A comparative study of global minimum energy conformations of hydrated peptides

J. L. Klepeis and C. A. Floudas*
Department of Chemical Engineering
Princeton University
Princeton, N.J. 08544-5263

Accepted for Publication
Journal of Computational Chemistry

*Author to whom all correspondence should be addressed.

Abstract

A global optimization method is described for identifying the global minimum energy conformation, as well as lower and upper bounds on the global minimum conformer of solvated peptides. In considering the effects of hydration, two independent continuum based solvation models are employed. The first method is based on the calculation of solvent–accessible surface areas, while the second method uses information on the solvent–accessible volume of hydration shells. The hydration effects predicted by the area and volume based models, using a variety of atomic solvation parameter (ASP) sets, are tested and compared by identifying global minimum energy structures of terminally blocked peptides and oligopeptides. Significant differences are observed, indicating that appropriate model selection is essential for accurately predicting hydrated peptide structures. Using this information, the applicability of these hydration models and ASP sets is discussed.

Keywords : Protein folding, Solvation, Global optimization, Met–Enkephalin, Leu–Enkephalin

1 Introduction

The prediction of native folded three-dimensional structures of proteins is an important problem in the field of computational chemistry. Success in this area would have major ramifications in both the theoretical and experimental study of proteins. It is widely accepted that, by employing the thermodynamic hypothesis[1,2,3], the folded conformation can be found by identifying the structure exhibiting the global minimum free energy. Therefore, a significant aspect of the protein folding problem involves the identification of the global minimum energy structure. In this direction, the ability to accurately model the effects of solvation is of primary importance.

The identification of the global minimum energy structure, with or without solvation, requires the use of efficient methods to search the nonconvex multi–dimensional conformation space. A large number of techniques have been developed, with varying degrees of success, to treat the multiple–minima problem. A number of recent review papers have surveyed the treatment of the protein conformation problem in terms of the global minimization of nonconvex energy functions[4,5,6,7,8]. This work addresses the protein folding problem through the use of a novel deterministic global optimization algorithm. This branch–and–bound based global optimization algorithm, αBB, is applicable to a large class of nonlinear optimization problems that have twice–differentiable functions[9,10,11,12,13].

Through the use of the αBB method, the identification of the global minimum energy structure, along with lower and upper bounds on this value, can be effectively identified so that the effects of solvation can be isolated and compared. Several continuum based solvation models, including a number of atomic solvation parameter (ASP) sets for both area and volume based models, are used to predict the global minimum structures of terminally blocked residues and the oligopeptides met–enkephalin and leu–enkephalin. The structural differences predicted by these models are then quantified and compared.

2 Potential Energy Model

There are a number of approaches that may be used to model protein interaction energies. In reality, the dynamics of atoms are governed by the quantum theory of their participating electrons. Using the Born-Oppenheimer approximation, one can determine the energy for fixed atomic nuclei from the smallest eigenvalue of the Hamiltonian of the electron system. These approximations and their derivatives are calculated using ab–initio methods. However, due to their computational complexity, such calculations are limited to extremely small molecules. Less detailed, semi-empirical methods are based on all atom representations of the peptide. In general, these models, also known as force fields, are expressed as summations of empirically derived potential functions, with the mathematical form of individual energy terms based on the phenomenological nature of that term. Finally, a variety of simplified models have been used to reduce the degrees of freedom associated with the conformational energy.

In this work, the ECEPP/3 (Empirical Conformational Energy Program for Peptides)[14] potential model is utilized. In this force field, it is assumed that the covalent bond lengths and bond angles are fixed at their equilibrium values. Then, the conformation is only a function of the independent torsional angles of the system, also known as dihedral angles. The total conformational energy is calculated as the sum of the electrostatic, nonbonded, hydrogen bonded, and torsional contributions. There is also a pseudo–potential for loop closing if the polypeptide contains two or more sulfur–containing residues. More recent work includes a revised treatment of prolyl and hydroxyprolyl residues[14]. For each prolyl or hydroxyprolyl residue contained in the polypeptide a fixed internal conformational energy for the pyrolidine ring is added. The main energy contributions (electrostatic, nonbonded, hydrogen bonded) are computed as the sum of terms for each atom pair (i,j) whose interatomic distance is a function of at least one dihedral angle. The general potential energy terms of ECEPP/3 are shown in Figure 1, while the development of the appropriate parameters is discussed and reported elsewhere[14].

3 Solvation Models

Solvation contributions are generally believed to be a significant force in stabilizing the native conformations of proteins. Explicit methods can be used to include solvation effects by actually surrounding the polypeptide with solvent molecules and calculating solvent–peptide and solvent–solvent interactions. Although these methods are conceptually simple, explicit inclusion of solvent molecules greatly increases the computational time needed to simulate the polypeptide system. Therefore, most simulations of this type are limited to restricted conformational searches. In addition, it is difficult to quantify the effect of hydrophobic interactions that result from the ordering of water molecules.

Methods for estimating solvent free energies have also been developed using both integral equations and continuum models. Integral equation methods can be used to evaluate solvent structure and thermodynamic properties. Typically, molecular dynamics or Monte Carlo simulations are used to calculate ensemble averages from which free energy differences can be obtained. A number of methods have been proposed to estimate these solvation free energies from simulations based on molecular dynamics and Monte Carlo averages[15,16,17].

The integral equation method has also been used to analyze the solvent structure of a protein system[18]. In contrast, continuum models use a simplified representation of the solvent environment by neglecting the molecular nature of the water molecules. Calculations of solvation free energies using electrostatic continuum models rely on numerical solutions to the Poisson–Boltzmann equation from which dielectric and ionic strength effects are obtained [19]. Other continuum models estimate free energies of solvation as a function of surface areas and volumes.

In this work, solvation contributions are included implicitly using empirical correlations with both surface area[20] and volume[21]. The main assumption of these models is that, for each functional group of the peptide, a hydration free energy can be calculated from an averaged free energy of interaction of the group with a layer of solvent known as the hydration shell. In addition, the total free energy of hydration is expressed as a sum of the free energies of hydration for each of the functional groups of the peptide; that is, an additive relationship is assumed.

3.1 Accessible Surface Area

Accessible surface area methods assume that the free energy of hydration is proportional to the solvent–accessible surface area of the peptide, as described by the following equation:

$$E_{HYD} = \sum_{i=1}^{N}(A_i)(\sigma_i) \tag{1}$$

In Equation (1), an additive relationship for N individual functional groups is assumed. (A_i) represents the solvent–accessible surface area for the functional group, and (σ_i) is an empirically derived free energy density parameter.

There are a number of ways to define the surface of a peptide. In developing these surfaces the peptide is represented by a union of spheres, with the radii of the spheres set by the van der Waals radii of the constituent atoms. A spherical test probe is then rolled over these spheres, thereby tracing out a surface. The molecular surface is set by direct contact between the probe sphere and the peptide spheres. In areas where the probe cannot make direct contact, the closest part of the probe is used. The solvent–accessible surface is defined by the surface traced by the center of the probe as the probe rolls over the peptide spheres. These areas depend on the radius of the probe sphere; when this radius is set to zero both the molecular and solvent–accessible surface areas become the van der Waals surface of the peptide.

Solvent–accessible surface areas are calculated using the MSEED[20] program, which employs algorithms developed by Connolly[22]. MSEED eliminates many unnecessary computations by considering only those convex faces that are on the accessible surface. Rigorous implementation of Connolly's method requires the calculation of interior surface areas, which are ultimately found to be zero. A full description of the MSEED algorithm is given elsewhere [20]. A number of other methods for calculating surface areas are also available[23,24,25].

One potential problem that may arise when calculating accessible surface areas is the appearance of gradient discontinuities. This may occur when a new vertex or edge appears on the surface. If the discontinuity is large, minimization techniques requiring gradients may fail to converge to the local minimum conformation. A complete analysis of all situations for which the gradient of the molecular surface area becomes discontinuous has been reported [26].

3.1.1 Atomic Solvation Parameter (ASP) Sets

Once the solvent–accessible surface areas have been calculated, these values must be multiplied by the appropriate (σ_i) parameters as shown in Equation (1). A variety of parameter sets have been developed to model the transfer of atoms from a gaseous to a hydrated environment. The parameter values for the five ASP sets used in this study are given in Table I.

The ASP sets WE1 and WE2, are taken from Table 3 of Wesson and Eisenberg (1992). These parameters are both derived from Wolfenden's measured free energies of transfer of amino acid side-chain analogs from vapor to water[28]. Both sets have been adjusted to correct for entropy of mixing effects based on solute and solvent size differences[29,30], although the applicability of these corrections has been criticized[31,32]. The parameters for these two sets are negative for all atoms excluding carbon. Qualitatively, this means that the nitrogen, oxygen and sulfur atoms are considered hydrophilic, that is they favor solvent exposure. Comparatively, the WE1 and WE2 parameters are similar, with the largest relative difference being a 3:1 ratio (WE1:WE2) for the σ_C parameter. Therefore, the hydrophobic character of these carbon atoms is stronger for the WE1 ASP set.

The OONS parameter set has been specifically developed to supplement the ECEPP/2 force field[33]. These parameters were derived by a least squares fitting to experimental free energies of gas to water transfer of small aliphatic and aromatic molecules. The most significant difference from the two previous ASP sets is a substantial increase in hydrophobic character for carboxyl (carbonyl) carbon atoms, which corresponds to a hundred-fold increase when compared to the same WE2 parameter. In addition, the free energy parameter becomes negative for aromatic carbons, which indicates a hydrophilic tendency. The three-fold decrease of the OONS values for carboxyl (carbonyl) and charged oxygen atom parameters, as compared to both the WE1 and WE2 ASP sets, is also significant.

Unlike the aforementioned models, the SCKS ASP set is not directly based on experimental free energies[34]. Instead, it is an optimized parameter set developed to complement the CHARMM[35] molecular mechanics force field. Specifically, through the use of experimental and molecular dynamics information, the relative weightings of solvation parameters were refined to provide the best correspondence between minimized and experimental structures. In comparing the individual free energy parameters, it is evident that the hydrophobic character of the carbon atoms is increased approximately three- and eight-fold over the WE1 and WE2 values, respectively. In contrast, the uncharged oxygen and nitrogen atom parameters are 6.5 times smaller (less hydrophilic) than those for the WE1 and WE2 ASP sets. This decrease does not apply to charged oxygen and nitrogen atoms, which possess extremely hydrophilic values.

The JRF ASP set was derived from NMR studies of low energy solvated configurations of 13 tetrapeptides[36]. This represents an important difference from other derivations because actual peptides, rather than simple model compounds, were used to develop the JRF parameters. An ensemble of low energy structures for these tetrapeptides was also produced using the ECEPP/2 potential function. Then, a nonlinear least–squares system was optimized for the best set of atomic solvation parameters. Although, the parameters for oxygen, nitrogen, and sulfur atoms are negative, their large absolute values indicate much larger hydrophilicities than corresponding atoms of any other ASP set. In addition, both the carboxyl (carbonyl) and aromatic carbon atoms possess strong hydrophilic parameters, which contradicts other free energy parameter values for these atoms. The single positive value belongs to the aliphatic carbon atom type, which, although larger than any other parameter for this atom type, possesses the smallest magnitude for the JRF ASP set. Furthermore, because it was developed from minimum energy conformations of peptides, the JRF ASP set has been shown to produce undesirable perturbations during local minimizations if the solvation energy contributions are added at every iteration. Therefore, unlike the aforementioned ASP sets, the JRF solvation energy effects are only included at local minimum conformations.

3.2 Accessible Volume of Hydration Shell

For volume shell models, the free energy of hydration is assumed to be proportional to the water–accessible volume of a hydration layer surrounding the peptide. This can be represented in the form:

$$E_{HYD} = \sum_{i=1}^{N}(VHS_i)(\delta_i) \tag{2}$$

An additive relationship for the N individual atoms of the peptide is assumed, and (VHS_i) represents the solvent–accessible volume of hydration shell for each atom i which is exposed to water. The (δ_i) parameters are empirically determined free energy of hydration densities for these atoms.

The hydration shell is defined by the volume inside a sphere of radius R_i^h but outside a sphere of radius R_i^v, with both radii centered on atom i. The larger radius, R_i^h, corresponds to the radius of the first hydration shell of atom i, while R_i^v is equal to the van der Waals radius. In order to calculate (VHS_i), the volume of a collection of overlapping hard spheres must be computed using:

$$V(\mathbf{R}) = \sum_i a_i S_i - \sum_{ij} b_{ij} D_{ij} + \sum_{ijk} c_{ijk} T_{ijk} - \sum_{ijkl} d_{ijkl} Q_{ijkl} \tag{3}$$

In Equation (3), S_i signifies the volume of a single sphere, while D_{ij}, T_{ijk} and Q_{ijkl} represent the volume of intersection of two, three and four spheres, respectively. This is sufficient because all higher order overlaps can be decomposed into the three types of intersections included in Equation (3). Therefore, the solvent–accessible volume of hydration can be written as:

$$(VHS_i) = V(R_i^h) - V(R_i^v) \tag{4}$$

The first term in Equation (4) is calculated using Equation (3) with the radii of all atoms set equal to their van der Waals radii, while the second term is calculated with the radius of atom i equal to R_i^h and the van der Waals radii of all the other atoms. A number of methods to compute hydration shell volumes have been proposed[21,37,38].

The form of Equation (3) is not suitable for force field models using pairwise intramolecular potential, such as ECEPP/3. Furthermore, direct truncation at the double–overlap term would lead to large errors. In this work, the RRIGS (Reduced Radius Independent Gaussian Sphere) approximation is used to efficiently calculate the exposed volume of the hydration shell[21]. This method uses a truncated form of Equation (3) but also artificially reduces the van der Waals radii of all atoms other than atom i when calculating (VHS_i). These reductions effectively decrease the contribution of the double overlap terms, leading to a cancellation of the error which results from neglecting the triple and higher overlap terms. In addition, the characteristic density of being inside the overlap volume of two intersecting spheres is not represented as a step function, but as a Gaussian function; this provides continuous derivatives of the hydration potential. Therefore, the solvation energy contributions can easily be added at every step of local minimizations since the RRIGS approximation has the same set of interactions as the ECEPP/3 potential.

Free energy density parameters for solvent accessible volumes have been developed for nonionic and charged organic solute molecules[39,40,41]. In this work, RRIGS specific (δ_i) parameters, which were developed by a least square fitting of experimental free energy of solvation data for 140 small organic molecules[21], are used (see Supplementary Material, Table SI). The classification of the RRIGS atom types is more fragmented than for the solvent accessible surface area ASP sets. The most hydrophilic values belong to the nitrogen and selected oxygen and hydrogen atom types. In addition, aromatic carbons tend to possess slightly hydrophilic values, whereas the carbonyl and aliphatic carbon atoms exhibit the most hydrophobic parameter values.

4 Global Optimization

The energy minimization problem can be formulated as a nonconvex nonlinear optimization problem. Let $i = 1, \ldots, N_{RES}$ be an indexed set describing the sequence of amino acid residues in the peptide chain. There are $\phi_i, \psi_i, \omega_i,\ i = 1, \ldots, N_{RES}$ dihedral angles along the backbone of this peptide. In addition, let K^i denote the number of dihedral angles for the side chain of the i^{th} residue and J^N and J^C denote the number of dihedral angles for the amino and carboxyl end groups, respectively. Using these definitions the optimization problem takes the following form:

$$\min \qquad E(\phi_i, \psi_i, \omega_i, \chi_i^k, \theta_j^N, \theta_j^C) \tag{5}$$

$$\begin{aligned}
\text{subject to} \quad -\pi &\leq \phi_i \leq \pi, \quad i = 1, \ldots, N_{RES} \\
-\pi &\leq \psi_i \leq \pi, \quad i = 1, \ldots, N_{RES} \\
-\pi &\leq \omega_i \leq \pi, \quad i = 1, \ldots, N_{RES} \\
-\pi &\leq \chi_i^k \leq \pi, \quad i = 1, \ldots, N_{RES},\ k = 1, \ldots, K^i \\
-\pi &\leq \theta_j^N \leq \pi, \quad j = 1, \ldots, J_N \\
-\pi &\leq \theta_j^C \leq \pi, \quad j = 1, \ldots, J_C
\end{aligned}$$

In general, E represents the total potential energy function and the free energy of solvation. However, in the case of the JRF ASP set, the potential energy function is minimized before adding the hydration energy contributions for this ASP set. In other words, gradient contributions from solvation are not considered. This approach is represented by the following equation:

$$E_{JRF}^{Total} = E_{Min}^{Unsol} + E_{JRF}^{Sol} \tag{6}$$

Even after reducing this optimization problem to a function of internal variables (dihedral angles), the multidimensional surface that describes the energy function has an astronomically large number of local minima. This has become known as the multiple–minima problem. In addition, evaluation of the potential, especially with the addition of solvation, is computationally expensive, which increases the computational time for even local minimization. Because the objective function has many local minima, using local optimization techniques depends on the initial points selected. Therefore, rigorous global optimization algorithms are needed to effectively locate the global minimum corresponding to the native state of the protein. A large number of techniques have been developed to search this nonconvex conformational space. In general, the major limitation is that these methods depend heavily on the supplied initial conformation. As a result, there is no guarantee for global convergence because large sections of the domain space may be bypassed.

