i < j-1} U_{p_i p_j}^{el} + \\
 & w_{tor} f_2(T) \sum_{i < j-1} U_{tor}(\gamma_i) + w_{tord} f_3(T) \sum_i U_{tord}(\gamma_i, \gamma_{i+1}) + w_b \sum_i U_b(\theta_i) + \\
 & w_{rot} \sum_i U_{rot}(\alpha_{SC_i}, \beta_{SC_i}) + w_{bond} \sum_i U_{bond}(d_i) + w_{corr}^{(3)} f_3(T) U_{corr}^{(3)} + \\
 & w_{corr}^{(4)} f_4(T) U_{corr}^{(4)} + w_{turn}^{(3)} f_3(T) U_{turn}^{(3)} + w_{turn}^{(4)} f_4(T) U_{turn}^{(4)} + \\
 & w_{ssbond} \sum_{nss} U_{ssbond}(d_{ss}) + w_{SC-corr} f_2(T) \sum_{m=1}^3 \sum_i U_{SC-corr}(\tau_i^{(m)})
 \end{aligned} \quad (1)$$

with

$$f_n(T) = \frac{\ln [\exp(1) + \exp(-1)]}{\ln \left\{ \exp \left[ (T/T_0)^{n-1} \right] + \exp \left[ - (T/T_0)^{n-1} \right] \right\}} \quad (2)$$

where each potential is multiplied by an appropriate weight,  $w.$ , and these weights are optimized. In Equation (1),  $U_{SC,SC}$  and  $U_{SC,p_j}$  are side-chain–side-chain and side-chain–peptide-group interaction potentials, respectively. The peptide-group–peptide-group interaction potential is split into the Lennard-Jones term ( $U_{p_i p_j}^{VDW}$ ) and the electrostatic term ( $U_{p_i p_j}^{el}$ ). The local properties of the polypeptide chain are described by  $U_{tor}$ ,  $U_{tor d}$ ,  $U_b$ ,  $U_{rot}$ , and  $U_{bond}$  potentials, which are torsional, double torsional, bending, rotameric, and virtual-bond-deformation terms, respectively.  $U_{corr}$  and  $U_{turn}$  are higher-order correlation terms that are necessary for the correct reproduction of secondary-structure elements [37],  $U_{ssbond}$  is a disulfide bond potential, and  $U_{SC-corr}$  is a recently implemented potential that couples the local positions of the backbone and side chains, which improves the predictive capacities of the UNRES force field [38,39]. Additionally, because the UNRES energy function originates from the PMF of polypeptide chains in water, in which the fine-grained degrees of freedom have been averaged out, it is temperature-dependent. The factors  $f_i$  arise from multiplying the terms of the respective order in the cluster-cumulant expansion of the PMF [37]. Because the current implementation of UNRES involves scoring the decoys, which correspond to folded structures and not folding simulations, we set the temperature at  $T = 300$  K, assuming that all proteins considered are folded at this temperature. The UNRES model uses an anisotropic potential for the interactions between side chains, which are represented by the Gay–Berne model [40]. This model allows for a more accurate approximation of the side-chain interactions than simpler spherical models.



**Figure 1.** Schematic representation of the UNited RESidue (UNRES) model of the polypeptide chain. There are two interaction sites per residue: united side-chain (SC) and united peptide group (p) are represented by light-gray ellipses and dark-gray circles, respectively.  $C^\alpha$  atoms (white circles) and the angles  $\beta$ ,  $\alpha$ ,  $\Theta$ , and  $\gamma$  define the positions of the backbone and side chains.

The energy-term weights in the initial version of the UNRES force field were optimized using only one  $\alpha$ -helical protein (PDB code: 1GAB) [27]. We shall term this version of UNRES ‘GB’. In later versions, the force field was re-parameterized using two training mini-proteins: the  $\alpha$ -helical tryptophan cage (PDB code: 1L2Y) and tryptophan zipper (PDB code: 1LE1) [28]. The latter force field was recently extended by the addition of the local torsional potentials [39] with very limited manual optimization of the weights of the torsional terms (Table 1). We term this version of UNRES ‘EL’. In the current force field, all the energy terms are physics-based except for side-chain–side-chain interaction terms, which were obtained by an analysis of the PDB [41]. Recently, a new approach to efficient force field optimization was developed [29] based on the maximum likelihood method [42]. However, even with the use of this method, only a very limited number of training proteins can be included in the optimization due to the high computational cost of the iterative procedure based on extensive folding simulations.

**Table 1.** Optimized weights for the EL variant of UNRES.

Weight Type	Value
$w_{SC}$	1.00000
$w_{SC_p}$	1.23315
$w_{pp}^{VDW}$	0.23173
$w_{pp}^{el}$	0.84476
$w_{tor}$	1.34316
$w_{tord}$	1.26571
$w_b$	0.62954
$w_{rot}$	0.10554
$w_{bond}$	1.00000
$w_{corr}^{(3)}$	0.37357
$w_{corr}^{(4)}$	0.19212
$w_{turn}^{(3)}$	1.40323
$w_{turn}^{(4)}$	0.64673
$w_{ssbond}$	1.00000
$w_{SC-corr}$	0.25000

The standard energy-term weights used in the UNRES energy function. These optimized weights correspond to the EL variant of UNRES. Abbreviations: Side Chain (SC), Van Der Waals (VDW), electrostatic (el), torsion (tor), toroidal (tord), bending (b), rotation (rot), bonding (bond), correlation (corr), turns (turn), and secondary structure bond (ssbond).

For the fold-recognition application reported in this paper, the side-chain–side-chain interaction ( $U_{SC_iSC_j}$ ), torsional ( $U_{tor}$ ), and side-chain-correlation ( $U_{SC-corr}$ ) terms of Equation (1) are the most important because they account for sequence-specific long- and short-range interactions. Therefore, in addition to the common weights for these terms, we introduced residue-pair-type-specific weights (a total of 400 for each of the three kinds of potentials). Residue-type-specific weights of the excluded-volume contributions ( $U_{SC_ip_j}$ ) were introduced, because these potentials control the size of the proteins and depend on a single residue type. Likewise, residue-type-specific weights were introduced for the contributions to the virtual-bond-angle ( $U_b$ ), side-chain-rotamer ( $U_{rot}$ ), and double-torsional ( $U_{tord}$ ) potentials, the type being that of the central residue. Because, for the decoys taken from the PDB database [41], the regular secondary structure is already present, the electrostatic ( $U_{p_i p_j}^{el}$ ) and correlation ( $U_{corr}^{(m)}$  and  $U_{turn}^{(m)}$ ) terms, which determine the regular secondary structure in free simulations, matter only as much as they contribute to the energy of the “bulk” of the secondary structure of different types. Therefore, only the weights corresponding to total  $U_{p_i p_j}^{el}$ ,  $U_{p_i p_j}^{VDW}$ ,  $U_{corr}^3$ ,  $U_{corr}^4$ ,  $U_{turn}^3$ , and  $U_{turn}^4$  were optimized (one weight per each kind of term).