In order to overcome these difficulties, the αBB global optimization approach[9,10,11,12,13] has been extended to identifying global minimum energy conformations of solvated peptides. The development of this branch and bound method was motivated by the need for an algorithm that could guarantee convergence to the global minimum of nonlinear optimization problems with twice–differentiable functions[42]. The application of this algorithm to the minimization of potential energy functions was first introduced for microclusters[43,44], and small acyclic molecules[45,46]. The αBB approach has also been applied to constrained optimization problems[10,11,12,13]. In more recent work, the algorithm has been shown to be successful for isolated peptide systems using the realistic ECEPP/3 potential energy model [47,48].

4.1 Minimization of Potential Energy using αBB

The αBB global optimization algorithm effectively brackets the global minimum solution by developing converging sequences of lower and upper bounds. These bounds are refined by iteratively partitioning the initial domain. Upper bounds on the global minimum are obtained by local minimizations of the original energy function, E. Lower bounds belong to the set of solutions of the convex lower bounding functions, which are constructed by augmenting E with the addition of separable quadratic terms. The lower bounding functions, L, of the energy hypersurface can be expressed in the following manner:

$$L \;=\; E + \sum_{i=1}^{N_{RES}} \alpha_{\phi,i} \left(\phi_i^L - \phi_i\right) \left(\phi_i^U - \phi_i\right) + \tag{7}$$

$$\sum_{i=1}^{N_{RES}} \alpha_{\psi,i} \left(\psi_i^L - \psi_i\right) \left(\psi_i^U - \psi_i\right) +$$
$$\sum_{i=1}^{N_{RES}} \alpha_{\omega,i} \left(\omega_i^L - \omega_i\right) \left(\omega_i^U - \omega_i\right) +$$
$$\sum_{i=1}^{N_{RES}} \sum_{k=1}^{K^i} \alpha_{\chi,i,k} \left(\chi_i^{k,L} - \chi_i^k\right) \left(\chi_i^{k,U} - \chi_i^k\right) +$$
$$\sum_{j=1}^{J^N} \alpha_{j,\theta^N} \left(\theta_j^{N,L} - \theta_j^N\right) \left(\theta_j^{N,U} - \theta_j^N\right) +$$
$$\sum_{j=1}^{J^C} \alpha_{j,\theta^C} \left(\theta_j^{C,L} - \theta_j^C\right) \left(\theta_j^{C,U} - \theta_j^C\right)$$

Here $\phi_i^L, \psi_i^L, \omega_i^L, \chi_i^{k,L}, \theta_j^{N,L}, \theta_j^{C,L}$ and $\phi_i^U, \psi_i^U, \omega_i^U, \chi_i^{k,U}, \theta_j^{N,U}, \theta_j^{C,U}$ represent lower and upper bounds on the dihedral angles $\phi_i, \psi_i, \omega_i, \chi_i^k, \theta_j^N, \theta_j^C$. The α parameters represent nonnegative parameters which must be greater or equal to the negative one–half of the minimum eigenvalue of the Hessian of E over the defined domain. Rigorous bounds on the α parameters can be obtained through a variety of approaches[9,10,49,45]. The overall effect of these terms is to overpower the nonconvexities of the original nonconvex terms by adding the value of 2α to the eigenvalues of the Hessian of E. The convex lower bounding functions, L, possess a number of important properties which guarantee global convergence[46]:

(i) L is a valid underestimator of E;

(ii) L matches E at all corner points of the box constraints;

(iii) L is convex in the current box constraints;

(iv) the maximum separation between L and E is bounded and proportional to α and to the square of the diagonal of the current box constraints. This property ensures that a feasibility and convergence tolerance can be reached for a finite size partition element;

(v) the underestimators L constructed over supersets of the current set are always less tight than the underestimator constructed over the current box constraints for every point within the current box constraints.

Once solutions for the upper and lower bounding problems have been established, the next step is to modify these problems for the next iteration. This is accomplished by successively partitioning the initial domain into smaller subdomains. For the protein conformation problems, it has been found that an effective partitioning strategy involves bisecting the same variable dimension across all nodes at a given level. In order to ensure non–decreasing lower bounds, the hyper–rectangle to be bisected is chosen by selecting the region which contains the infimum of the minima of lower bounds. A non–increasing sequence for the upper bound is found by solving the nonconvex problem, E, locally and selecting it to be the minimum over all the previously recorded upper bounds. Obviously, if the single minimum of L for any hyper–rectangle is greater than the current upper bound, this hyper–rectangle can be discarded because the global minimum cannot be within this subdomain (fathoming step). The computational requirement of the αBB algorithm depends on the number of variables (global) on which branching occurs. Therefore, these global variables need to be chosen carefully.

4.2 Algorithmic Description

The determination of the global minimum energy conformation using αBB requires the interfacing of a number of programs (αBB[9,10,11,12,13], PACK[50], NPSOL[51] and potential and solvation energy modules). PACK, a peptide generation program, is called once directly by αBB in order to initialize the current problem. In subsequent steps PACK is called through NPSOL[51], a local nonlinear optimization solver used to solve both the upper and lower bounding problems. PACK internally transforms to and from Cartesian and internal coordinate systems, and provides potential energy and gradient contributions for the ECEPP/3 potential model at every step of the local minimizations. When considering surface–accessible solvation, surface areas are calculated using MSEED[20]; whereas volumes of hydration shells are determined using the RRIGS module[21]. Finally, an additional module, UBC (Upper Bound Check), is used to verify the quality of the upper bound solutions. The overall interface is shown schematically in Figure 2.

The basic steps of the algorithm are as follows:

(1) The initial best upper bound is set to an arbitrarily large value. The original domain is partitioned along one of the global variables.

(2) A convex function (L) is constructed in each hyper–rectangle and minimized using NPSOL, with calls (through PACK) to both ECEPP/3 and one of the two solvation modules. If a solution is greater than the current best upper bound the entire subregion can be fathomed, otherwise the solution is stored.

(3) The local minima solutions for L are used as initial starting points for local minimizations of the upper bounding function (E) in each hyper–rectangle. Again, the appropriate calls are made to PACK and the potential and solvation energy modules. In solving the upper bounding problems, all variable bounds are expanded to

[-180,180]. These solutions are upper bounds on the global minimum solution in each hyper–rectangle.

(4) The current best upper bound is updated to be the minimum of those thus far stored. If a new upper bound (from step 3) is selected, the upper bound check, UBC, module is called. UBC checks that the absolute value of each gradient in the objective function gradient vector is below a specified tolerance (kcal/mole/deg). If a gradient does not satisfy this check the corresponding variable bounds are incrementally increased and the problem is solved with the previous point used as the initial starting point. This process is repeated until the gradient constraints are satisfied or an iteration limit is exceeded. UBC also employs algorithms to calculate the second derivative matrix[52] which is used to verify that the upper bound solution is a local minimum, that is the Hessian matrix is positive semi–definite. If the matrix is not positive semi–definite or the gradient checks are not satisfied, the upper bound solution is rejected.

(5) The hyper–rectangle with the current minimum value for L is selected and partitioned along one of the global variables.

(6) If the best upper and lower bounds are within the ϵ tolerance the program will terminate, otherwise it will return to Step 2.

5 Computational Studies

5.1 Terminally–Blocked Residues

The single residue examples were defined as terminally blocked by using acetyl (amino) and methyl (carboxyl) end groups. All dihedral angles were treated as global variables, excluding the three θ angles of the end groups. The relative convergence was set to 10^{-2}. For these examples, all dihedral angles, excluding those of the end groups, were treated as global variables. The remaining variables were treated locally; that is, they were allowed to vary during minimization, but their domain space was not partitioned. When using the RRIGS and JRF models the global variables were assigned initial α values of 3.0. For the other solvation models, the α values were increased to 5.0. The results for these examples, which provide energetic information on the global minima, are provided as Supplementary Material (Tables SII through SVII).

For a number of residues, the JRF global minimum solutions possess ω angles in the range of [-30,30] with the corresponding ϕ and ψ angles near the [-150,80] region. Additional runs were conducted in which the ω angles were constrained to the range of [160,200]. In all cases, with the exception of serine, this constraint led to increases in solvation energies and decreases in potential energy terms while the structures became either β–sheet–like or α–helical (see Table SVIII of Supplementary Material). Without exception, the ω angles for the all other global minimum energy solutions were within the [160,200] range. The remaining analysis in this section refers to these constrained (ω within [160,200]) minima for the JRF ASP set. This is appropriate not only in comparing the JRF results with other solvation results, but it also makes the analysis relevant for the oligopeptide studies because similar ω bounds are typically used.

The results of the solvation models are more clearly evaluated when examining energy differences (data shown in Tables SIX and SX of Supplementary Material). For example, ΔE^{POT} ($\Delta E^{POT} = E^{POT}_{ASP}$ - E^{POT}_{RRIGS}) refers to the change in potential energy of an area based global minimum (E^{POT}_{ASP}) and the RRIGS global minimum (E^{POT}_{RRIGS}) solution for a given terminally blocked residue. This difference is positive in almost all cases, which indicates that the potential energy of the RRIGS structure is always lower and provides more stabilization at the corresponding global minimum solution. In most cases, this difference is very small, especially for the OONS and SCKS ASP sets. In fact, for both these ASP sets several residues, most noticeably phenylalanine and tyrosine, have more potential energy stabilization at their corresponding global minima. However, for five peptides, namely phenylalanine, serine, threonine, tyrosine and leucine, the JRF potential energy is more than 10 kcal/mole less stabilizing. This set of residues includes the three residues (serine, threonine and tyrosine) that contain hydroxyl groups among the side chain atoms, as well as the two regular aromatic residues (phenylalanine and tyrosine). It is also interesting to note that these atom types (i.e., hydroxyl oxygen and aromatic carbon) correspond to two of the most hydrophilic type atoms in the JRF ASP set. Finally, the leucine ΔE^{POT} seems to be abnormally high because of the large torsional contribution at the JRF global minimum conformation.

The results for ΔE^{HYD} ($\Delta E^{HYD} = E^{HYD}_{ASP}$ - E^{HYD}_{RRIGS}), which refers to the change in hydration energy between an area based global minimum (E^{HYD}_{ASP}) and the RRIGS global minimum (E^{HYD}_{RRIGS}) solutions, are especially interesting. These differences are positive in most cases, which indicates that the hydration energy of the RRIGS structure is generally lower. However, when examining the JRF results ΔE^{HYD} is negative for four examples, namely histidine, phenylalanine, tryptophan and tyrosine. Excluding the special case of proline, these four residues correspond to the naturally occurring residues which possess ringed side chain structures. Other trends are also apparent. The most positive ΔE^{HYD} values for the JRF ASP set are provided by the aliphatic residues. In addition, the acidic residues, glutamic and aspartic acid, and the amide forms of these residues, glutamine and asparagine, have comparable values for ΔE^{HYD}.

For the other (gradient inclusive) ASP sets the ΔE^{HYD} of different residues are less varied. However, it is important to consider that for all residues, excluding tyrosine, the ASP sets follow a WE2, WE1, OONS and SCKS order when ranked beginning with the most stabilizing hydration energy. Low hydration energies are expected for WE2 because of the consistently small hydrophobic and relatively large hydrophilic parameters. In most cases, the WE1 ΔE^{HYD} are only slightly larger than for WE2. This can be directly attributed to the increased hydrophobicity of the free energy parameter for the carbon atoms of the WE1 ASP set. When comparing the OONS and WE1 ASP sets the increased ΔE^{HYD} is more noticeable, which is most likely a result of the combined effects of the strong hydrophobic value of the carboxyl (carbonyl) carbon parameter and the decreased hydrophilic value of the carboxyl (carbonyl) oxygen parameter for the OONS ASP set. However, for aromatic residues (i.e., phenylalanine, tryptophan and tyrosine) these effects are partially offset by introducing a hydrophilic character for aromatic carbons in the OONS ASP set. In fact, for tyrosine this change is strong enough to cause the OONS ASP set to produce a more stabilizing hydration energy than the WE1 ASP set. A comparison between the OONS and SCKS reveals the largest increase in ΔE^{HYD} values. This can partly be attributed to the relatively large value of the free energy parameters for carbon atoms of the SCKS ASP set. The increase is also enhanced for aromatic residues because of the hydrophilic nature of the aromatic carbon atoms for the OONS ASP set. In addition, for residues with nitrogen containing side chains, the ΔE^{HYD} increase is heightened because of a subsequent decrease in the value of the free energy parameter for nitrogen atoms in the SCKS ASP set. Finally, a comparison of other surface accessible solvation results to the JRF results is qualitatively similar to those made between the RRIGS and JRF models. Specifically, the strong hydration energy stabilization of ring containing residues, as well as the decreased stabilization provided by aliphatic residues, is evident.

A more detailed analysis was performed by generating adiabatically relaxed ϕ–ψ maps for N–acetyl–N'–methyl–alanineamide (shown in Figures S1 to S7 of Supplementary Material). The adiabatic curves define regions within a given energy of the global minimum value. The first map corresponds to an adiabatically relaxed map for the unsolvated form of the peptide. This was calculated by fixing the ϕ and ψ angles at 3 degree increments and using a local minimization solver to minimize the ECEPP/3 potential energy by varying the remaining dihedral angles. The other maps were constructed by a similar procedure, although the minimized energy now included both ECEPP/3 and the appropriate hydration free energy. In generating the data for the JRF, the ECEPP/3 energy was first minimized in the absence of solvent at each point and the map was generated by adding the solvation free energy for the JRF model at the minimized conformation.

These maps reveal several important effects of including solvation (notation for the peptide regions are given in Table SXI of Supplementary Materials). Experimental data for the alanine peptide suggests that more than one conformation is present in solution, and NMR coupling constants indicate a large population of conformations with -70 > ϕ > -80[53]. It is also expected that hydration may weaken intrapeptide hydrogen bonding. The unsolvated map indicates well defined regions for intramolecular hydrogen bonding (C_7) and for right–handed α–helices (α_R). The global minimum occurs within the C_7 region. The RRIGS map retains some features of the unsolvated map, with the global minimum in the C_7 region and a very strong α_R region. However, there is a broadening of the β–sheet (C_5) region as well as a less distinct C_7 minimum. This can be contrasted with both the WE1 and WE2 adiabatic maps, which exhibit large C_5 regions and significant decreases in the size of both the C_7 and α_R regions. The OONS map contains an even larger low energy region that connects the C_5 and C_7 domains. The α_R low energy region is also broader than either of those indicated by the WE1 or WE2 map. In all three cases (WE1, WE2 and OONS) the global minimum is shifted to the β–sheet domain. In contrast, the SCKS adiabatic map is more similar to the RRIGS map because of its smaller and disjoint C_5 and C_7 regions, as well as the location of the global minimum in the C_7 well. The largest disparity between these maps exists with the JRF adiabatic map, which indicates a complete shift away from the C_7 minimum towards the C_5 region.

Qualitatively, similar trends are observed for the ϕ–ψ distribution of other terminally blocked amino acids (see Table SXII of Supplementary Material). The RRIGS model predicts a majority of global minima in the C_7 region, which indicates a tendency to preserve certain potential energy effects. As expected, the majority of WE1 and WE2 global minima lie within the C_5 domain, with the same distribution for each parameter set. The most uniform distribution of global minima belongs to the OONS ASP set, for which there are an almost equal number of C_5 and C_7 global minimum structures. This agrees with the large low energy regions displayed on the N–acetyl–N'–methyl–alanineamide adiabatic map. The large population of C_7 global minima for the SCKS ASP set is also suggested by the strong C_7 region on the N–acetyl–N'–methyl–alanineamide map. In accordance with the distinct implementation of the JRF model, these results are less predictable. Specifically, although almost half of the JRF global minima lie in the C_5 domain, a significant number also exhibit α-helical type structures, which contrasts with the ϕ-ψ map of N–acetyl–N'–methyl–alanineamide.

5.2 Met–enkephalin

Met–enkephalin (H–Tyr–Gly–Gly–Phe–Met–OH) is an endogenous opioid pentapeptide found in the human brain, pituitary, and peripheral tissues and is involved in a variety of physiological processes. The peptide consists of 24 independent torsional angles and a total of 75 atoms, and has played the role of a benchmark molecular conformation problem. The energy hypersurface is extremely complex with the number of local minima estimated on the order of 10^{11} [54]. Based on a previous study, the unsolvated global minimum potential energy conformation, with an ECEPP/3 energy of -11.707 kcal/mole, was shown to exhibit a type II' β–bend along the N–C' peptidic bond of Gly^3 and Phe^4 [48].

In studying the effects of solvation on the structure of met–enkephalin, the results for the unsolvated structure were verified by employing the algorithm outlined in Section 4.2. A major difference from the previous implementation[48] is the addition of the UBC module, as well as the expansion of all variable bounds (to [-180,180]) when solving the upper bounding problems. Since the backbone dihedral angles (i.e., ϕ and ψ) are the most influential variables in defining the backbone structure, the corresponding 10 backbone dihedral angles were treated as global variables for the enkephalin problems. Although they were not partitioned during the global search, all other variables (i.e., ω and all χ) were allowed to vary during local minimizations. The global variables were assigned initial α values of 5.0 when using the unsolvated, RRIGS and JRF models, and values of 10.0 for all other models. In the case of unsolvated met–enkephalin, the structural and energetic results of the previously identified global minimum energy structure[48] were confirmed.

Experimental results have indicated that met–enkephalin in aqueous solution does not possess an unique structure[55]. In general, experimentally determined aqueous conformations are found to exhibit characteristics of extended random–coil polypeptide with no discernible secondary structure. When considering the effects of hydration, the competition for backbone hydrogen bonding (with water), which contributes to the bending of the unsolvated conformation, should result in a more extended structure.

The RRIGS model predicts a more extended structure than the global minimum structure reported for the unsolvated case[48]. In fact, although a slight turn occurs near the N–terminus, the structure possesses no hydrogen bonds (< 2.2 $\mathring{A}$) and an overall end–to–end C^α distance of 10.16 $\mathring{A}$. In addition, there exists close proximity of the Tyr and Phe aromatic rings, as shown in Figure 3. The centroids of these rings are separated by 4.16 $\mathring{A}$, which is slightly closer than the preferential aromatic–aromatic interaction distance of 4.5 to 7 $\mathring{A}$[56]. Furthermore, the aromatic rings are essentially in a parallel, as opposed to the more common orthogonal, orientation. This suggests an attempt to balance the slightly hydrophilic nature of the aromatic carbon atoms, as given by the RRIGS δ_i, and the favorable hydrophobic interactions between the two rings. The values of the dihedral angles for the global minimum energy conformation are given in Table II.

The global minimum structures for the area based hydration models (gradient inclusive) are less extended, as exhibited by Figures 4 and 5 (as well as Figures S8 and S9 of Supplementary Material). The lowest energy structures for the WE1 and WE2 models are very similar, with an the end–to–end C^{α} distance of 5.85 $\AA$ for both solvation models. In addition, the bend near the N–termini is stabilized by a hydrogen bond between the CO of the tyrosine residue and the NH proton of the phenylalanine residue (approximately 1.98 $\AA$). This bend is similar to the type II' β–bend of the unsolvated global minimum energy structure, although it is shifted to the Gly^2–Gly^3 backbone region. The aromatic ring separation is wider (approximately 6.48 $\AA$ for both models) than for the RRIGS global minimum structure, although the side chain orientations are similar. The values of the dihedral angles for the WE1 and WE2 global minimum structures are given in Tables III and IV respectively.

The lowest energy conformation for the OONS ASP is also similar to the WE1 and WE2 structures. In this case, the end–to–end C^{α} distance is again 5.85 $\AA$. The bending near the N–termini is again similar to a type II' β–bend along the Gly^2–Gly^3 backbone, although in this case it is stabilized by a slightly weaker hydrogen bond between the CO of the tyrosine residue and the NH proton of the phenylalanine residue (approximately 2.01 $\AA$). The 6.60 $\AA$ aromatic ring separation is also slightly larger, which may be attributed to the slightly hydrophilic character of the aromatic carbon parameters as compared to the WE1 and WE2 ASP sets. The values of the dihedral angles for the global minimum structure are given in Tables V.

The SCKS global minimum structure is even less extended, as shown in Figure 5. Although the aromatic ring separation becomes wider (8.13 $\AA$) the overall end–to–end C^{α} distance decreases to 5.80 $\AA$. In this structure, there are two stabilizing hydrogen bonds – a 1.86 $\AA$ hydrogen bond between the NH proton of the first glycine residue and the CO of the methionine residue, and a 2.02 $\AA$ hydrogen bond between the CO of the first glycine residue and the NH proton of the phenylalanine residue. This backbone structure exhibits a type II' β–bend around the Gly^3 and Phe^4 residues, which is similar to the global minimum energy conformation for unsolvated met–enkephalin. This compact structure is consistent with the relatively strong hydrophobic values of all carbon atom free energy parameters, as well as the relatively weak hydrophobic values of the oxygen and nitrogen atoms for the SCKS ASP set. The values of dihedral angles corresponding to the global minimum energy structure are given in Table VI.

In contrast, the JRF global minimum energy structure resembles a more extended conformation, with an overall end–to–end C^{α} distance of 9.56 $\AA$. The plot of this structure, given in Figure 6, shows that the residues near the N–terminus are almost fully extended, although there is slight turn near the C–terminus. This bending is stabilized by the formation of 2.10 $\AA$ hydrogen bond between the CO of the second glycine residue and the NH proton of the methionine residue. In addition, the structure displays a large 14.87 $\AA$ separation between the centroids of the Phe and Tyr aromatic rings. This can be partly attributed to the strongly hydrophilic character of the aromatic and carboxyl (carbonyl) carbons parameters for the JRF ASP set. The values of dihedral angles corresponding to the global minimum energy are given in Table VII.

The structures were further analyzed by comparing energy evaluations at corresponding global minimum solutions. This information is given in Tables VIII and IX. In all cases, excluding the SCKS model, the JRF global minimum energy structure provides that most stabilizing values for the hydration energy. However, these stabilizing hydration energies are generally offset by the relatively high value for potential energy at the JRF global minimum conformation (5.06 kcal/mole, obtained by calculating E_{TOT} - E_{HYD} from Tables VIII and IX). In fact, the high potential energy causes the JRF structure to exhibit the highest values for overall energy, excluding the case of the JRF model. Only when considering the JRF model do these stabilizing hydration free energies tend to dominate the prediction of the global minimum structure. This is evidenced by the fact that the JRF structure provides an overall energy, more than 100 kcal/mole lower than any other total energy, which can be directly attributed to the differences in hydration energy. When using the SCKS model, the only case for which the JRF conformation does not produce the most stabilizing hydration energy, the JRF structure provides the least stabilizing hydration energy. This indicates that unlike the other hydration models, the SCKS model does not provide more hydration energy stabilization for extended conformations. This agrees with the prediction of the SCKS global minimum energy structure, which exhibits the most folded conformation. The SCKS structure also closely resembles the unsolvated global minimum energy structure and it exhibits the lowest potential energy contribution, -11.63 kcal/mole, which is only .08 kcal/mole higher than the global minimum potential energy. This suggests that low potential energy conformations are not only favored but also enhanced by hydration effects for the SCKS model. Excluding the SCKS model, the other models predict relatively large hydration energies at the SCKS structure. In fact, for the RRIGS, WE1 and WE2 models, the SCKS structure produces the highest values for the hydration energies. For the OONS and JRF model, the hydration energies are only smaller than those for the RRIGS structure. This is consistent with the hydrophilic nature of the aromatic carbons for the OONS and JRF models. Specifically, because the aromatic ring separation is smallest for the RRIGS structure, the OONS and JRF hydration models tend to provide higher hydration energies for this structure. Although hydration energies for the RRIGS structure are typically high, the RRIGS model predicts a stabilizing hydration energy for this structure, second only to the JRF structure. It is this hydration energy contribution, when coupled with a relatively low potential energy (-8.59 kcal/mole), that sets the RRIGS global minimum energy structure. For the other hydration models, low potential energy contributions (-9.77, -9.75 and -9.22 kcal/mole for WE2, WE1 and OONS, respectively) seem to be more important in the prediction of relatively compact structures. In these cases the relative weighting of the hydration energy contributions does not favor extended conformations. However, these models also do not provide low hydration energies at the most compact structures, such as the SCKS global minimum energy structure. This indicates an interplay of hydration and potential energy contributions, although the prediction of relatively compact structures suggest the importance of low potential energy contributions.

5.3 Leu–enkephalin

Like met–enkephalin, leu–enkephalin (H–Tyr–Gly–Gly–Phe–Leu–OH) is an endogenous pentapeptide in which the methionine residue has been replaced by a leucine residue. The algorithmic details (i.e., global variables, α values) are identical to those outlined form met–enkephalin. In addition, the unsolvated model was reexamined using the implementation described in Section 4.2. An unsolvated global minimum energy conformation with a slightly lower ECEPP/3 energy (-9.349 kcal/mole) than the minimum energy previously reported (-9.332 kcal/mole), where α values equal to 3.5 were used[48], was identified. The new unsolvated global minimum energy structure exhibits a type II' β–bend centered around the Gly^2–Gly^3 backbone region. The previously reported structure also possessed a type II' β–bend, although the bend was shifted to the N–C' peptidic bond of Gly^3–Phe^4. A plot of the unsolvated global minimum energy structure and its corresponding dihedral angles are given in Figure S10 and Table SXIII of Supplementary Material.

Qualitatively, the results for the hydrated forms of leu–enkephalin are similar to those for met–enkephalin (see Tables SXIV to SXIX for dihedral angles and Figures S11 to S16 for plots of global minima included in Supplementary Material). The RRIGS and JRF models provide the most extended backbone conformations, with the central difference between the two models being the proximity and orientation of the two aromatic rings. The more compact structures predicted using the WE1, WE2 and OONS hydration models all share a type II' β–bend along the Gly^2–Gly^3 backbone region as a dominant conformational characteristic, although the structures of the terminal residues are different. In addition, the aromatic ring separation is larger for the OONS global minimum structure, which again may be attributed to the slightly hydrophilic nature of the aromatic carbon parameters. For the SCKS model the leu–enkephalin structure retains the type II' β–bend around the Gly^2 and Gly^3 residues. This conformation also exhibits the best correspondence to the unsolvated global minimum energy conformation, which is consistent with both the met–enkephalin results and the more hydrophobic nature of the free energy solvation parameters for the SCKS ASP set.

In general, the information obtained from energy evaluations at corresponding global minimum are very similar to those obtained for met–enkephalin (see Tables SXX and SXXI in Supplementary Material). The JRF global minimum energy structure provides that most stabilizing values for the hydration energy for all hydration models (excluding SCKS), although the potential energy contribution is large at this conformation (5.17 kcal/mole). Only when considering the JRF model do these stabilizing hydration free energies tend dominate the prediction of the global minimum structure. In fact, in all other cases, the total energy for the JRF model is the lowest of all hydration models. The other disparate result is that the JRF structure provides the least stabilizing hydration energy for the SCKS model. As previously mentioned, the SCKS model does not provide hydration energy stabilization for extended conformations. This is consistent with the highly folded global minimum energy conformation that is predicted for the SCKS model. As with met–enkephalin, the SCKS structure also resembles the unsolvated global minimum energy structure. It also exhibits the lowest potential energy contribution, -9.29 kcal/mole, which is only 0.059 kcal/mole higher than the global minimum potential energy. Excluding the SCKS model, the other models predict larger hydration energies at the SCKS structure, although these differences are smaller than for met–enkephalin. This is expected, as the conformations for leu–enkephalin for the WE1, WE2, OONS and SCKS all exhibit the same backbone β–bend. In all cases, except when considering the results of the RRIGS model, the highest hydration energy contributions are calculated using the RRIGS structure. In contrast, the RRIGS model predicts a stabilizing hydration energy for this structure, second only to the JRF structure. Therefore, the proximity of the aromatic rings is a beneficial component for the RRIGS hydration energy. For other models, the decrease in hydration energy due to an extended structure (such as the JRF structure), is obviously outweighed by the positioning of the aromatic rings in the RRIGS structure. In addition, the potential energy contribution for the RRIGS structure is relatively low, (-7.56 kcal/mole). This is similar to the potential energy contribution for the OONS global minimum energy conformation (-7.56 kcal/mole). For these two models, the interplay between hydration and potential energy contributions is evident. That is, the potential energy contribution does not determine the global minimum energy structure, which seems to be the case for the SCKS model. For example, in comparison to the OONS structure, the WE1 and WE2 models predict slightly stronger hydrogen bonds and smaller aromatic ring separation. For the WE1 and WE2 hydration models, the potential energy contributions are more stabilizing (-9.77 and -9.75 kcal/mole, respectively), and these contributions seem to be more important in the prediction of relatively compact structures.

6 Conclusions

In this paper, an approach was presented for identifying the global minimum energy of hydrated peptides. In order to determine the effects of different solvation models, the global minimum energy conformations were located and compared for a variety of hydration models. When considering solvent accessible surface area models, the MSEED algorithm was used to calculate the appropriate areas, and five ASP sets were considered. In general, hydration free energy and gradient calculations were performed at each step of local minimization. However, for the case of the JRF parameter set, the hydration free energy was added at local minima only. The RRIGS solvent accessible volume of hydration shell model was also implemented. For this method, function and gradient information was used at each step of local minimizations. The procedure was applied to the naturally occurring amino acids, and the pentapeptides met–enkephalin and leu–enkephalin.

Important comparisons can be made using the terminally blocked single residue analysis. In general, the analysis of ΔE^{POT} ($\Delta E^{POT} = E^{POT}_{ASP}$ - E^{POT}_{RRIGS} refers to the change in potential energy of an area based global minimum (E^{POT}_{ASP}) and the RRIGS global minimum (E^{POT}_{RRIGS})) indicates that the RRIGS global minima have more stabilizing potential energy contributions at their corresponding global minima. For most residues, other hydration models also predict potential energy contributions similar to the RRIGS values. In several cases, most notably for the SCKS model, structures exhibit lower potential energy contributions at the corresponding global minimum energy conformations. In contrast, when using the JRF model, the potential energy contribution for five residues (including phenylalanine and tyrosine) is more than 10 kcal/mole less stabilizing. Therefore, it may be expected that global minimum energy conformations of larger peptides modeled with the SCKS would possess the lowest potential energy contributions, whereas the JRF global minimum energy structures would exhibit relatively high potential energy contributions.

The analysis of ΔE^{HYD} ($\Delta E^{HYD} = E^{HYD}_{ASP}$ - E^{HYD}_{RRIGS} refers to the change in hydration energy of an area based global minimum (E^{HYD}_{ASP}) and the RRIGS global minimum (E^{HYD}_{RRIGS})) also indicated that the RRIGS structures often possessed the most stabilizing hydration energy components. The four exceptions are those residues possessing ringed side chains and modeled using the JRF ASP set. In fact, for these four residues (histidine, phenylalanine, tryptophan and tyrosine) the JRF hydration energy contributions were much more stabilizing at the corresponding global minimum energy structures. The other gradient inclusive area based models exhibited less stabilizing hydration energies than the RRIGS model, with the least stabilization occurring at the SCKS global minimum energy structures. The other three models, WE1, WE2, OONS possessed similar hydration energy stabilization, especially for residues containing ringed side chains. Expected trends for larger peptides include relatively low hydration energy contributions at RRIGS global minimum energy conformations when compared to the WE1, WE2, OONS and SCKS ASP set. In addition, the highest values for hydration energy would be expected for the SCKS global minimum structure. In contrast, the lowest and most stabilizing hydration energies should occur for peptides containing aromatic residues and modeled using the JRF parameter set.

Such trends were evident when comparing global minimum energy conformations for both met– and leu– enkephalin. Specifically, the most compact global minimum energy structures were predicted using the SCKS ASP set. In fact, both the met– and leu–enkephalin structures closely resembled the unsolvated global minimum energy structures. In both cases, the potential energy contributions were less than 0.1 kcal/mole higher than the corresponding global minimum potential energy. Unlike the other hydration models, the SCKS model also predicted positive values for the hydration energies with the least stabilizing hydration energies belonging to the most extended structures.

In contrast, the most extended structures were predicted using the RRIGS and JRF models. Although both models gave extended backbones, the orientation of the aromatic rings of the phenylalanine and tyrosine side chains differed. The aromatic side chains for the JRF model were widely separated, which is attributable to the strong hydrophilic character of the aromatic carbons for the JRF ASP set. Although the potential energy contribution was relatively high for the JRF structures, the large stabilizing hydration energy dominated the overall energy. When considering the RRIGS global minimum energy structures the aromatic rings exhibit an almost parallel orientation, with a small separation distance. In contrast to the JRF model, these favorable nonbonded interactions result in lower potential energy contributions.

In all cases, excluding the SCKS model, the JRF structure provides the most stabilizing hydration energies. This indicates that hydration effects (except in the SCKS case) favor extended conformations. However, only the RRIGS model predicts very stabilizing hydration energies for the RRIGS global minimum energy structure. Therefore, the RRIGS model favors both extended conformations, although the lower potential energy at the global minimum conformation indicates the importance of balancing both potential energy and hydration effects. When considering other hydration models (i.e., WE1, WE2 and OONS) the low potential energy contributions are more important in the prediction of relatively compact structures. In addition, the relative weighting of the hydration effects are smaller than for the RRIGS model.