Multiple types of calculations can be performed with UNRES, from single-point energy calculations and energy minimization in internal and external coordinates to Monte Carlo and Molecular Dynamics calculations of various variants and modifications. Including serial (sequential)

and parallel runs with scaling up to 70% on 16K cores [43]. Example of UNRES usage include: Conformational Space Annealing (CSA) [44], Hybrid Monte Carlo (HMC) [45], Replica Exchange Molecular Dynamics (REMD) [46], and Multiplexed Replica Exchange Molecular Dynamics (MREMD) [47]. UNRES has been successfully used for studies of protein folding pathways, thermodynamics, and kinetics [48–50]; in studies of multimeric systems [51,52] with the use of periodic boundary conditions [53]; and in systems with nonstandard amino acids [54] and links [55]. More detail on UNRES can be found elsewhere [36].

## 2.2. Reoptimization

As stated in the introduction, the current version of UNRES was optimized to run free simulations in which the potential clashes are removed; however, in general, this is not the case for scoring fixed decoys. Therefore, first, we modified the potential to limit its repulsive components. Specifically, we imposed a cutoff on the repulsive parts of the potential to limit the maximum repulsion to 3 kcal/mol for a given interaction type between each pair of interaction centers. Only with such an approach can the UNRES force field be used without the prior energy minimization of a system because, otherwise, even slight overlapping of the interaction centers can outweigh all other energy components. Another possibility would be to use soft potentials, such as the 8–6 Lennard-Jones potential, but, even then, short energy minimization is needed [56]. We then reoptimized the force field to accommodate this change and to customize it to the task of decoy scoring.

We employed the neural network technique for optimization. In the implementation of the UNRES model developed in this work, although energy is a linear function of the parameters, the error function to minimize is expressed as:

$$F = \sum_{p=1}^{totp} \sum_{i=0}^{N_p-1} \left[ TMS_p^{pred}(\{U\}_{pi}) - TMS_{pi}^{decoy} \right]^2 \quad (3)$$

where index  $p$  runs over the training proteins, index  $i$  runs over the decoys corresponding to a given training protein (including the native structure, which has an index of 0),  $totp$  is the total number of training proteins,  $N_p$  is the total number of decoys for protein  $p$  (including the native structure of this protein),  $TMS^{pred}$  and  $TMS^{decoy}$  denote the predicted TM-scores and those calculated from the respective decoy and native structures, and the input features are  $\{U\}_{pi}$ , the set of UNRES energy components (Equation 1) calculated for decoy  $i$  of protein  $p$ . As described later, other input features will be used alongside  $\{U\}_{pi}$ . In this work, we used the back-propagation neural network method to approximate the values of  $TMS_p^{pred}$  that will minimize Equation (3). The neural network we used is a nonlinear function from the input features ( $\{U\}_{pi}$ ) to the output feature ( $TMS^{pred}$ ). We started with a random set of neural network weights and passed the input features through the neural network weights to calculate  $TMS^{pred}$ . This part of the process is known as feed-forward. We then calculated the error associated with said prediction and, by the steepest descent method, modified the neural network weights to reduce this error. This process is known as back-propagation. It was carried out repeatedly until an overfit-protection test was violated. In this case, for the overfit protection, we left out part of the data from training and chose weights that gave the best results for this left-out set.

From UNRES, we first extracted information characterizing the state of the protein as an initial step toward UNRES reoptimization. Besides giving the overall UNRES energy value and the values for each of the components in Equation (1), we also split these components into their residue-type-specific contributions. For example, for paired interactions, we calculated separate values for the contributions from interactions between one type of residue and another. In total, we have the following input features from UNRES: one overall energy, nine single-value components, four 20-valued components that are residue-type-dependent, and three 400-value components that are dependent on the type of a residue pair. See Table 2 for a list and description of these labels. Additionally, the weights of each kind of UNRES energy component and that of the total UNRES energy were also optimized. In

total, we have  $1 + 15 + 4 \times 20 + 3 \times 400 = 1296$  values characterizing the UNRES energy function in this study. It should be noted that the parameters of the pairwise side-chain–side-chain interaction energies are not symmetric. The reason for this is the directionality of the protein chain (from the N to the C terminus). For example, a parameter for an ‘AC’ pair of side chains means that alanine precedes cysteine in sequence. Such non-symmetry of interactions is quite commonly used in fold-recognition studies [1–6] and partially accounts for the “through sequence” long-range interactions.

**Table 2.** Labels for UNRES characterization.

Label	Size	Description
$U_{SC_iSC_j}$	400	Side-chain hydrophobic (hydrophilic) interaction
$U_{SC_iP_j}$	20	Excluded-volume side-chain–peptide-group interactions
$U_{P_iP_j}^{vdW}$	1	Lennard-Jones interaction energy between peptide-group centers
$U_{P_iP_j}^{el}$	1	Electrostatic energy between peptide-group dipoles
$U_{tor}$	400	Virtual-bond dihedral angle torsional terms
$U_{tor d}$	20	Virtual-bond dihedral angle double-torsional terms
$U_b$	20	Virtual-bond angle bending terms
$U_{rot}$	20	Side-chain rotamer term
$U_{bond}$	1	Virtual-bond-deformation term
$U_{corr}^{(3)}$	1	Third-order correlation term
$U_{corr}^{(4)}$	1	Multibody coupling backbone-local and backbone electrostatic interactions
$U_{turn}^{(3)}$	1	Correlation contributions involving three consecutive peptide groups
$U_{turn}^{(4)}$	1	Correlation contributions involving four consecutive peptide groups
$U_{ssbond}$	1	Residue-dependent side-chain rotamer terms
$U_{SC-crr}$	400	Physics-based side-chain backbone correlation potentials

Labels for the energy decomposition used as input for representing UNRES to the neural network and their description. Size refers to the number of components a given label has. Some, such as the overall energy, are single-valued (1), four depend on the residue type and are 20-valued (20), and three depend on the type of a residue pair and are 400-valued (400).

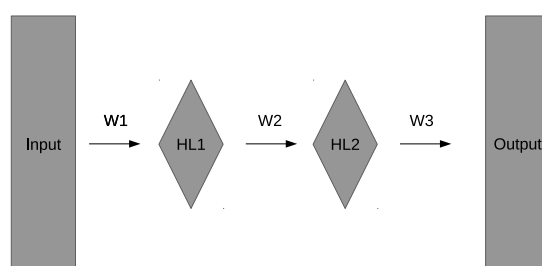
The same approach we used in the Seder1 scoring function [18] was used here for scoring a model of a given sequence based on its similarity to the native PDB [41] structure of the sequence, as measured by the TMS, which provides a normalized value for training our networks. In addition, we established a formula for transforming the TM-score,  $TMS' = 1 - 2 \times TMS$ , to achieve a distribution of values more fitting to our bipolar selection of neural networks and to align the directionality between our score and the energy values. With our transformation, a native structure scores a ‘−1’.

We used a two-layer feed-forward neural network with momentum, recently described in detail [30]. A diagram of the neural network architecture is given in Figure 2. HL1 and HL2 refer to the first and second hidden layers, respectively, and W1, W2, and W3 refer to the weights connecting the different layers. We used an all-connected network in which weights connect all the nodes of one layer to all the nodes of the next layer. The number of weights for a given network will depend on the number of inputs, as this will determine the number of weights in W1. For an all-connected network, the number of weights connecting layer L1 with  $h_1$  nodes (plus a bias node) to layer L2 with  $h_2$  nodes is  $(h_1 + 1) \times h_2$ . At each node, the weighted sum of the previous layer is passed through an activation function to give the value of that node. We used a bipolar hyperbolic tangent activation function. Momentum refers to the contribution of the gradient calculated at the previous time step to the correction of the weights.

We used the steepest-descent back-propagation algorithm to optimize the neural network weights. We started the analyses with a randomly selected training set of 22,805 protein chain models from the full training set of 296,381. The use of a small training set enabled us to optimize the architecture. After that initial optimization, deviations from the optimized values were tested for the full set. The final optimized values are given in Table 3. From the full training set, we selected 30% of the proteins at random for an overfit-protection set. A total of six such random training/overfit sets were



used to train different realizations of the neural network. For each of the approaches to optimizing UNRES in Table 4, the initial weights and the order of the training proteins were randomized for each of the six neural network realizations used for the corresponding approach. We chose to use six realizations based on experience from previous work [18,30,57,58]. To obtain the final prediction, we averaged the results from the six realizations. This also gives an estimate for the stability of the prediction through the standard deviation. We obtained the full training and overfit-protection sets from three sources (number of models given in parenthesis): server models submitted to CASP4 through CASP10 (123,634) [59], native models from the PDB (54,084) [41], and native models from the UCSF database of protein models (118,663) [60]. The three sources are treated in more detail in our previous publication [18]. Combining these three sources resulted in 296,381 proteins. From these, we randomly set aside 30% for the overfit-protection set and used the rest for training. We used the published results from CASP11 [59] as the testing set for the results here. This set contains 83 proteins that were selected by the CASP organizers to represent a variety of protein structures. Our approach here simulates participation in CASP11.



**Figure 2.** General architecture of the neural network. HL1 and HL2 refer to the first and second hidden layers, respectively, and W1, W2, and W3 refer to the weights connecting the different layers.

**Table 3.** Optimized parameters for neural networks.

Parameter	Value
<i>HL1</i>	6
<i>HL2</i>	7
<i>a</i>	0.2
$\mu$	0.003973
<i>P</i>	0.4
$N_{max}$	10,000
$N_{stop}$	400

*HL1* and *HL2* are the number of neurons in the first and second hidden layers, respectively. An additional bias neuron is used in each layer. *a*,  $\mu$ , and *P* are the activation parameter, the learning rate, and the momentum, respectively.  $N_{max}$  and  $N_{stop}$  are the maximum number of epochs and the number of epochs necessitating a stop after no improvement on the overfit-protection set, respectively.

Both the EL and GB versions of UNRES were optimized, resulting in the OUNRES (Optimized UNRES) versions. We also tried to integrate additional external information into UNRES, such as the number of residues, the number of atoms, and the scores from DFire2 [61] and Seder1 [18]. All input features were z-scored. We used cutoff values, defined by the values of the top and bottom 1% of the data, to limit the effect of outliers in the UNRES values.

Our testing method consisted of simulating participation in the CASP11 competition [59]. All of our training instances were restricted to models available before the release of CASP11 targets. We collected a total of 83 surviving CASP11 targets and the top 150 server predictions for them. This resulted in a total testing dataset of 12,240 structures. No information from these structures, such as overfit protection, parameter optimization, or any others, was used in the training of the neural network for the testing reported here. However, the version of OUNRES released with this work (and available from <http://mamiris.com/services.html>) used CASP11 information to select the top-performing networks among possible candidates.

### 3. Results

We analyzed the results of the optimization of the partial UNRES energies for both the GB and EL methods, with and without added input features. Protein structure scoring functions have several uses of interest. Among these functions is the ability to select the closest model to a native structure and, in the scope of protein folding, to select models along the folding pathway that approach a native state. Testing the first application of selecting the model closest to the native structure is relatively easy; testing the second use is considerably more challenging. We used the Pearson correlation with the real TMS and several other self-developed methods to estimate the effectiveness of selecting paths along the folding pathway; however, these results are more difficult to interpret since TMS does not account for charge distributions, as mentioned earlier.

We began by comparing the mean TMS of the top five selected models according to the various methods. We first ranked the models according to the prediction of a given method and then calculated the mean TMSs of the top five models for each of the 83 CASP11 targets. The mean and standard deviation (STD) of these 83 values were calculated for each method. The results of these calculations are given in Table 4.

**Table 4.** Top five mean Template Modeling Scores (TMSs).

EL			GB		
Method	Top5TMS	STD	Method	Top5TMS	STD
UNRES	0.596	0.074	UNRES	0.596	0.074
OUNRES	0.599	0.076	OUNRES	0.601	0.068
OUNRES+Seder1	0.600	0.072	OUNRES+Seder1	0.600	0.079
OUNRES+length	0.605	0.076	OUNRES+length	0.605	0.077
OUNRES+both	0.606	0.073	OUNRES+both	0.606	0.079
OU+both+DFire2	0.605	0.081	OU+both+DFire2	0.603	0.077
DFire2 alone	0.607	0.074	Seder1 alone	0.611	0.085

EL and GB UNRES variants tested relative to Optimized UNRES (OUNRES), OUNRES with Seder1 as neural network input (OUNRES+Seder1), OUNRES with residue length and number of atoms (OUNRES+length), OUNRES with both Seder1 and length (OUNRES+both), and OUNRES with Seder1, length, and DFire2.0 (OU+both+DFire2). Top5TMS is the mean of the TMSs of the top five models selected by each method. STD is the standard deviation and was calculated over the 83 CASP11 targets.

The results of the correlation were also calculated. Pearson correlations were calculated between the TMS to native structure of a model and its prediction according to the different methods. The correlation was also calculated between the TMS to native and the UNRES energy. This calculation was done per target. Then, the mean and STD of the correlations calculated for each of the 83 targets were obtained. These results are presented in Table 5. The Pearson coefficients are low, but it should be noted that the coefficients obtained for the OUNRES variants are higher than those for Seder1 and DFire2 alone; this means that OUNRES can rank the bulk of the decoys better than those two methods.