When dealing with small, flexible peptides (i.e., met– and leu–enkephalin) it is expected that extended conformations should prevail. Therefore, the use of both the RRIGS and JRF models seems appropriate. However, it is important to notice that the JRF seems to be hydration energy dominated, whereas the RRIGS method tends to balance potential energy and hydration energy terms. This confirms earlier work for which the RRIGS method retained α–helical characteristics for global minimum energy structures of decaglycine and Ac–Ala$_4$–Pro–NHMe, while the JRF always predicted random coil conformations. In contrast, the SCKS model is potential energy dominated, and the prediction of compact is even enhanced by the hydration energy effects. The remaining hydration models provided intermediate results; that is, they did not predict the most compact conformations, but these conformations were also not extended.

Acknowledgments

The authors gratefully acknowledge financial support from the National Science Foundation, Air Force Office of Scientific Research, and the National Institutes of Health (R01 GM52032).

References

1. C. B. Anfinsen, E. Haber, M. Sela, and F. H. White, *Proc. Natl. Acad. Sci. USA*, 47, 1309 (1961).

2. C. B. Anfinsen, *Science*, 181, 223 (1973).

3. C. B. Anfinsen and H. A. Scheraga, *Adv. Protein Chem.*, 29, 205 (1975).

4. C. A. Floudas, J. L. Klepeis, and P. M. Pardalos, In *DIMACS Series in Discrete Mathematics and Theoretical Computer science*. American Mathematical Society, 1998, (to appear).

5. A. Neumaier, *SIAM Rev.*, 39, 407 (1997).

6. P. M. Pardalos, D. Shalloway, and G. Xue, editors, *Global Minimization of Nonconvex Energy Functions: Molecular Conformation and Protein Folding*, volume 23 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*. Amer. Math. Soc., 1996.

7. M. Vásquez, G. Némethy, and H. A. Scheraga, *Chemical Reviews*, 94, 2183 (1994).

8. H. A. Scheraga, In K. B. Lipkowitz and D. B. Boyd, editors, *Reviews in Computational Chemistry*, volume 3, page 73. VCH Publishers, 1992.

9. C. S. Adjiman, I. P. Androulakis, and C. A. Floudas, *Comput. Chem. Eng.*, 21, S445 (1997).

10. C. S. Adjiman, S. Dallwig, C. A. Floudas, and A. Neumaier, *Comput. Chem. Eng.* (1998), (in press).

11. C. S. Adjiman, I. P. Androulakis, and C. A. Floudas, *Comput. Chem. Eng.* (1998), (in press).

12. C. S. Adjiman, I. P. Androulakis, C. D. Maranas, and C. A. Floudas, *Comput. Chem. Eng.*, 20, S419 (1996).

13. I. P. Androulakis, C. D. Maranas, and C. A. Floudas, *J. Glob. Opt.*, 7, 337 (1995).

14. G. Némethy, K. D. Gibson, K. A. Palmer, C. N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, and H. A. Scheraga, *J. Phys. Chem.*, 96, 6472 (1992).

15. A. Dejaegere and M. Karplus, *J. Phys. Chem.*, 100, 11148 (1996).

16. P. Kollman, *Chem. Rev.*, 93, 2395 (1993).

17. T. P. Straatsma and J. A. McCammon, *Annu. Rev. Phys. Chem.*, 43, 407 (1992).

18. A. Kitao, F. Hirata, and N. Go, *J. Phys. Chem.*, 97, 10223 (1993).

19. B. Honig, K. Sharp, and A. Yang, *J. Phys. Chem.*, 97, 1101 (1993).

20. G. Perrot, B. Cheng, K. D. Gibson, K. A. Palmer J. Vila, A. Nayeem, B. Maigret, and H. A. Scheraga, *J. Comp. Chem*, 13, 1 (1992).

21. J. D. Augspurger and H. A. Scheraga, *J. Comp. Chem*, 17, 1549 (1996).

22. M. L. Connolly, *J. Appl. Cryst.*, 16, 548 (1983).

23. B. von Freyberg and W. Braun, *J. Comp. Chem.*, 14, 510 (1993).

24. F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander, and M. Scharf, *J. Comp. Chem.*, 16, 273 (1995).

25. F. Eisenhaber and P. Argos, *J. Comp. Chem.*, 14, 1272 (1993).

26. R. J. Wawak, K. D. Gibson, and H. A. Scheraga, *J. Math. Chem.*, 15, 207 (1994).

27. L. Wesson and D. Eisenberg, *Protein Science*, 1, 227 (1992).

28. R. Wolfenden, L. ANdersson, P. M. Cullis, and C. C. B. Southgate, *Biochem.*, 20, 849 (1981).

29. J. Kyte and R. F. Doolittle, *J. Mol. Bio.*, 157, 105 (1982).

30. R. Friedman K. A. Sharp, A. Nicholls and B. Honig, *Science*, 252, 106 (1991).

31. A. H. Juffer, F. Eisenhaber, S. J. Hubbard, D. Walther, and P. Argos, *Protein Sci.*, 4, 2499 (1995).

32. A. Ben-Naim and R. M. Mazo, *J. Phys. Chem.*, 97, 10829 (1993).

33. T. Ooi, M. Oobatake, G. Némethy, and H. A. Scheraga, *Proc. Natl. Acad. Sci. USA*, 84, 3086 (1987).

34. C. A. Schiffer, J. W. Caldwell, P. A. Kollman, and R. M. Stroud, *Mol. Sim.*, 10, 121 (1993).

35. B. Brooks, R. Bruccoleri, B. Olafson, D. States, S. Swaminathan, and M. Karplus, *J. Comp. Chem.*, 4, 187 (1983).

36. R. L. Williams, J. Vila, G. Perrot, and H. A. Scheraga, *Proteins*, 14, 110 (1992).

37. A. J. Hopfinger, *Macromolecules*, 4, 731 (1971).

38. Y. K. Kang, G. Némethy, and H. A. Scheraga, *J. Phys. Chem.*, 91, 4105 (1987).

39. Y. K. Kang, G. Némethy, and H. A. Scheraga, *J. Phys. Chem.*, 91, 4109 (1987).

40. Y. K. Kang, G. Némethy, and H. A. Scheraga, *J. Phys. Chem.*, 91, 4118 (1987).

41. Y. K. Kang, K. D. Gibson, G. Némethy, and H. A. Scheraga, *J. Phys. Chem.*, 92 (1988).

42. C.A. Floudas, In L.T. Biegler, T.F. Coleman, A.R. Conn, and F.N. Santosa, editors, *Large Scale Optimization with Applications, Part II: Optimal Design and Control*, volume 93, page 129. IMA Volumes in Mathematics and its Applications, Springer–Verlag, 1997.

43. C. D. Maranas and C. A. Floudas, *Annals of Operations Research*, 42, 85 (1993).

44. C. D. Maranas and C. A. Floudas, *J. Chem. Phys.*, 97, 7667 (1992).

45. C. D. Maranas and C. A. Floudas, *J. Chem. Phys.*, 100, 1247 (1994).

46. C. D. Maranas and C. A. Floudas, *J. Glob. Opt.*, 4, 135 (1994).

47. C. D. Maranas, I. P. Androulakis, and C. A. Floudas, In *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, volume 23, page 133. American Mathematical Society, 1996.

48. I. P. Androulakis, C. D. Maranas, and C. A. Floudas, *J. Glob. Opt.*, 11, 1 (1997).

49. C. S. Adjiman and C. A. Floudas, *J. Glob. Opt.*, 9, 23 (1996).

50. H.A. Scheraga, *PACK: Programs for Packing Polypeptide Chains*, 1996, online documentation.

51. P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, *NPSOL 4.0 User's Guide*, Systems Optimization Laboratory, Dept. of Operations Research, Stanford University, CA., 1986.

52. T. Noguti and N. Go, *J. Phys. Soc. Japan*, 52, 3685 (1983).

53. V. Madison and K. D. Kopple, *J. Am. Chem. Soc.*, 102, 4855 (1980).

54. Z. Li and H. A. Scheraga, *J. Mol. Struct. (Theochem.)*, 179, 333 (1988).

55. W. H. Graham, E. S. Carter II, and R. P. Hicks, *Biopolymers*, 32, 1755 (1992).

56. S. K. Burley and G. A. Petsko, *Science*, 229, 23 (1985).

Table I: Free energy density of solvation parameters for the ASP set employed with the solvent accessible surface area model. The first column describes the atom type, whereas the remaining columns provide the solvation parameters in cal/(mol $\AA^2$) for the corresponding ASP set.

Atom Type	**WE1**	**WE2**	**OONS**	**SCKS**	**JRF**
C aliphatic	12.0	4.0	8.0	32.5	216.0
C carboxyl, carbonyl	12.0	4.0	427.0	32.5	-732.0
C aromatic	12.0	4.0	-8.0	32.5	-678.0
N noncharged	-116.0	-113.0	-132.0	-17.5	-312.0
N charged	-186.0	-169.0	-132.0	-217.5	-312.0
O carboxyl, carbonyl	-116.0	-113.0	-38.0	-17.5	-262.0
O hydroxyl	-116.0	-113.0	-172.0	-17.5	-910.0
O charged	-175.0	-169.0	-38.0	-280.0	-910.0
S all	-18.0	-17.0	-21.0	-9.0	-281.0

Table II: Dihedral angles at the global minimum energy conformation of met–enkephalin, using the RRIGS model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-168.32	-30.81	178.52	-173.58	-101.26	18.83	
Gly	78.83	-86.96	182.73				
Gly	162.94	91.72	172.83				
Phe	-150.72	162.32	181.50	66.66	92.68		
Met	-77.80	106.79	181.63	-67.82	178.91	180.01	-60.01

Table III: Dihedral angles at the global minimum energy conformation of met–enkephalin, using the WE1 model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-162.65	-43.34	-177.43	-173.76	-90.62	2.61	
Gly	66.15	-86.62	172.92				
Gly	-152.31	32.40	-178.49				
Phe	-157.59	154.87	179.36	52.02	-96.19		
Met	-90.62	128.89	-179.18	-169.29	176.88	180.14	-59.99

Table IV: Dihedral angles at the global minimum energy conformation of met–enkephalin, using the WE2 model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-162.70	-43.23	-177.47	-173.94	-90.83	2.63	
Gly	66.15	-86.59	173.03				
Gly	-152.49	32.41	-178.55				
Phe	-157.84	154.97	179.26	52.12	-96.11		
Met	-89.96	129.19	-179.17	-169.47	176.75	180.13	-59.99

Table V: Dihedral angles at the global minimum energy conformation of met–enkephalin, using the OONS model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-166.11	-50.84	-176.25	-188.97	-102.81	2.45	
Gly	63.86	-86.04	175.39				
Gly	-151.94	33.86	-178.80				
Phe	-159.47	153.41	179.46	50.93	-96.43		
Met	-79.75	148.31	-178.93	-68.16	181.45	178.08	59.70

Table VI: Dihedral angles at the global minimum energy conformation of met–enkephalin, using the SCKS model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-82.91	154.09	-176.27	-172.88	79.47	-166.08	
Gly	-151.61	81.91	168.71				
Gly	84.09	-72.41	-169.54				
Phe	-137.07	18.52	-173.06	57.94	-86.04		
Met	-162.71	158.63	-179.76	51.94	173.67	179.21	-58.18

Table VII: Dihedral angles at the global minimum energy conformation of met–enkephalin, using the JRF model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-84.96	160.74	179.09	-59.83	100.80	-179.29	
Gly	-160.26	151.83	-177.53				
Gly	159.50	-157.94	178.71				
Phe	-76.55	76.23	-178.05	-61.87	108.68		
Met	-132.90	147.47	-179.83	-65.17	-175.99	-84.91	59.38

Table VIII: Comparison of hydration energies for met–enkephalin. The first column refers to the hydration model used in the function evaluations, which are performed at the global solutions for the hydration model listed in the second column. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}. The last column provides the heavy atom root mean squared deviation between the global minimum energy structures of the hydration models listed in the first two columns.

	Global of	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$	(RMSD)
RRIGS	RRIGS	-50.01	-41.42	21.84	-31.46	1.02	0.00
	WE1	-47.87	-38.12	22.09	-32.61	0.78	2.83
	WE2	-47.91	-38.14	22.09	-32.63	0.76	2.83
	OONS	-47.17	-37.95	22.25	-32.13	0.66	2.66
	SCKS	-47.24	-35.61	21.47	-35.40	2.30	4.04
	JRF	-41.63	-46.69	23.29	-19.13	0.90	4.83
WE1	RRIGS	-26.60	-18.00	21.84	-31.46	1.02	2.83
	WE1	-30.31	-20.56	22.09	-32.61	0.78	0.00
	WE2	-30.31	-20.53	22.09	-32.63	0.76	0.01
	OONS	-29.01	-19.79	22.25	-32.13	0.66	0.80
	SCKS	-27.80	-16.17	21.47	-35.40	2.30	3.33
	JRF	-19.49	-24.55	23.29	-19.13	0.90	4.33
WE2	RRIGS	-29.87	-21.27	21.84	-31.46	1.02	2.83
	WE1	-33.26	-23.52	22.09	-32.61	0.78	0.01
	WE2	-33.27	-23.49	22.09	-32.63	0.76	0.00
	OONS	-32.01	-22.79	22.25	-32.13	0.66	0.80
	SCKS	-30.77	-19.15	21.47	-35.40	2.30	3.33
	JRF	-22.93	-27.99	23.29	-19.13	0.90	4.32

Table IX: Comparison of hydration energies for met–enkephalin. The first column refers to the hydration model used in the function evaluations, which are performed at the global solutions for the hydration model listed in the second column. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}. The last column provides the heavy atom root mean squared deviation between the global minimum energy structures of the hydration models listed in the first two columns.

	Global of	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$	(RMSD)
OONS	RRIGS	-24.18	-15.59	21.84	-31.46	1.02	2.66
	WE1	-31.08	-21.33	22.09	-32.61	0.78	0.80
	WE2	-31.09	-21.31	22.09	-32.63	0.76	0.80
	OONS	-31.45	-22.23	22.25	-32.13	0.66	0.00
	SCKS	-29.57	-17.95	21.47	-35.40	2.30	3.38
	JRF	-19.60	-24.66	23.29	-19.13	0.90	4.12
SCKS	RRIGS	3.43	12.02	21.84	-31.46	1.02	4.04
	WE1	0.90	10.65	22.09	-32.61	0.78	3.33
	WE2	0.89	10.67	22.09	-32.63	0.76	3.33
	OONS	1.66	10.88	22.25	-32.13	0.66	3.38
	SCKS	-0.62	11.00	21.47	-35.40	2.30	0.00
	JRF	17.44	12.38	23.29	-19.13	0.90	3.78
JRF	RRIGS	-139.36	-130.77	21.84	-31.46	1.02	4.83
	WE1	-180.59	-170.84	22.09	-32.61	0.78	4.33
	WE2	-180.57	-170.79	22.09	-32.63	0.76	4.32
	OONS	-181.70	-172.48	22.25	-32.13	0.66	4.12
	SCKS	-171.67	-160.04	21.47	-35.40	2.30	3.78
	JRF	-283.76	-288.82	23.29	-19.13	0.90	0.00

List of figures

Figure 1: Potential energy terms in ECEPP/3 force field. r_{ij} refers to the interatomic distance of the atomic pair (ij). Q_i and Q_j are dipole parameters for the respective atoms, in which the dielectric constant of 2 has been incorporated. F_{ij} is set equal to 0.5 for 1–4 interactions and 1.0 for 1–5 and higher interactions. A_{ij}, C_{ij}, A'_{ij} and B_{ij} are nonbonded and hydrogen bonded parameters specific to the atomic pair. $E_{o,k}$ and $E_{o,l}$ are parameters corresponding to torsional barrier energies for a given dihedral angle. θ_k represents any dihedral angle. c_k takes the values -1,1, and n_k refers to the symmetry type for the particular dihedral angle. The cystine loop–closing term is calculated as a penalty term of three distances involved in loop–closing, where r_{il} represents the actual distance and r_{io} represents the required distance. B_i, the penalty parameter, is set equal to 100. Finally, E_p is a fixed internal energy that is added for each proline residue in the protein.

Figure 2: Interface for global optimization.

Figure 3: Plot of met–enkephalin conformation. Global minimum energy of -50.01 kcal/mole using the RRIGS model for hydration.

Figure 4: Plot of met–enkephalin conformation. Global minimum energy of -30.31 kcal/mole using the WE1 model for hydration.

Figure 5: Plot of met–enkephalin conformation. Global minimum energy of -0.62 kcal/mole using the SCKS model for hydration.