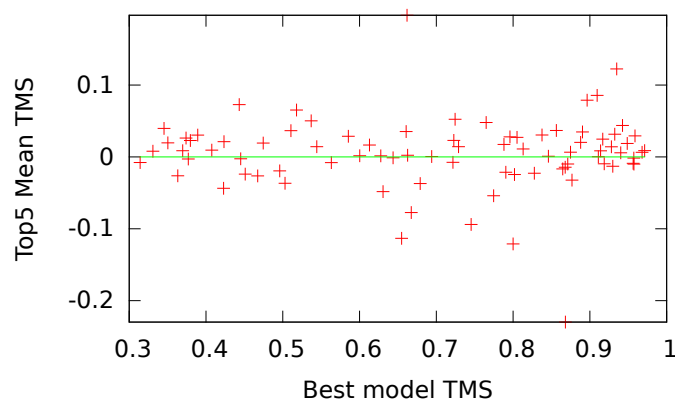


**Table 5.** Pearson Correlation.

EL			GB		
Method	Correlation	STD	Method	Correlation	STD
UNRES	0.151	0.189	UNRES	0.150	0.177
OUNRES	0.322	0.173	OUNRES	0.315	0.176
OUNRES+Seder1	0.320	0.182	OUNRES+Seder1	0.326	0.186
OUNRES+length	0.316	0.201	OUNRES+length	0.332	0.194
OUNRES+both	0.327	0.175	OUNRES+both	0.326	0.195
both+DFire2	0.320	0.186	both+DFire2	0.334	0.199
DFire2 alone	0.247	0.215	Seder1 alone	0.280	0.185

Refer to Table 4 for the legend. STD is the standard deviation and was calculated over the 83 CASP11 targets.

To better understand the result of our optimization, in Figure 3 we give the differences in the top five mean TMSs between OUNRES and UNRES as a function of the TMS to native of the best available decoy for that target (topTMS). In most cases (51/83), with both easy and hard targets, we find that OUNRES is an improvement over UNRES. In some cases, the optimization seems to reduce the quality of the top selected models, represented by a highly negative y-axis. We looked at the two worst cases: T0782 with a topTMS 0.85 and T0765 with a topTMS of 0.8. In both cases, OUNRES appears to perform significantly worse than UNRES, judging by the top five mean TMSs. If we observe the resulting protein structures in Figure 4, we see that in the case of T0782, UNRES seems to better model the beta barrel. However, in the case of T0765, it seems that OUNRES produces a better structure, while the increase in TMS for UNRES is mostly due to the structure being more compact. In Figure 5, we plot the change in correlation upon optimization as a function of the topTMS. We see that the correlation improves for most (68/83) targets. Additionally, it seems that only targets with a high topTMS (easy targets) are made worse by using OUNRES over UNRES. For hard targets, it seems that the correlation is always improving.



**Figure 3.** Difference in the top five mean TMSs between OUNRES and UNRES as a function of the TMS to native of the best available decoy for that target submitted during the CASP11 experiment. Smaller values on the x-axis indicate harder targets, while the y-axis measures the success of UNRES optimization, with higher values indicating greater success.

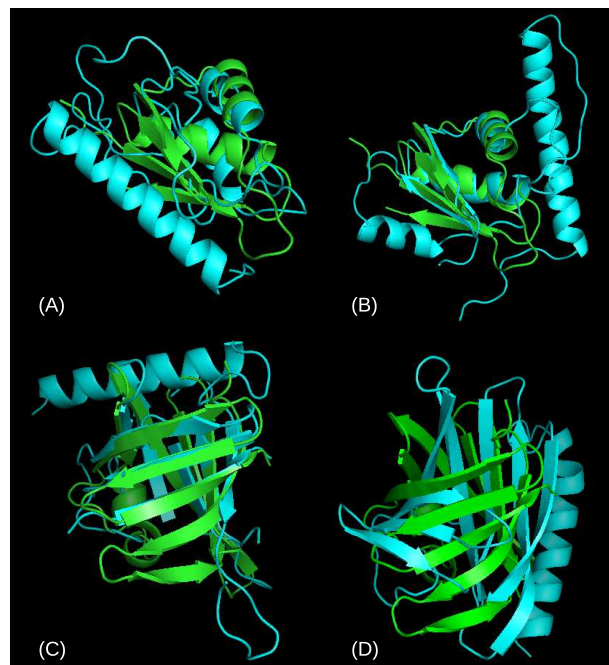
We also tested the directional accuracy of the different methods in two ways. First, we calculated the mean TMS for the top 1–5 (top1-5) and top 10–15 (top10-15) real model TMS to native ranking. We then calculated the average score/energy for the top1-5 and top10-15 sets according to the different methods and calculated the change. If the change was appropriate for a given score/energy, i.e., it points to the top1-5 being more favorable, we assigned a '+1' for this CASP11 target. If the score was inappropriate, we assigned a '-1' to this target. We then calculated the average assignment of the

83 targets and the STD. We term this parameter the *Directional Accuracy* (DA). Results for the DA are presented in Table 6.

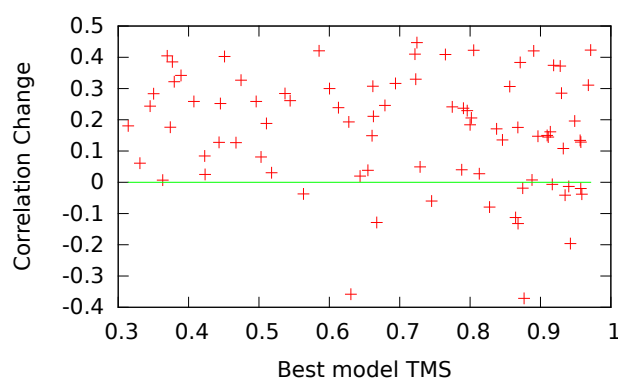
**Table 6.** Directional Accuracy (DA).

EL			GB		
Method	DA	STD	Method	DA	STD
UNRES	0.084	0.996	UNRES	0.084	0.996
OUNRES	0.181	0.984	OUNRES	0.181	0.984
OUNRES+Seder1	0.253	0.967	OUNRES+Seder1	0.229	0.973
OUNRES+length	0.253	0.967	OUNRES+length	0.157	0.988
OUNRES+both	0.181	0.984	OUNRES+both	0.205	0.979
both+DFire2	0.181	0.984	both+DFire2	0.205	0.979

Refer to Table 4 for the legend. STD is the standard deviation and was calculated over the 83 CASP11 targets. Note that it is artificially high because of the discrete nature of this test.



**Figure 4.** Structures for the top scored models for the two worst cases in Figure 3 using UNRES and OUNRES. Native structures are given in green, decoy model structures in cyan. (A) UNRES best for target T0765, model Alpha-Gelly-Server-TS5; (B) OUNRES best for target T0765, model RBO-Aleph-TS3; (C) UNRES best for target T0782, model SAM-T08-server-TS2; (D) OUNRES best for target T0782, model RBO-Aleph-TS3.



**Figure 5.** Difference in correlation between OUNRES and UNRES as a function of the TMS to native of the best available decoy for that CASP11 target. Smaller values on the x-axis indicate harder targets, while the y-axis measures the success of UNRES optimization, with higher values indicating greater success.

This test can also be done in the reverse order. The models can be ranked according to a score/energy, then the real TMS difference between the top1-5 and the top10-15 can be calculated. We calculated this for the 83 CASP11 targets we used and averaged the results to arrive at a single value per method employed. We also calculated the STD for this test. We term this parameter the *Second Directional Accuracy* (DA2). Results for DA2 are given in Table 7.

We also tested a path to native accuracy. We can imagine that the server models we collected from the CASP11 experiments are a folding pathway in the configuration space to the native structure. To obtain this pathway, we started with the native model for a given target sequence. We then found the nearest structure, as measured by the real TMS, and repeated the process. The used structures were excluded until all models for the target were exhausted. Following the consecutively closest structures, a folding pathway was obtained, for which energies and scores were calculated. In a similar fashion to that above, we assigned a '+1' if the change in energy or score was consistent with the direction of the path, i.e., decreasing or not changing as the native state is approached, and a '-1' otherwise. We then averaged these values along the path for a given CASP11 target and then averaged the resultant means to arrive at a single value per method. We call this the *Path to Native Accuracy* or PNA. These results are given in Table 8.

**Table 7.** Second Directional Accuracy (DA2).

EL		GB			
Method	DA2	STD	Method	DA2	STD
UNRES	-0.009	0.065	UNRES	-0.008	0.067
OUNRES	0.058	0.073	OUNRES	0.058	0.076
OUNRES+Seder1	0.064	0.070	OUNRES+Seder1	0.056	0.068
OUNRES+length	0.064	0.074	OUNRES+length	0.070	0.068
OUNRES+both	0.061	0.078	OUNRES+both	0.066	0.066
both+DFire2	0.061	0.083	both+DFire2	0.065	0.074

Refer to Table 4 for the legend. STD is the standard deviation and was calculated over the 83 CASP11 targets. Note that, in this case, for the optimized methods, the signal is almost entirely positive, which indicates that the methods were directionally successful for almost all protein targets (72/83 for OUNRES+length).

**Table 8.** Path to Native Accuracy (PNA).

EL			GB		
Method	PNA	STD	Method	PNA	STD
UNRES	0.017	0.051	UNRES	0.019	0.057
OUNRES	0.028	0.051	OUNRES	0.020	0.057
OUNRES+Seder1	0.018	0.054	OUNRES+Seder1	0.010	0.052
OUNRES+length	0.034	0.053	OUNRES+length	0.024	0.046
OUNRES+both	0.023	0.049	OUNRES+both	0.013	0.047
both+DFire2	0.026	0.055	both+DFire2	0.021	0.059

Refer to Table 4 for the legend. STD is the standard deviation and was calculated over the 83 CASP11 targets.

#### 4. Discussion

We see an overall consistent improvement with the optimization and the addition of input features across all tests undertaken. We did not find any significant advantage of the EL or GB approaches to UNRES over the other. For the mean TMS of the top1-5 models, we see a consistent improvement, with the optimization adding about a percent of relative accuracy and inclusion of the sequence length and number of atoms adding another relative percent to the accuracy. Improvements over UNRES of OUNRES+length (EL or GB) have a statistical confidence of more than 99% according to a two-sided Student's *t*-test. The addition of information from DFire2 does not seem to improve the accuracy for this case.

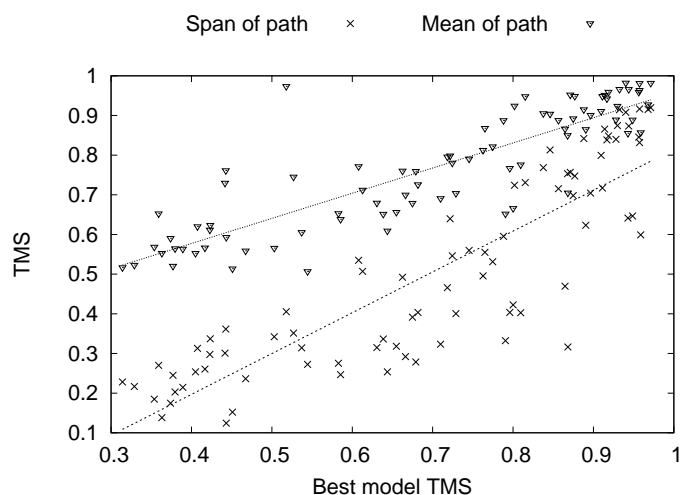
We also calculated the mean TMS for the top five ranked structures using only DFire2 or Seder1 to perform the ranking. In both cases, we get slightly better results than the neural networks trained on parameters extracted from UNRES, with and without Seder1 or DFire2 or both as inputs. This seems to be due to the optimization of the scoring for models farther from the native state than the top five. This can be seen in terms of correlations, where using either Seder1 or DFire2 yields worse correlations, with Seder1 outperforming DFire2 both for correlations and for top five mean TMS. These observations seem to indicate that although we seem to have improved UNRES by optimization, there is information yet to be picked up by the neural network for close to native structures.

A significant effect is observed for the Pearson correlation, where a 2-fold increase is observed upon the optimization of UNRES. Slight fluctuation around this improvement is observed with the introduction of additional input features; however, there is no significant improvement above a correlation of 0.32–0.33. As indicated earlier, the strong improvement upon optimization could be in part due to UNRES's consideration of charge distributions and the fact that the correlation is calculated over the entire sample of models for a given target.

For DA1 and DA2, we again see a strong response upon optimization and some additional response to the inclusion of additional input features. In both cases, we expect that the better the score/energy, the more consistently appropriate will be the change in values between the top1-5 to top10-15 models. More than a 2-fold increase in accuracy is observed upon optimization, and an additional significant improvement in accuracy is observed if additional input features are introduced. Improvement due to optimization in both cases has a confidence greater than 99% according to a two-sided Student's *t*-test. One should note that due to the choice of variables for test DA1 (1, −1) and its discrete nature, the STD in this case is exaggerated. Note that for DA2, for the optimized methods, the signal is almost entirely positive; i.e., the mean minus the STD is greater than zero. This indicates that the optimized methods were directionally successful for almost all protein targets. For the most successful method, OUNRES+length, 72 out of 83 targets had the correct directionality.

For the PNA, we do not see a strong signal. The fluctuations in the PNA are quite large, indicating that, for many targets, the directional assignments were rendered more erroneous. However, there is enough signal to observe a significant improvement from the optimization and additional input features. In this case, there is no clear consistency in the improvements though, and that could be due to the nature of the path to the native structure. It is interesting to note that the path to native is

intimately related to the hardness of the target. In Figure 6, we give the mean and span of the path as a function of the best model TMS submitted to CASP11. The 83 points in the plot correspond to the 83 CASP11 targets we used. The mean of the path is defined as the average over the TMS between consecutive models along the path. The span of the path is defined as the TMS between the native structure and the final model structure in the path until all others have been excluded. The Pearson correlation coefficients between the best model TMS and the mean/span of the path are 0.868/0.869, respectively. The correlation between the mean and the span of the path is 0.922. We also calculated the best fit lines between the best model TMS and the mean/span of the path. The line for the span of the path is given by  $l(x) = 1.03x - 0.21$ . The line for the mean of the path is given by  $l(x) = 0.63x + 0.32$ .