Figure 6: Plot of met–enkephalin conformation. Global minimum energy of -283.76 kcal/mole using the JRF model for hydration.

$$
\begin{aligned}
E = & \sum_{(i,j)\in ES} 332.0\frac{q_i\, q_j}{D r_{ij}} && \text{(Electrostatic)}\\
& + \sum_{(i,j)\in NB} F\, \frac{A}{r_{ij}^{12}} - \frac{C}{r_{ij}^{6}} && \text{(Nonbonded)}\\
& + \sum_{(hx)\in HX} F\, \frac{A'}{r_{hx}^{12}} - \frac{B}{r_{hx}^{10}} && \text{(Hydrogen bonded)}\\
& + \sum_{k\in TOR} \left(\frac{E_0}{2}\right)(1 \pm \cos n_k \theta_k) && \text{(Torsional)}\\
& + \sum_{l\in LOOP} B_L \sum_{il=1}^{il=3} (r_{il} - r_{io})^2 && \text{(Cystine Loop-Closing)}\\
& + \sum_{l\in LOOP} A_L\, (r_{4l} - r_{4o})^2 && \text{(Cystine Torsional)}
\end{aligned}
$$

Figure 1:

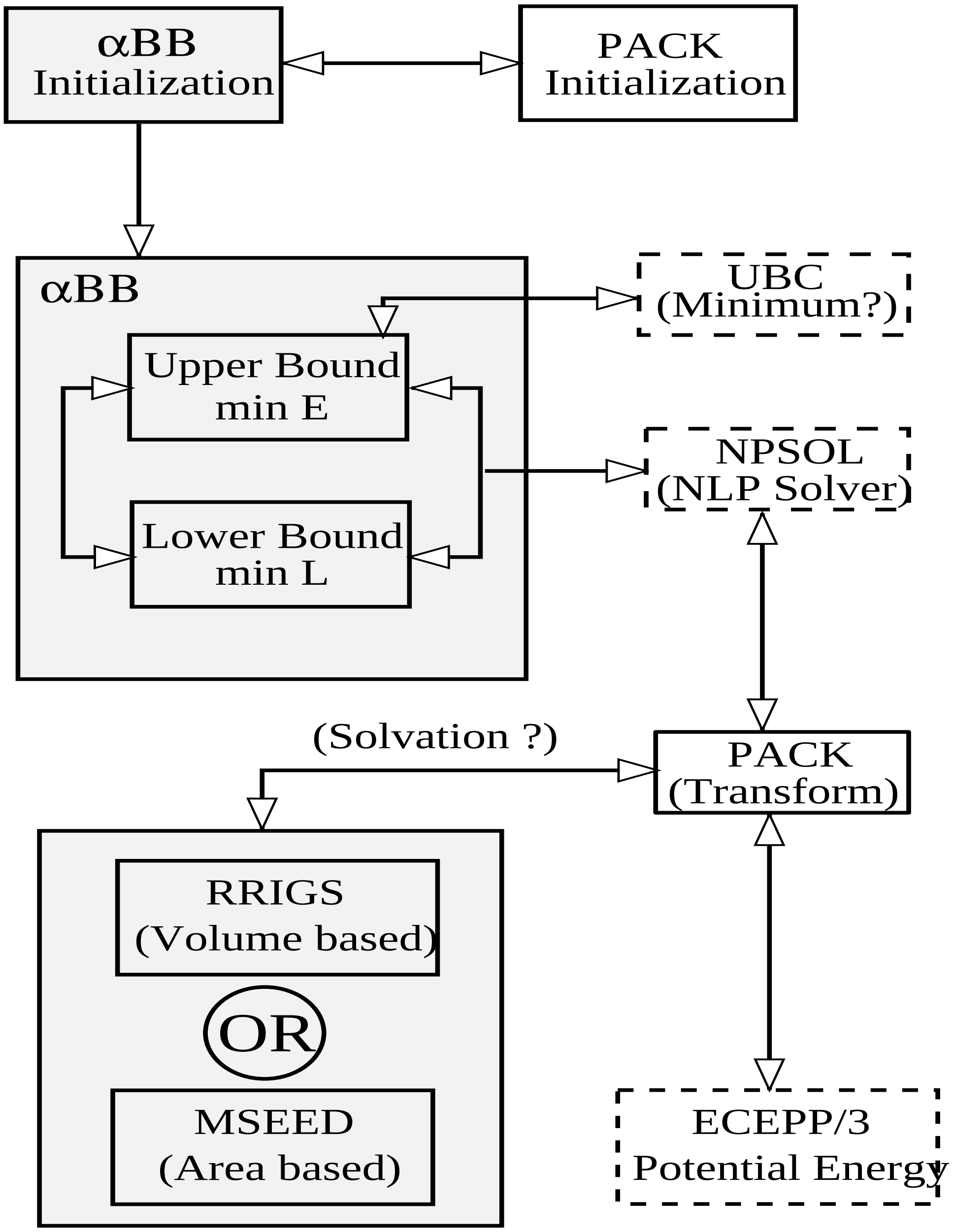

αBB
Initialization
PACK
Initialization
αBB
UBC
(Minimum?)
Upper Bound
min E
NPSOL
(NLP Solver)
Lower Bound
min L
(Solvation ?)
PACK
(Transform)
RRIGS
(Volume based)
OR
MSEED
(Area based)
ECEPP/3
Potential Energy

Figure 2:

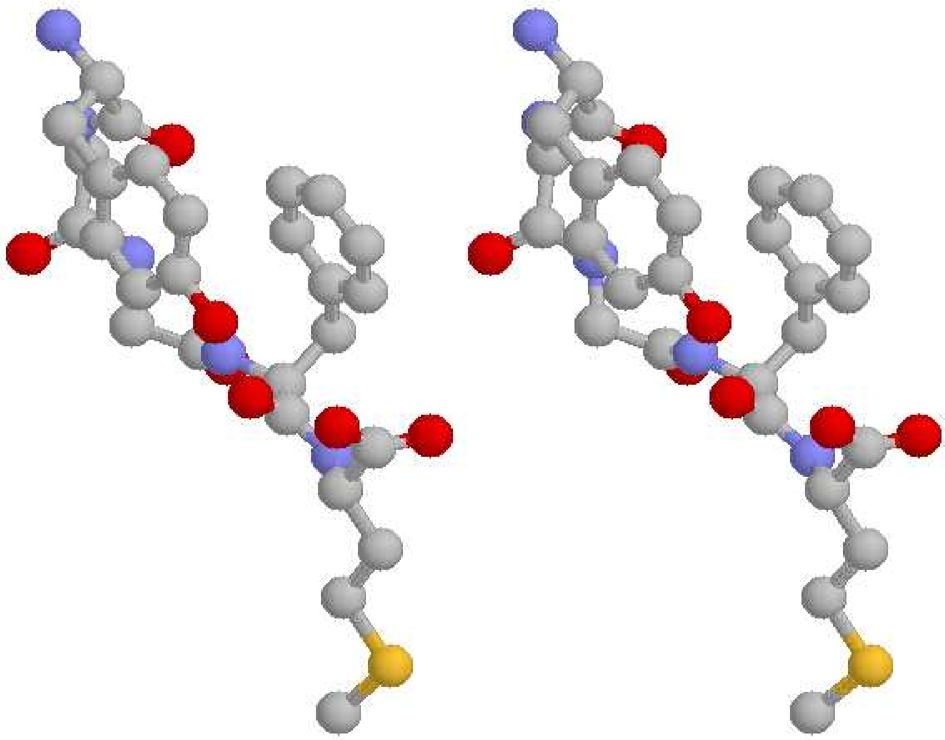


Figure 3:

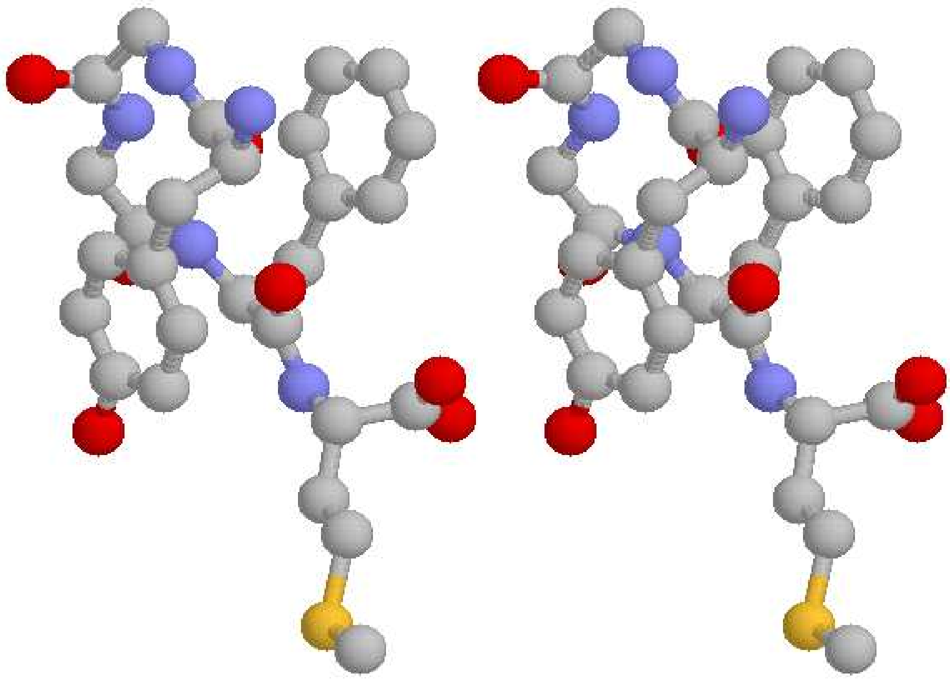


Figure 4:

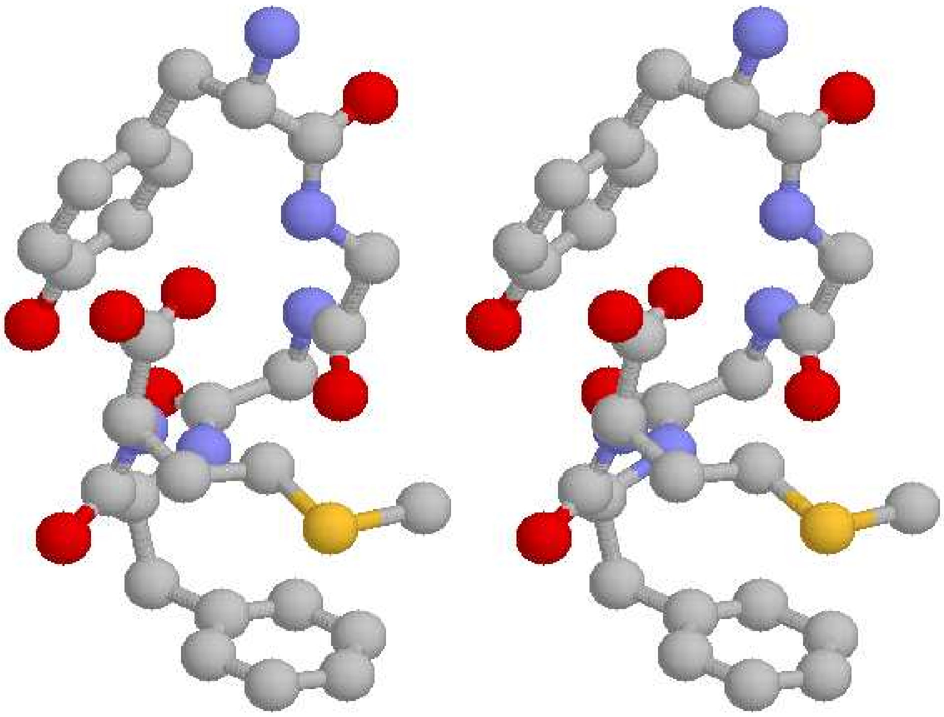


Figure 5:

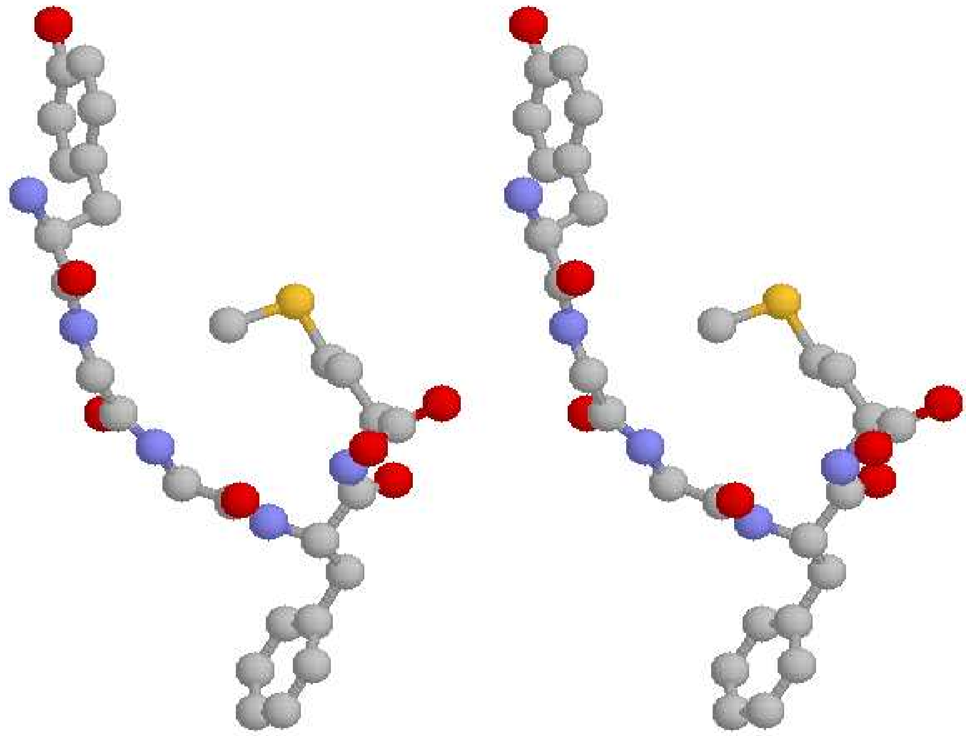


Figure 6:

Supplementary Material

A comparative study of global minimum energy conformations of hydrated peptides

J. L. Klepeis and C. A. Floudas
Department of Chemical Engineering
Princeton University
Princeton, N.J. 08544-5263

Tables SI to SXXI
Figures S1 to S16

Table SI: Free energy density of solvation parameters employed in the RRIGS model. The second column provides the solvation parameters in cal/(mol $\mathring{A}^2$), and the last two columns correspond to the van der Waals and hydration radii ($\mathring{A}$), respectively.

Atom Type	δ	R^v	R^h
H hydroxyl, amino	-10.35	1.415	4.17
H acid	-3.206	1.415	4.17
H amide	-7.714	1.415	4.17
H thiol	2.709	1.415	4.17
C aliphatic CH_3	1.319	2.125	5.35
C aliphatic CH_2	0.2374	2.225	5.35
C aliphatic CH	-1.271	2.375	5.35
C other aliphatic	-2.297	2.060	5.35
C cyclic CH	0.2890	2.375	5.35
C aromatic CH	-0.2137	2.100	5.35
C aromatic CR	-1.713	1.850	5.35
C branched aromatic C	-1.910	1.850	5.35
C aromatic COH	-0.6063	1.850	5.35
C carbonyl	2.696	1.870	5.35
N primary amine	-1.149	1.755	5.05
N secondary amine	-10.28	1.755	5.05
N aromatic	-10.48	1.755	5.05
N amide	-7.332	1.755	5.05
O hydroxyl, ether	-7.396	1.620	4.95
O acid, ester	0.07897	1.620	4.95
O ketone, carbonyl	-15.70	1.560	4.95
O acid, amide carbonyl	-15.56	1.560	4.95
S thiol, disulfide	-4.706	2.075	5.37

Table SII: Global minimum energies of terminally blocked peptides using the RRIGS solvation model and ASP set. The amino end group is specified as N–Acetyl–amino; the carboxyl end group is specified as Carboxyl–$CONHCH_3$. The total number of dihedral angles is indicated in the column headed by # DA. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}.

Residue	**# DA**	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$
Gly	6	-22.46	-16.14	-3.71	-2.62	0.01
Ala	7	-20.82	-15.64	-3.92	-1.28	0.02
Cys	7	-23.51	-17.67	-4.66	-1.21	0.03
His	8	-34.47	-25.57	-6.78	-2.21	0.09
Phe	8	-24.72	-16.55	-7.23	-0.94	0.00
Ser	8	-28.32	-20.47	-5.40	-2.49	0.04
Trp	8	-31.48	-21.92	-8.99	-0.59	0.02
Asn	9	-49.07	-26.47	-5.16	-17.47	0.03
Asp	9	-39.96	-20.94	-6.29	-12.74	0.01
Thr	9	-29.18	-19.59	-5.74	-4.19	0.34
Tyr	9	-30.11	-21.90	-6.65	-1.57	0.01
Val	9	-18.92	-14.74	-3.11	-1.16	0.09
Gln	10	-46.49	-27.70	-5.38	-13.49	0.08
Glu	10	-36.11	-20.92	-5.42	-9.85	0.08
Ile	10	-17.11	-14.57	-2.80	-0.52	0.78
Leu	10	-20.22	-14.53	-4.16	-1.88	0.35
Met	10	-23.93	-17.02	-4.62	-2.40	0.11
Lys	11	-28.15	-20.17	-5.91	-2.17	0.10
Arg	13	-63.84	-32.38	-6.21	-25.36	0.11

Table SIII: Global minimum energies of terminally blocked peptides using the WE1 ASP set. The amino end group is specified as N–Acetyl–amino; the carboxyl end group is specified as Carboxyl–$CONHCH_3$. The total number of dihedral angles is indicated in the column headed by # DA. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}.