**Figure 6.** Mean (triangles) and span (x) of the path as a function of the best server model TMS submitted to CASP11. The Pearson correlation coefficients between the best model TMS and the mean/span of the path are 0.868/0.869, respectively. The correlation between the mean and the span of the path is 0.922. The best fit lines are presented to aid the eye. The line for the span of the path is given by  $l(x) = 1.03x - 0.21$ . The line for the mean of the path is given by  $l(x) = 0.63x + 0.32$ .

## 5. Conclusions

We reoptimized the UNRES energy function for protein decoy model quality assessment and achieved consistently better results in a number of tests. We find the bulk of the improvement for this round of optimization is from the improved scoring for the bulk of the models. This is seen by the large increase in the correlation of OUNRES relative to UNRES. This bias toward improving the bulk of the data results from the choice of neural network architecture and approach. It should be noted that this is the first attempt at using UNRES for scoring fixed decoy sets. A very early version of UNRES was used for threading [9], but in that work, the decoys were subjected to restrained energy minimization with UNRES, and minimized energies were used for scoring.

We introduced several quantities to help compare the energy/score functions of proteins. The Top5TMS measures the average TMS to native of the top five picked by the method. DA1 and DA2 provide a measure of the directional successes of the energy/score function in terms of the path to native. Finally, the PNA is a measure of the directional success of an energy/score function along the folding pathway of a protein.

We find that additional information in the form of additional input features tends to improve the accuracy of UNRES and OUNRES in picking the closest to native models and in assigning the direction toward a closer to native structure. In this respect, it seems that simply adding the z-scaled number of residues and atoms improves the performance of OUNRES most significantly. However, we find that DFire2 does not seem to improve the performance of UNRES, possibly due to an existing PMF in UNRES. On the other hand, the DA1, DA2, and PNA measures suggest that OUNRES has a substantial power of energy-ranking the bulk of the decoys and can, therefore, be used for selecting

the decoys for further processing, rather than picking prediction candidates. Thus, OUNRES seems to be of advantage when decoys need to be pre-selected for further processing, rather than for the selection of the final models.

**Author Contributions:** Conceptualization, E.F., A.L. and A.K.; Data curation, P.K. and M.A.M.; Formal analysis, E.F.; Funding acquisition, A.L. and A.K.; Investigation, E.F., P.K. and M.A.M.; Methodology, E.F.; Resources, P.K. and M.A.M.; Software, E.F., P.K. and M.A.M.; Supervision, E.F., A.L. and A.K.; Validation, E.F.; Visualization, E.F.; Writing—original draft, E.F.; Writing—review & editing, E.F., P.K., M.A.M., A.L. and A.K.

**Funding:** This work was supported in part by the National Institute of Health through Grant No. R01GMGM127701, by The Research Institute at Nationwide Children’s Hospital, by the Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute by the Indiana METACyt Initiative and by the National Science Foundation under Grant No. CNS-0521433 and CNS-0723054. Additional computational resources for the initial design of UNRES were provided by the PL-GRID infrastructure, the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdańsk, the Interdisciplinary Center of Mathematical and Computer Modeling (ICM) at the University of Warsaw, and the Beowulf cluster at the Faculty of Chemistry, University of Gdańsk. AL was supported in part by grant DEC-2013/10/M/ST4/00640 from the National Science Center of Poland (Narodowe Centrum Nauki). MM was supported in part by grant DEC-2015/17/N/ST4/03935 from the National Science Center of Poland (Narodowe Centrum Nauki), and PK was supported in part by grant DEC-2015/17/N/ST4/03937 from the National Science Center of Poland (Narodowe Centrum Nauki).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sippl, M.J.; Némethy, G.; Scheraga, H.A. Intermolecular potentials from crystal data. 6. Determination of empirical potentials for O-H...O-C hydrogen bonds from packing configurations. *J. Phys. Chem.* **1984**, *88*, 6231–6233.
2. Sippl, M.J. Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins. *J. Mol. Biol.* **1990**, *213*, 859–883. [[CrossRef](#)]
3. Casari, G.; Sippl, M.J. Structure-derived hydrophobic potential. Hydrophobic potential derived from X-ray structures of globular proteins is able to identify native folds. *J. Mol. Biol.* **1992**, *224*, 725–732. [[PubMed](#)]
4. Sippl, M.J. Boltzmann’s principle, knowledge-based mean fields and protein folding. An approach to the computational determination of protein structures. *J. Comput.-Aid. Mol. Des.* **1993**, *7*, 473–501. [[CrossRef](#)]
5. Böhm, H.J. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 243–256. [[CrossRef](#)]
6. Sippl, M.J. Knowledge-based potentials for proteins. *Curr. Opin. Struct. Biol.* **1995**, *5*, 229–235. [[CrossRef](#)]
7. Thomas, P.D.; Dill, K.A. Statistical potentials extracted from protein structures: How accurate are they? *J. Mol. Biol.* **1996**, *257*, 457–469. [[CrossRef](#)]
8. Ben-Naim, A. Statistical potentials extracted from protein structures: Are these meaningful potentials? *J. Chem. Phys.* **1997**, *107*, 3698–3706. [[CrossRef](#)]
9. Liwo, A.; Pincus, M.R.; Wawak, R.J.; Rackovsky, S.; Ołdziej, S.; Scheraga, H.A. A United-Residue Force Field for Off-Lattice Protein-Structure Simulations. II: Parameterization of Local Interactions and Determination of the Weights of Energy Terms by Z-score Optimization. *J. Comput. Chem.* **1997**, *18*, 874–887. [[CrossRef](#)]
10. Koppensteiner, W.; Sippl, M.J. Knowledge-based potentials—back to the roots. *Biochem. Biokhimiia* **1998**, *63*, 247.
11. Liwo, A.; Lee, J.; Ripoll, D.R.; Pillardy, J.; Scheraga, H.A. Protein Structure Prediction by Global Optimization of a Potential Energy Function. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5482–5485. [[CrossRef](#)] [[PubMed](#)]
12. Samudrala, R.; Levitt, M. Decoys ‘R’ Us: A database of incorrect conformations to improve protein structure prediction. *Protein Sci.* **2000**, *9*, 1399–1401.
13. Bonneau, R.; Baker, D. Ab initio protein structure prediction: Progress and prospects. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 173–189. [[CrossRef](#)] [[PubMed](#)]
14. Summa, C.M.; Levitt, M.; DeGrado, W.F. An atomic environment potential for use in protein structure prediction. *J. Mol. Biol.* **2005**, *352*, 986–1001. [[CrossRef](#)]



15. Pierce, B.; Weng, Z. ZRANK: Reranking protein docking predictions with an optimized energy function. *Proteins Struct. Funct. Bioinform.* **2007**, *67*, 1078–1086. [[CrossRef](#)] [[PubMed](#)]
16. Benkert, P.; Biasini, M.; Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **2011**, *27*, 343–350. [[CrossRef](#)] [[PubMed](#)]
17. Rodrigues, J.P.G.L.M.; Levitt, M.; Chopra, G. KoBaMIN: A knowledge-based minimization web server for protein structure refinement. *Nucleic Acids Res.* **2012**, *40*, W323–W328. [[CrossRef](#)] [[PubMed](#)]
18. Faraggi, E.; Kloczkowski, A. A global machine learning based scoring function for protein structure prediction. *Proteins Struct. Funct. Bioinform.* **2014**, *82*, 752–759. [[CrossRef](#)]
19. McGuffin, L.J.; Buenavista, M.T.; Roche, D.B. The ModFOLD4 server for the quality assessment of 3D protein models. *Nucleic Acids Res.* **2013**, *41*, W368–W372. [[CrossRef](#)]
20. McGuffin, L.J.; Atkins, J.D.; Salehe, B.R.; Shuid, A.N.; Roche, D.B. IntFOLD: An integrated server for modelling protein structures and functions from amino acid sequences. *Nucleic Acids Res.* **2015**, *43*, W169–W173. [[CrossRef](#)]
21. Maghrabi, A.H.A.; McGuffin, L.J. ModFOLD6: An accurate web server for the global and local quality estimation of 3D protein models. *Nucleic Acids Res* **2017**, *45*, W416–W421. [[PubMed](#)]
22. Yang, J.; Zhang, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* **2015**, *43*, W174–W181. [[CrossRef](#)] [[PubMed](#)]
23. Wang, Y.; Virtanen, J.; Xue, Z.; Zhang, Y. I-TASSER-MR: Automated molecular replacement for distant-homology proteins using iterative fragment assembly and progressive sequence truncation. *Nucleic Acids Res.* **2017**, *45*, W429–W434. [[CrossRef](#)] [[PubMed](#)]
24. Buchan, D.W.A.; Jones, D.T. EigenTHREADER: Analogous protein fold recognition by efficient contact map threading. *Bioinformatics* **2017**, *33*, 2684–2690. [[CrossRef](#)] [[PubMed](#)]
25. Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A.E.; Kolinski, A. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.* **2016**, *116*, 7898–7936. [[CrossRef](#)] [[PubMed](#)]
26. Zhang, Y.; Skolnick, J. TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **2005**, *33*, 2302–2309. [[CrossRef](#)] [[PubMed](#)]
27. Liwo, A.; Khalili, M.; Czaplewski, C.; Kalinowski, S.; Ołdziej, S.; Wachucik, K.; Scheraga, H. Modification and Optimization of the United-Residue (UNRES) Potential Energy Function for Canonical Simulations. I. Temperature Dependence of the Effective Energy Function and Tests of the Optimization Method with Single Training Proteins. *J. Phys. Chem. B* **2007**, *111*, 260–285. [[CrossRef](#)] [[PubMed](#)]
28. He, Y.; Xiao, Y.; Liwo, A.; Scheraga, H.A. Exploring the Parameter Space of the Coarse-Grained UNRES Force Field by Random Search: Selecting a Transferable Medium-Resolution Force Field. *J. Comput. Chem.* **2009**, *30*, 2127–2135. [[PubMed](#)]
29. Zaborowski, B.; Jagieła, D.; Czaplewski, C.; Hałabis, A.; Lewandowska, A.; Żmudzińska, W.; Ołdziej, S.; Karczyńska, A.; Omieczynski, C.; Wirecki, T.; et al. A Maximum-Likelihood Approach to Force-Field Calibration. *J. Chem. Inf. Model.* **2015**, *55*, 2050–2070. [[CrossRef](#)] [[PubMed](#)]
30. Faraggi, E.; Kloczkowski, A. GENN: A GENERAL NEURAL NETWORK FOR LEARNING TABULATED DATA WITH EXAMPLES FROM PROTEIN STRUCTURE PREDICTION. *Methods Mol. Biol.* **2014**, *1260*, 165–178.
31. Liwo, A.; Pincus, M.R.; Wawak, R.J.; Rackovsky, S.; Scheraga, H.A. Prediction of Protein Conformation on the Basis of a Search for Compact Structures; Test on Avian Pancreatic Polypeptide. *Protein Sci.* **1993**, *2*, 1715–1731. [[PubMed](#)]
32. Liwo, A.; Ołdziej, S.; Pincus, M.R.; Wawak, R.J.; Rackovsky, S.; Scheraga, H.A. A United-Residue Force Field for Off-Lattice Protein-Structure Simulations. I. Functional Forms and Parameters of Long-Range Side-Chain Interaction Potentials from Protein Crystal Data. *J. Comput. Chem.* **1997**, *18*, 849–873. [[CrossRef](#)]
33. Liwo, A.; Czaplewski, C.; Ołdziej, S.; Rojas, A.V.; Kaźmierkiewicz, R.; Makowski, M.; Murarka, R.K.; Scheraga, H.A. Simulation of protein structure and dynamics with the coarse-grained UNRES force field. In *Coarse-Graining of Condensed Phase and Biomolecular Systems*; Voth, G., Ed.; CRC Press: Boca Raton, FL, USA, 2008; Chapter 8, pp. 1391–1411.
34. Liwo, A.; Baranowski, M.; Czaplewski, C.; Gołaś, E.; He, Y.; Jagieła, D.; Krupa, P.; Maciejczyk, M.; Makowski, M.; Mozolewska, M.A.; et al. A Unified Coarse-Grained Model of Biological Macromolecules Based on Mean-Field Multipole-Multipole Interactions. *J. Mol. Model.* **2014**, *20*, 2306. [[CrossRef](#)] [[PubMed](#)]