Residue	**# DA**	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$
Gly	6	-12.93	-6.68	-3.71	-2.54	0.00
Ala	7	-10.98	-6.59	-3.91	-0.48	0.00
Cys	7	-12.80	-8.19	-4.15	-0.47	0.01
His	8	-19.12	-11.52	-6.59	-1.02	0.02
Phe	8	-12.51	-4.38	-7.22	-0.94	0.04
Ser	8	-16.46	-11.24	-3.74	-1.48	0.00
Trp	8	-16.54	-7.03	-8.98	-0.56	0.04
Asn	9	-38.52	-16.03	-6.16	-16.34	0.01
Asp	9	-32.68	-14.21	-5.65	-12.83	0.01
Thr	9	-16.61	-7.35	-5.55	-3.98	0.26
Tyr	9	-15.46	-8.66	-5.79	-1.03	0.03
Val	9	-8.89	-5.26	-2.94	-0.77	0.07
Gln	10	-34.91	-17.36	-5.01	-12.67	0.14
Glu	10	-28.88	-15.08	-5.00	-8.92	0.13
Ile	10	-6.80	-4.30	-2.75	-0.51	0.76
Leu	10	-9.92	-5.22	-3.79	-1.07	0.16
Met	10	-12.62	-6.50	-4.63	-1.64	0.14
Lys	11	-19.56	-13.06	-5.26	-1.31	0.07
Arg	13	-51.46	-20.51	-6.39	-24.62	0.07

Table SIV: Global minimum energies of terminally blocked peptides using the WE2 ASP set. The amino end group is specified as N–Acetyl–amino; the carboxyl end group is specified as Carboxyl–$CONHCH_3$. The total number of dihedral angles is indicated in the column headed by # DA. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}.

Residue	**# DA**	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$
Gly	6	-14.46	-8.20	-3.71	-2.55	0.00
Ala	7	-12.73	-8.33	-3.92	-0.48	0.00
Cys	7	-14.18	-9.54	-4.17	-0.48	0.01
His	8	-20.97	-13.33	-6.64	-1.03	0.02
Phe	8	-15.10	-6.96	-7.22	-0.94	0.03
Ser	8	-17.88	-12.58	-3.79	-1.51	0.00
Trp	8	-19.11	-9.59	-8.98	-0.56	0.03
Asn	9	-39.76	-17.26	-6.16	-16.35	0.01
Asp	9	-34.03	-15.52	-5.66	-12.84	0.00
Thr	9	-18.48	-9.14	-5.60	-4.01	0.27
Tyr	9	-19.54	-11.36	-6.64	-1.57	0.03
Val	9	-11.05	-7.37	-2.99	-0.76	0.07
Gln	10	-36.14	-17.88	-6.01	-12.33	0.09
Glu	10	-30.37	-16.56	-5.02	-8.92	0.13
Ile	10	-9.19	-6.68	-2.76	-0.51	0.76
Leu	10	-12.28	-7.58	-3.79	-1.07	0.15
Met	10	-14.73	-8.60	-4.64	-1.64	0.14
Lys	11	-21.46	-14.90	-5.31	-1.32	0.07
Arg	13	-52.94	-21.98	-6.41	-24.62	0.07

Table SV: Global minimum energies of terminally blocked peptides using the OONS ASP set. The amino end group is specified as N–Acetyl–amino; the carboxyl end group is specified as Carboxyl–$CONHCH_3$. The total number of dihedral angles is indicated in the column headed by # DA. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}.

Residue	**# DA**	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$
Gly	6	-7.52	-1.24	-3.72	-2.58	0.01
Ala	7	-5.59	-1.18	-3.93	-0.48	0.01
Cys	7	-8.24	-2.61	-4.64	-1.02	0.03
His	8	-15.70	-8.65	-6.07	-0.99	0.01
Phe	8	-10.49	-2.07	-7.55	-0.88	0.00
Ser	8	-13.32	-5.53	-5.43	-2.39	0.02
Trp	8	-15.07	-5.54	-8.97	-0.58	0.02
Asn	9	-30.84	-8.31	-6.14	-16.41	0.02
Asp	9	-27.55	-9.07	-5.80	-12.69	0.01
Thr	9	-13.84	-4.36	-5.72	-4.09	0.33
Tyr	9	-17.64	-9.18	-6.94	-1.52	0.00
Val	9	-4.58	-0.51	-3.04	-1.11	0.08
Gln	10	-26.80	-8.31	-5.22	-13.43	0.15
Glu	10	-23.36	-8.89	-5.95	-8.60	0.08
Ile	10	-2.39	0.12	-2.74	-0.51	0.74
Leu	10	-5.50	0.18	-4.12	-1.76	0.20
Met	10	-7.94	-1.83	-4.63	-1.63	0.14
Lys	11	-14.03	-6.20	-5.89	-2.04	0.10
Arg	13	-46.79	-15.77	-6.52	-24.58	0.08

Table SVI: Global minimum energies of terminally blocked peptides using the SCKS ASP set. The amino end group is specified as N–Acetyl–amino; the carboxyl end group is specified as Carboxyl–$CONHCH_3$. The total number of dihedral angles is indicated in the column headed by # DA. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}.

Residue	**# DA**	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$
Gly	6	-0.50	5.82	-3.70	-2.62	0.00
Ala	7	2.01	7.18	-3.91	-1.27	0.01
Cys	7	-0.50	5.33	-4.66	-1.19	0.02
His	8	-1.07	7.29	-7.53	-0.86	0.04
Phe	8	1.69	10.11	-7.56	-0.87	0.02
Ser	8	-1.81	6.04	-5.39	-2.48	0.01
Trp	8	0.39	9.92	-8.98	-0.58	0.03
Asn	9	-18.37	4.57	-6.33	-16.63	0.02
Asp	9	-14.09	5.25	-6.49	-12.85	0.00
Thr	9	-2.36	7.22	-5.72	-4.18	0.32
Tyr	9	0.07	8.54	-6.95	-1.53	0.02
Val	9	4.51	8.65	-3.10	-1.13	0.09
Gln	10	-13.53	5.24	-5.38	-13.47	0.08
Glu	10	-9.27	5.93	-5.41	-9.87	0.08
Ile	10	6.81	9.34	-2.77	-0.51	0.75
Leu	10	3.99	9.69	-4.11	-1.80	0.20
Met	10	1.75	8.65	-4.62	-2.38	0.10
Lys	11	-0.40	7.58	-5.90	-2.15	0.07
Arg	13	-26.07	5.26	-7.11	-24.33	0.11

Table SVII: Global minimum energies of terminally blocked peptides using the JRF ASP set. The amino end group is specified as N–Acetyl–amino; the carboxyl end group is specified as Carboxyl–$CONHCH_3$. The total number of dihedral angles is indicated in the column headed by # DA. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}.

Residue	**# DA**	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$
Gly	6	15.99	14.29	3.01	-1.54	0.23
Ala	7	29.71	24.81	2.36	-0.24	2.78
Cys	7	0.26	-4.22	2.51	-0.21	2.18
His	8	-50.22	-42.99	-6.87	-0.46	0.10
Phe	8	-83.47	-86.53	0.12	-0.82	3.76
Ser	8	-5.72	-9.62	2.68	-1.39	2.61
Trp	8	-105.88	-98.00	-7.91	0.03	0.00
Asn	9	-20.76	-20.86	8.13	-16.55	8.52
Asp	9	-41.14	-31.91	2.31	-12.95	1.41
Thr	9	6.56	2.82	0.31	-2.16	5.59
Tyr	9	-102.43	-105.52	1.02	-1.43	3.50
Val	9	46.54	39.84	2.53	-0.86	5.03
Gln	10	-13.89	-6.55	1.93	-12.69	3.42
Glu	10	-33.55	-19.61	-5.18	-8.93	0.17
Ile	10	53.61	56.15	-2.80	-0.52	0.78
Leu	10	47.62	29.61	8.37	-0.54	10.18
Met	10	26.33	21.35	2.10	-1.61	4.49
Lys	11	26.65	22.85	0.40	-1.45	4.85
Arg	13	-34.88	-4.57	-6.12	-24.39	0.20

Table SVIII: Local minimum energies of terminally blocked peptides using the JRF ASP set with constrained ω bounds [160,200]. The amino end group is specified as N–Acetyl–amino; the carboxyl end group is specified as Carboxyl–$CONHCH_3$. The total number of dihedral angles is indicated in the column headed by # DA. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}.

Residue	**# DA**	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$
Ala	7	32.97	37.44	-4.00	-0.47	0.00
Cys	7	2.34	7.27	-5.17	0.24	0.00
Ser	8	-4.75	-11.02	0.69	-0.28	5.86
Asn	9	-19.13	3.43	-6.13	-16.43	0.00
Asp	9	-39.35	-20.76	-6.13	-12.47	0.01
Val	9	46.71	50.62	-3.26	-0.77	0.12
Gln	10	-13.51	4.20	-5.17	-12.74	0.20
Leu	10	49.68	41.57	1.58	-0.66	7.19
Met	10	27.04	32.32	-4.41	-1.65	0.78
Lys	11	26.96	33.71	-5.51	-1.34	0.10

Table SIX: Comparison of unsolvated components for all accessible area based ASP set global minima and the RRIGS global minima (of terminally blocked peptides). $\Delta E^{POT} = E^{POT}_{ASP} - E^{POT}_{RRIGS}$ at corresponding global minima, for the ASP set as listed.

Residue	WE1	WE2	OONS	SCKS	JRF
Gly	0.07	0.06	0.04	0.00	8.02
Ala	0.79	0.78	0.77	0.00	0.71
Cys	1.22	1.20	0.21	0.01	0.91
His	1.31	1.25	1.85	0.55	1.67
Phe	0.04	0.04	-0.26	-0.25	11.23
Ser	2.64	2.56	0.06	0.00	14.13
Trp	0.04	0.03	0.02	0.02	1.67
Asn	0.10	0.10	0.06	-0.35	0.04
Asp	0.55	0.52	0.55	-0.32	0.44
Thr	0.32	0.25	0.11	0.01	13.32
Tyr	1.42	0.03	-0.24	-0.25	11.30
Val	0.55	0.50	0.11	0.04	0.27
Gln	1.23	0.53	0.29	0.01	1.07
Glu	1.39	1.37	0.71	-0.02	1.24
Ile	0.04	0.03	0.03	0.01	0.00
Leu	1.00	0.99	0.02	-0.01	13.80
Met	0.79	0.78	0.79	0.01	1.63
Lys	1.48	1.43	0.15	0.01	1.23
Arg	0.52	0.50	0.44	0.13	1.15

Table SX: Comparison of unsolvated components for all accessible area based ASP set global minima and the RRIGS global minima (of terminally blocked peptides). $\Delta E^{HYD} = E^{HYD}_{ASP}$ - E^{HYD}_{RRIGS} at corresponding global minima, for the ASP set as listed.

Residue	WE1	WE2	OONS	SCKS	JRF
Gly	9.46	7.94	14.90	21.96	30.43
Ala	9.05	7.31	14.46	22.82	53.08
Cys	9.48	8.13	15.06	23.00	24.94
His	14.05	12.25	16.92	32.86	-17.42
Phe	12.17	9.59	14.48	26.66	-69.98
Ser	9.23	7.89	14.93	26.51	9.45
Trp	14.89	12.33	16.39	31.85	-76.07
Asn	10.44	9.21	18.16	31.04	29.90
Asp	6.72	5.41	11.87	26.19	0.17
Thr	12.25	10.45	15.23	26.81	22.42
Tyr	13.24	10.54	12.72	30.44	-83.62
Val	9.48	7.37	14.23	23.39	65.36
Gln	10.35	9.82	19.40	32.95	31.90
Glu	5.84	4.37	12.03	26.85	1.32
Ile	10.26	7.88	14.69	23.91	70.72
Leu	9.31	6.95	14.71	24.22	56.10
Met	10.53	8.43	15.20	25.68	49.34
Lys	7.11	5.27	13.97	27.75	53.88
Arg	11.87	10.40	16.61	37.64	27.80

Table SXI: Approximate dihedral angles and nomenclature for ϕ–ψ regions.

Conformer	**ϕ,ψ**	**Protein structure**
C_5	-150, 150	β–sheet
P_{II}	-80, 150	polyproline II
C_7	-80, 80	γ–turn
α_R	-80, -50	α–helix (right)

Table SXII: Distribution of global minima for terminally blocked amino acids using solvation model listed in first column.

Model	**C_5**	**P_{II}**	**C_7**	α_R	Other
RRIGS	5	0	14	0	0
WE1	13	1	2	2	1
WE2	13	1	2	2	1
OONS	6	2	8	2	1
SCKS	6	0	11	2	0
JRF	9	2	0	7	1

Table SXIII: Dihedral angles at the global minimum energy conformation of unsolvated leu–enkephalin.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-163.07	-42.29	182.31	-174.79	90.16	-177.26	
Gly	65.90	-88.33	174.19				
Gly	-150.83	31.94	181.29				
Phe	-158.73	157.23	178.02	53.24	84.40		
Leu	-77.83	123.31	181.31	-179.81	64.47	172.58	179.43

Table SXIV: Dihedral angles at the global minimum energy conformation of leu–enkephalin, using the RRIGS model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-168.37	-30.66	178.49	-173.40	78.69	-161.13	
Gly	78.92	-87.17	-177.32				
Gly	163.21	91.51	172.72				
Phe	-150.66	161.54	-178.44	66.75	-86.84		
Leu	-75.45	105.32	-178.26	179.51	63.84	172.22	179.31

Table SXV: Dihedral angles at the global minimum energy conformation of leu–enkephalin, using the WE1 model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-162.77	-43.63	-177.55	-174.64	88.60	182.98	
Gly	66.15	-86.15	173.14				
Gly	-152.77	32.53	181.90				
Phe	-158.33	156.18	179.06	51.73	83.41		
Leu	-86.58	124.93	-179.00	182.65	69.00	54.85	-59.77

Table SXVI: Dihedral angles at the global minimum energy conformation of leu–enkephalin, using the WE2 model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-162.81	-43.49	-177.59	-174.74	88.48	182.98	
Gly	66.15	-86.18	173.23				
Gly	-152.91	32.54	181.81				
Phe	-158.47	156.24	179.03	51.81	83.50		
Leu	-86.17	125.26	-178.99	182.37	68.61	54.60	-59.84

Table SXVII: Dihedral angles at the global minimum energy conformation of leu–enkephalin, using the OONS model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-166.40	-51.73	-175.64	-189.82	75.41	182.41	
Gly	63.68	-85.81	175.24				
Gly	-152.32	33.95	181.27				
Phe	-159.65	153.94	-180.61	51.12	83.54		
Leu	-84.06	148.49	-179.04	-63.09	160.85	59.05	62.99

Table SXVIII: Dihedral angles at the global minimum energy conformation of leu–enkephalin, using the SCKS model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-162.84	-43.09	182.46	-174.41	-89.51	2.80	
Gly	65.94	-88.13	173.67				
Gly	-150.11	31.76	181.73				
Phe	-157.73	156.45	178.12	53.21	83.97		
Leu	-80.71	123.90	181.21	-178.83	65.57	-186.90	-180.38

Table SXIX: Dihedral angles at the global minimum energy conformation of leu–enkephalin, using the JRF model for hydration.

	ϕ	ψ	ω	χ_1	χ_2	χ_3	χ_4
Tyr	-84.88	160.00	178.30	-60.54	100.49	-179.22	
Gly	-160.78	140.99	-178.01				
Gly	144.15	-152.83	177.03				
Phe	-79.95	71.30	-176.06	-60.97	108.26		
Leu	-83.98	138.62	-179.24	-53.91	176.56	-178.84	69.81

Table SXX: Comparison of hydration energies for leu–enkephalin. The first column refers to the hydration model used in the function evaluations, which are performed at the global solutions for the hydration model listed in the second column. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}. The last column provides the heavy atom root mean squared deviation between the global minimum energy structures of the hydration models listed in the first two columns.