35. Khalili, M.; Liwo, A.; Rakowski, F.; Grochowski, P.; Scheraga, H. Molecular Dynamics with the United-Residue Model of Polypeptide Chains. I. Lagrange Equations of Motion and Tests of Numerical Stability in the Microcanonical Mode. *J. Phys. Chem. B* **2005**, *109*, 13785–13797. [[PubMed](#)]
36. Krupa, P.; Mozolewska, M.A.; Wiśniewska, M.; Yin, Y.; He, Y.; Sieradzan, A.K.; Ganzynkiewicz, R.; Lipska, A.G.; Karczyńska, A.; Ślusarz, M.; et al. Performance of protein-structure predictions with the physics-based UNRES force field in CASP11. *Bioinformatics* **2016**, *32*, 3270–3278. [[PubMed](#)]
37. Liwo, A.; Czaplewski, C.; Pillardy, J.; Scheraga, H.A. Cumulant-Based Expressions for the Multibody Terms for the Correlation between Local and Electrostatic Interactions in the United-Residue Force Field. *J. Chem. Phys.* **2001**, *115*, 2323–2347. [[CrossRef](#)]
38. Krupa, P.; Sieradzan, A.K.; Rackovsky, S.; Baranowski, M.; Ołdziej, S.; Scheraga, H.A.; Liwo, A.; Czaplewski, C. Improvement of the Treatment of Loop Structures in the UNRES Force Field by Inclusion of Coupling between Backbone- and Side-Chain-Local Conformational States. *J. Chem. Theory Comput.* **2013**, *9*, 4620–4632. [[CrossRef](#)]
39. Sieradzan, A.K.; Krupa, P.; Scheraga, H.A.; Liwo, A.; Czaplewski, C. Physics-Based Potentials for the Coupling between Backbone- and Side-Chain-Local Conformational States in the United Residue (UNRES) Force Field for Protein Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 817–831. [[CrossRef](#)]
40. Gay, J.G.; Berne, B.J. Modification of the overlap potential to mimic a linear site-site potential. *J. Chem. Phys.* **1981**, *74*, 3316–3319. [[CrossRef](#)]
41. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
42. Seber, G.A.; Wild, C.J. *Nonlinear Regression*; Wiley: New York, NY, USA, 1989; pp. 228–236.
43. Liwo, A.; Ołdziej, S.; Czaplewski, C.; Kleinerman, D.S.; Blood, P.; Scheraga, H.A. Implementation of molecular dynamics and its extensions with the coarse-grained UNRES force field on massively parallel systems: Toward millisecond-scale simulations of protein structure, dynamics, and thermodynamics. *J. Chem. Theory Comput.* **2010**, *6*, 890–909. [[CrossRef](#)] [[PubMed](#)]
44. Lee, J.; Liwo, A.; Scheraga, H.A. Energy-based *de novo* protein folding by conformational space annealing and an off-lattice united-residue force field: Application to the 10-55 fragment of staphylococcal protein A and to apo-calbindin D9K. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 2025–2030. [[CrossRef](#)] [[PubMed](#)]
45. Duane, S.; Kennedy, A.D.; Pendleton, B.J.; Roweth, D. Hybrid monte carlo. *Phys. Lett. B* **1987**, *195*, 216–222. [[CrossRef](#)]
46. Nancias, M.; Chinchio, M.; Ołdziej, S.; Czaplewski, C.; Scheraga, H.A. Protein structure prediction with the UNRES force-field using Replica-Exchange Monte Carlo-with-Minimization; Comparison with MCM, CSA, and CFMC. *J. Comput. Chem.* **2005**, *26*, 1472–1486. [[CrossRef](#)] [[PubMed](#)]
47. Czaplewski, C.; Kalinowski, S.; Liwo, A.; Scheraga, H.A. Application of Multiplexing Replica Exchange Molecular Dynamics Method to the UNRES Force Field: Tests with  $\alpha$  and  $\alpha + \beta$  Proteins. *J. Chem. Theor. Comput.* **2009**, *5*, 627–640. [[CrossRef](#)] [[PubMed](#)]
48. Maisuradze, G.G.; Senet, P.; Czaplewski, C.; Liwo, A.; Scheraga, H.A. Investigation of protein folding by coarse-grained molecular dynamics with the UNRES force field. *J. Phys. Chem. A* **2010**, *114*, 4471–4485. [[CrossRef](#)] [[PubMed](#)]
49. Maisuradze, G.G.; Medina, J.; Kachlishvili, K.; Krupa, P.; Mozolewska, M.A.; Martin-Malpartida, P.; Maisuradze, L.; Macias, M.J.; Scheraga, H.A. Preventing fibril formation of a protein by selective mutation. *Proc. Natl. Acad. Sci. USA* **2015**, 13549–13554. [[CrossRef](#)] [[PubMed](#)]
50. Zhou, R.; Maisuradze, G.G.; Suñol, D.; Todorovski, T.; Macias, M.J.; Xiao, Y.; Scheraga, H.A.; Czaplewski, C.; Liwo, A. Folding kinetics of WW domains with the united residue force field for bridging microscopic motions and experimental measurements. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 18243–18248. [[CrossRef](#)]
51. Mozolewska, M.A.; Krupa, P.; Scheraga, H.A.; Liwo, A. Molecular modeling of the binding modes of the iron-sulfur protein to the Jac1 co-chaperone from *Saccharomyces cerevisiae* by all-atom and coarse-grained approaches. *Proteins Struct. Funct. Bioinform.* **2015**, *83*, 1414–1426. [[CrossRef](#)]
52. Lipska, A.G.; Sieradzan, A.K.; Krupa, P.; Mozolewska, M.A.; Dž019Auria, S.; Liwo, A. Studies of conformational changes of an arginine-binding protein from *Thermotoga maritima* in the presence and absence of ligand via molecular dynamics simulations with the coarse-grained UNRES force field. *J. Mol. Model.* **2015**, *21*, 1–11.

53. Sieradzan, A.K. Introduction of Periodic Boundary Conditions into UNRES force field. *J. Comput. Chem.* **2015**, *36*, 940–946. [[PubMed](#)]
54. Sieradzan, A.K.; Hansmann, U.H.E.; Scheraga, H.A.; Liwo, A. Extension of UNRES Force Field to Treat Polypeptide Chains with d-Amino Acid Residues. *J. Chem. Theory Comput.* **2012**, *8*, 4746–4757. [[PubMed](#)]
55. Mozolewska, M.A.; Sieradzan, A.K.; Niadzvedstki, A.; Czaplewski, C.; Liwo, A.; Krupa, P. Role of the sulfur to  $\alpha$ -carbon thioether bridges in thurincin H. *J. Biomol. Struct. Dyn.* **2017**, *35*, 2868–2879. [[CrossRef](#)] [[PubMed](#)]
56. Kynast, P.; Derreumaux, P.; Strodel, B. Evaluation of the coarse-grained OPEP force field for protein-protein docking. *BMC Biophys.* **2016**, *9*, 4. [[CrossRef](#)] [[PubMed](#)]
57. Faraggi, E.; Zhang, T.; Yang, Y.; Kurgan, L.; Zhou, Y. SPINE X: Improving protein secondary structure prediction by multistep learning coupled with prediction of solvent accessible surface area and backbone torsion angles. *J. Comput. Chem.* **2012**, *33*, 259–267. [[CrossRef](#)] [[PubMed](#)]
58. Faraggi, E.; Zhou, Y.; Kloczkowski, A. Accurate single-sequence prediction of solvent accessible surface area using local and global features. *Proteins Struct. Funct. Bioinform.* **2014**, *82*, 3170–3176. [[CrossRef](#)]
59. Moult, J.; Fidelis, K.; Krysztafowych, A.; Schwede, T.; Tramontano, A. Critical assessment of methods of protein structure prediction: Progress and new directions in round XI. *Proteins Struct. Funct. Bioinform.* **2016**, *84*, 4–14. [[CrossRef](#)]
60. Eramian, D.; Eswar, N.; Shen, M.Y.; Sali, A. How well can the accuracy of comparative protein structure models be predicted? *Protein Sci.* **2008**, *17*, 1881–1893. [[CrossRef](#)]
61. Yang, Y.; Zhou, Y. Ab initio folding of terminal segments with secondary structures reveals the fine difference between two closely related all-atom statistical energy functions. *Protein Sci.* **2008**, *17*, 1212–1219. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).