	Global of	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$	(RMSD)
RRIGS	RRIGS	-46.56	-39.00	22.31	-30.95	1.07	0.00
	WE1	-44.70	-35.68	22.66	-32.43	0.75	2.56
	WE2	-44.75	-35.69	22.67	-32.45	0.72	2.55
	OONS	-43.14	-35.58	22.75	-31.55	1.25	2.66
	SCKS	-44.79	-35.50	22.67	-32.58	0.62	2.60
	JRF	-39.10	-44.27	23.77	-19.07	0.46	4.64
WE1	RRIGS	-24.51	-16.94	22.31	-30.95	1.07	2.56
	WE1	-28.37	-19.35	22.66	-32.43	0.75	0.00
	WE2	-28.36	-19.31	22.67	-32.45	0.72	0.01
	OONS	-26.58	-19.03	22.75	-31.55	1.25	0.86
	SCKS	-28.19	-18.90	22.67	-32.58	0.62	0.77
	JRF	-18.76	-23.92	23.77	-19.07	0.46	3.98
WE2	RRIGS	-28.00	-20.44	22.31	-30.95	1.07	2.55
	WE1	-31.49	-22.47	22.66	-32.43	0.75	0.01
	WE2	-31.50	-22.44	22.67	-32.45	0.72	0.00
	OONS	-29.74	-22.19	22.75	-31.55	1.25	0.86
	SCKS	-31.33	-22.04	22.67	-32.58	0.62	0.77
	JRF	-22.35	-27.51	23.77	-19.07	0.46	3.97

Table SXXI: Comparison of hydration energies for leu–enkephalin. The first column refers to the hydration model used in the function evaluations, which are performed at the global solutions for the hydration model listed in the second column. The total energy, E_{TOT}, is provided along with the contributions from hydration, E_{HYD}, nonbonded interactions (including hydrogen bonding), E_{NB}, electrostatic interactions, E_{ES}, and torsion, E_{TOR}. The last column provides the heavy atom root mean squared deviation between the global minimum energy structures of the hydration models listed in the first two columns.

	Global of	$\mathbf{E}_{TOT}$	$\mathbf{E}_{HYD}$	$\mathbf{E}_{NB}$	$\mathbf{E}_{ES}$	$\mathbf{E}_{TOR}$	(RMSD)
OONS	RRIGS	-21.63	-14.06	22.31	-30.95	1.07	2.66
	WE1	-28.61	-19.59	22.66	-32.43	0.75	0.86
	WE2	-28.62	-19.56	22.67	-32.45	0.72	0.86
	OONS	-28.77	-21.21	22.75	-31.55	1.25	0.00
	SCKS	-28.37	-19.08	22.67	-32.58	0.62	1.14
	JRF	-19.23	-24.39	23.77	-19.07	0.46	3.90
SCKS	RRIGS	5.57	13.14	22.31	-30.95	1.07	2.60
	WE1	2.54	11.56	22.66	-32.43	0.75	0.77
	WE2	2.53	11.58	22.67	-32.45	0.72	0.77
	OONS	4.15	11.71	22.75	-31.55	1.25	1.14
	SCKS	2.36	11.65	22.67	-32.58	0.62	0.00
	JRF	18.35	13.19	23.77	-19.07	0.46	4.06
JRF	RRIGS	-112.59	-105.02	22.31	-30.95	1.07	4.64
	WE1	-152.64	-143.62	22.66	-32.43	0.75	3.98
	WE2	-152.61	-143.56	22.67	-32.45	0.72	3.97
	OONS	-158.44	-150.88	22.75	-31.55	1.25	3.90
	SCKS	-149.29	-140.00	22.67	-32.58	0.62	4.06
	JRF	-263.14	-268.31	23.77	-19.07	0.46	0.00

List of figures

Figure S1: Adiabatic ϕ–ψ map for unsolvated N–acetyl–N'–methyl–alanineamide. The adiabatic curves define regions within a given energy (1, 2, 5, 9 kcal/mole) of the global minimum value, and the (*) represents the location of the global minimum.

Figure S2: Adiabatic ϕ–ψ map for solvated N–acetyl–N'–methyl–alanineamide, using the RRIGS solvation model. The adiabatic curves define regions within a given energy (1, 2, 5, 9 kcal/mole) of the global minimum value, and the (*) represents the location of the global minimum.

Figure S3: Adiabatic ϕ–ψ map for solvated N–acetyl–N'–methyl–alanineamide, using the WE1 ASP set. The adiabatic curves define regions within a given energy (1, 2, 5, 9 kcal/mole) of the global minimum value, and the (*) represents the location of the global minimum.

Figure S4: Adiabatic ϕ–ψ map for solvated N–acetyl–N'–methyl–alanineamide, using the WE2 ASP set. The adiabatic curves define regions within a given energy (1, 2, 5, 9 kcal/mole) of the global minimum value, and the (*) represents the location of the global minimum.

Figure S5: Adiabatic ϕ–ψ map for solvated N–acetyl–N'–methyl–alanineamide, using the OONS ASP. The adiabatic curves define regions within a given energy (1, 2, 5, 9 kcal/mole) of the global minimum value, and the (*) represents the location of the global minimum.

Figure S6: Adiabatic ϕ–ψ map for solvated N–acetyl–N'–methyl–alanineamide, using the SCKS ASP set. The adiabatic curves define regions within a given energy (1, 2, 5, 9 kcal/mole) of the global minimum value, and the (*) represents the location of the global minimum.

Figure S7: Adiabatic ϕ–ψ map for solvated N–acetyl–N'–methyl–alanineamide, using the JRF ASP set. The adiabatic curves define regions within a given energy (1, 2, 5, 9 kcal/mole) of the global minimum value, and the (*) represents the location of the global minimum.

Figure S8: Plot of met–enkephalin conformation. Global minimum energy of -33.27 kcal/mole using the WE2 model for hydration.

Figure S9: Plot of met–enkephalin conformation. Global minimum energy of -31.45 kcal/mole using the OONS model for hydration.

Figure S10: Plot of unsolvated leu–enkephalin conformation. Global minimum energy of -9.349 kcal/mole.

Figure S11: Plot of leu–enkephalin conformation. Global minimum energy of -46.57 kcal/mole using the RRIGS model for hydration.

Figure S12: Plot of leu–enkephalin conformation. Global minimum energy of -28.37 kcal/mole using the WE1 model for hydration.

Figure S13: Plot of leu–enkephalin conformation. Global minimum energy of -31.50 kcal/mole using the WE2 model for hydration.

Figure S14: Plot of leu–enkephalin conformation. Global minimum energy of -28.77 kcal/mole using the OONS model for hydration.

Figure S15: Plot of leu–enkephalin conformation. Global minimum energy of 2.35 kcal/mole using the SCKS model for hydration.

Figure S16: Plot of leu–enkephalin conformation. Global minimum energy of -263.14 kcal/mole using the JRF model for hydration.

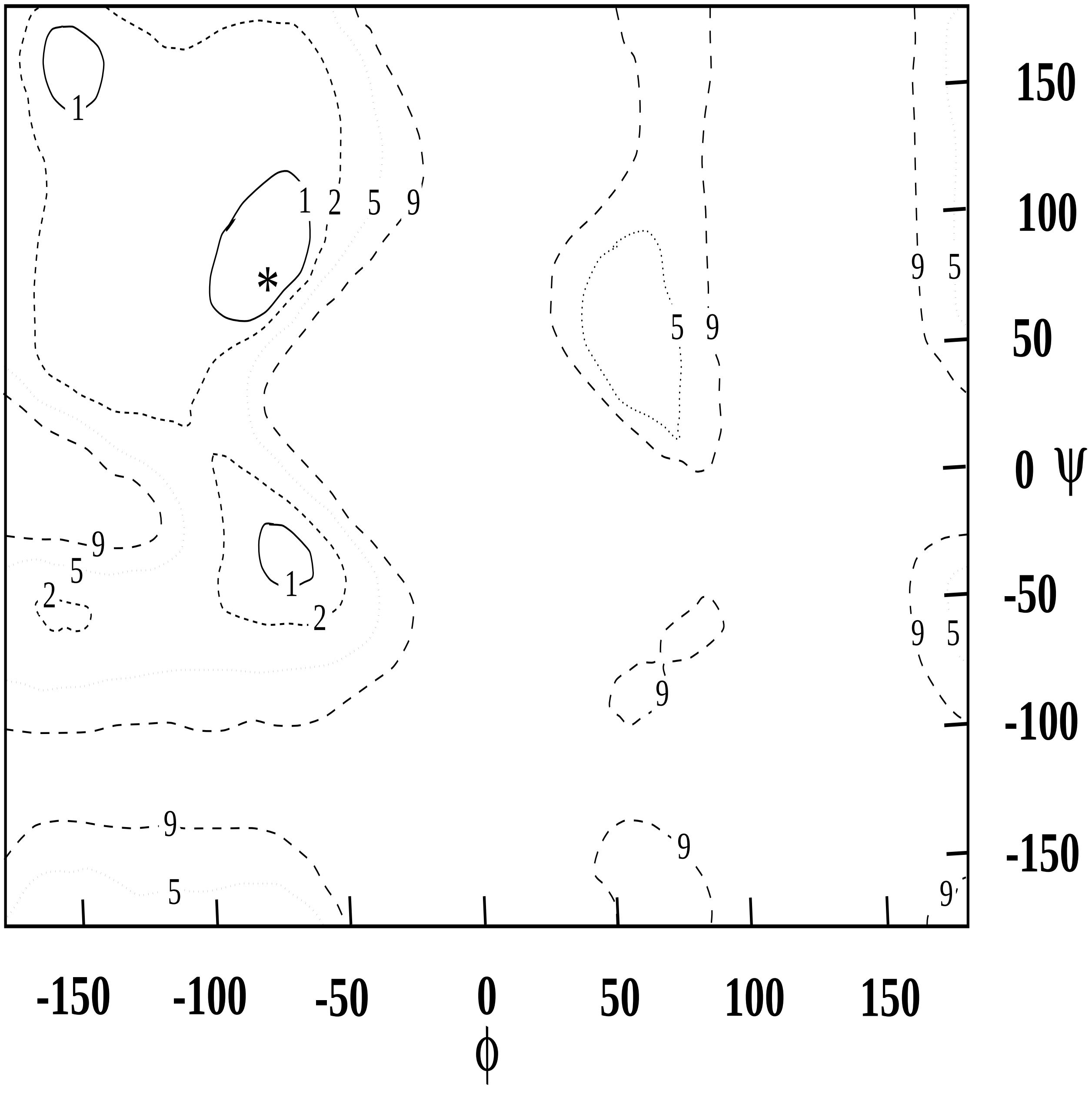

150
100
50
0 ψ
-50
-100
-150
-150
-100
-50
0
50
100
150
ϕ
1
1 2 5 9
*
9 5
5 9
9
5
2
1
2
9 5
9
9
5
9
9

Figure S1:

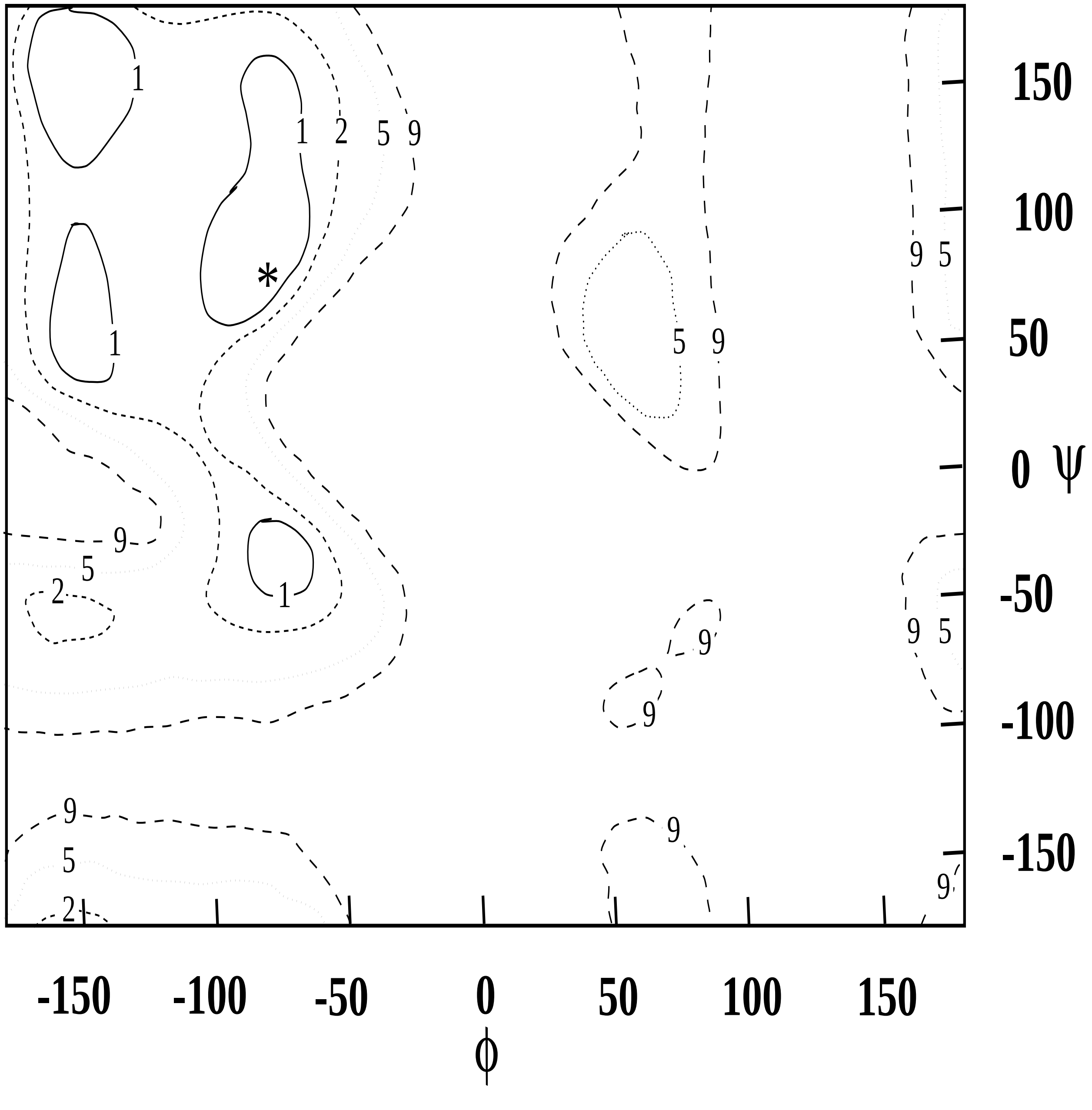

1
1 2 5 9
*
1
5 9
9 5
9
5
2
1
9
9 5
9
9
5
2
9
9
150
100
50
0 ψ
-50
-100
-150
-150 -100 -50 0 50 100 150
φ

Figure S2:

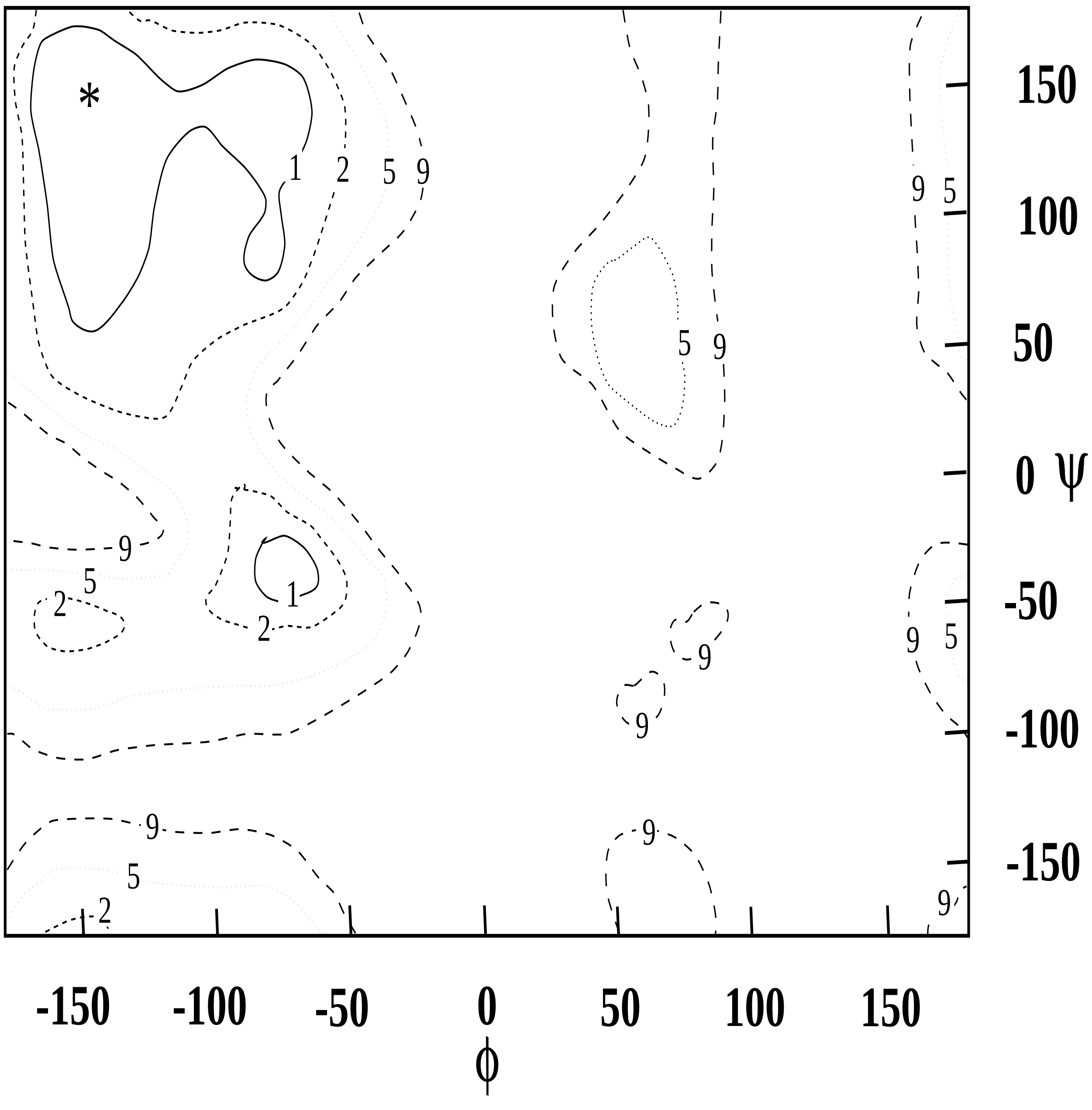


Figure S3:

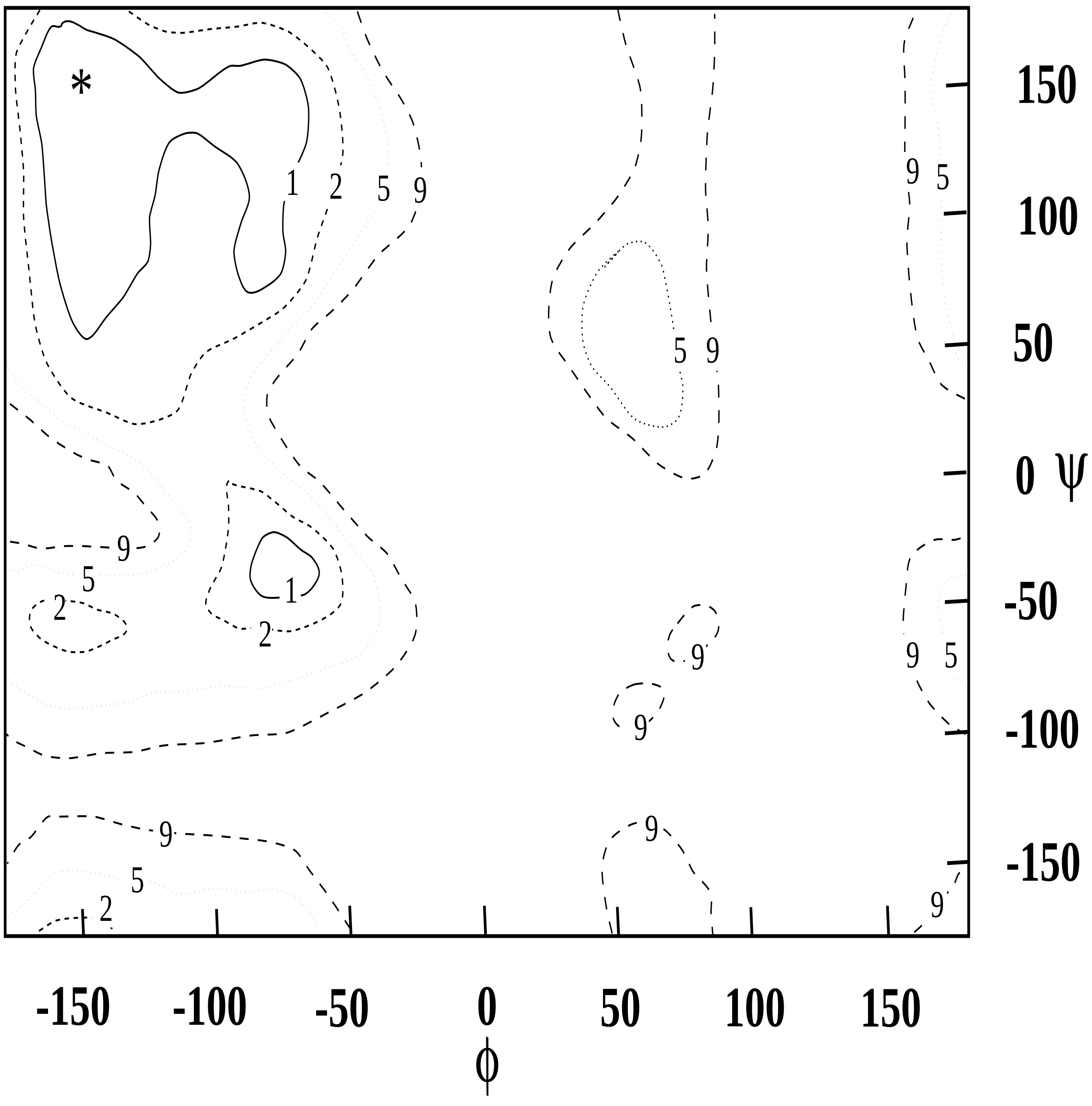
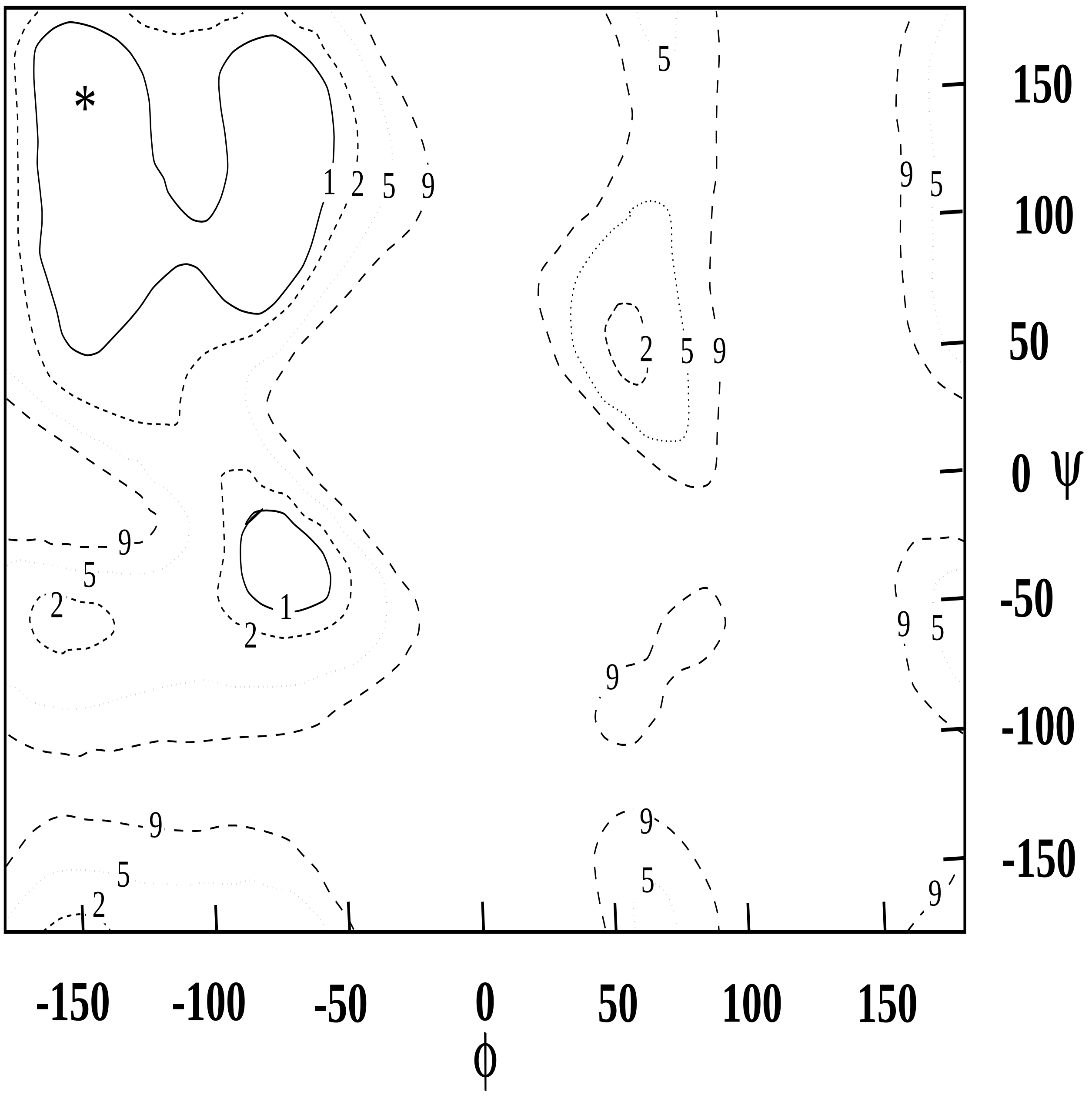

*
1
2
5
9
9
5
5
9
9
5
1
2
2
9
9
5
9
9
9
5
2
9
150
100
50
0
ψ
-50
-100
-150
-150
-100
-50
0
50
100
150
ϕ

Figure S4:



*
1 2 5 9
5
2 5 9
9 5
9
5
2
1
2
9 5
9
9
5
2
9
5
9
150
100
50
0 ψ
-50
-100
-150
-150 -100 -50 0 50 100 150
ϕ

Figure S5:

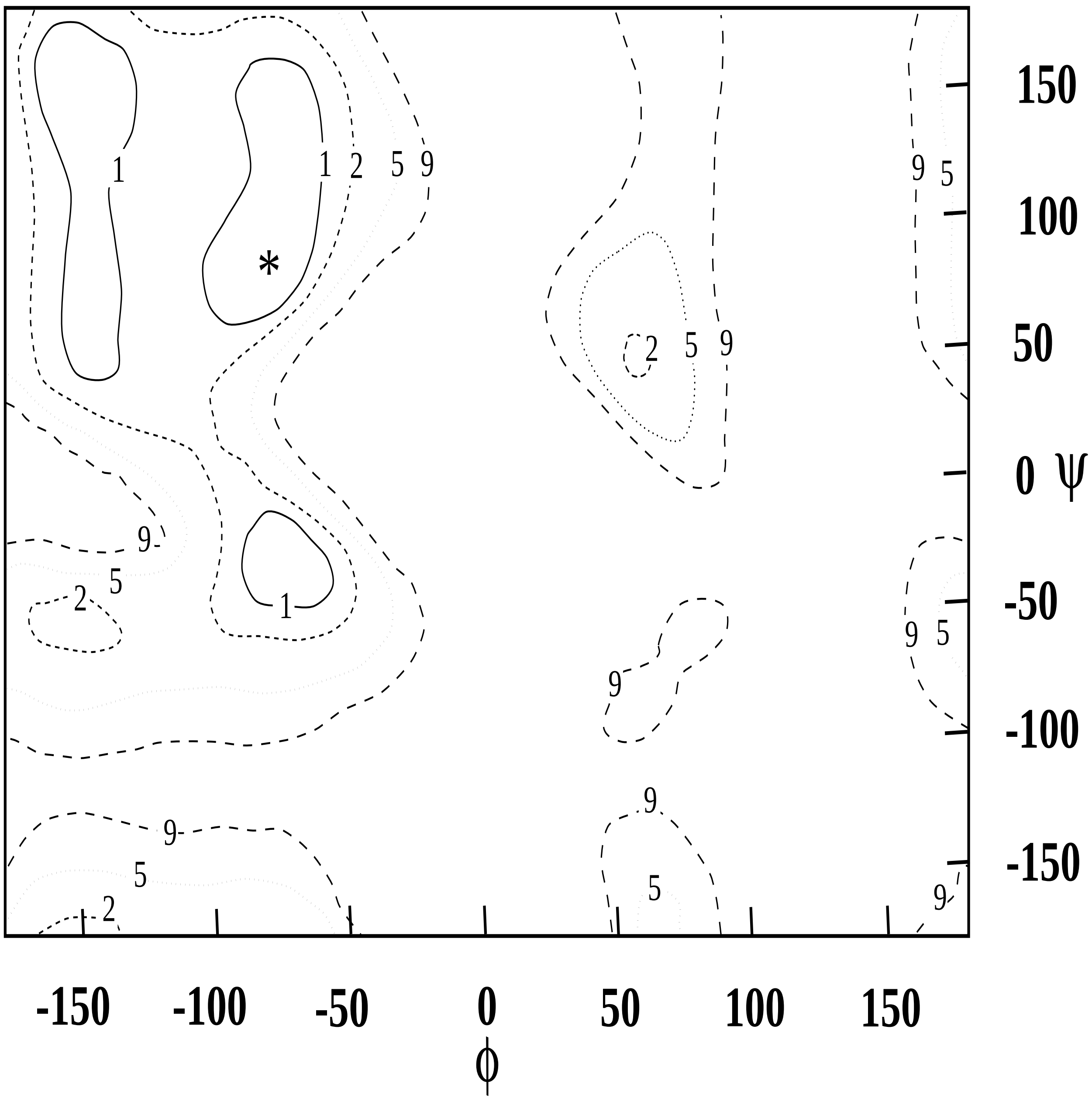

150
100
50
0 ψ
-50
-100
-150
-150
-100
-50
0
50
100
150
ϕ
1
1
2
5
9
*
2
5
9
9
5
9
2
5
1
9
9
9
5
9
5
2
9
9
5

Figure S6:

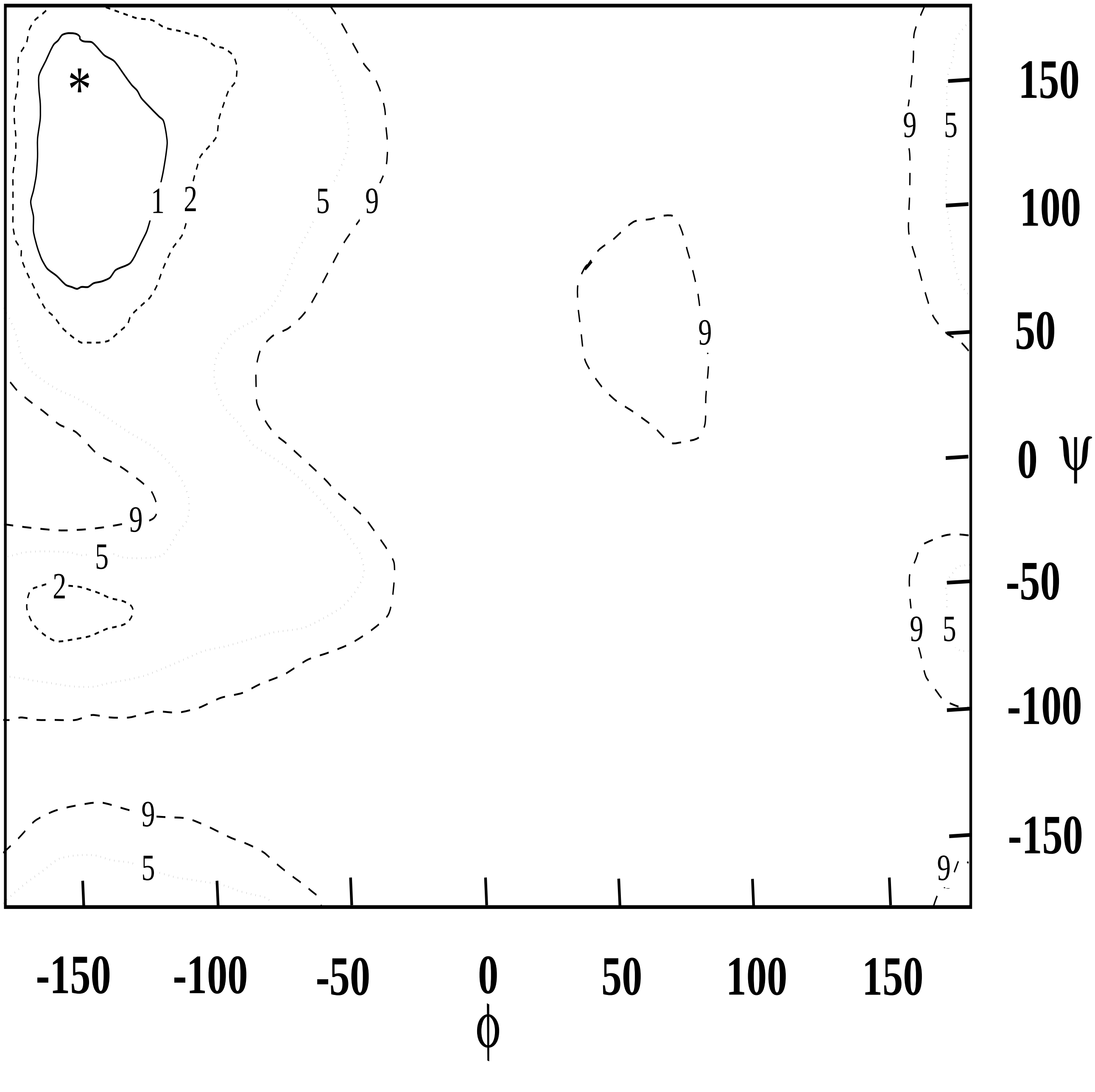

*
1
2
5
9
2
5
9
9
5
9
5
9
9
5
9
150
100
50
0
-50
-100
-150
ψ
-150
-100
-50
0
50
100
150
φ

Figure S7:

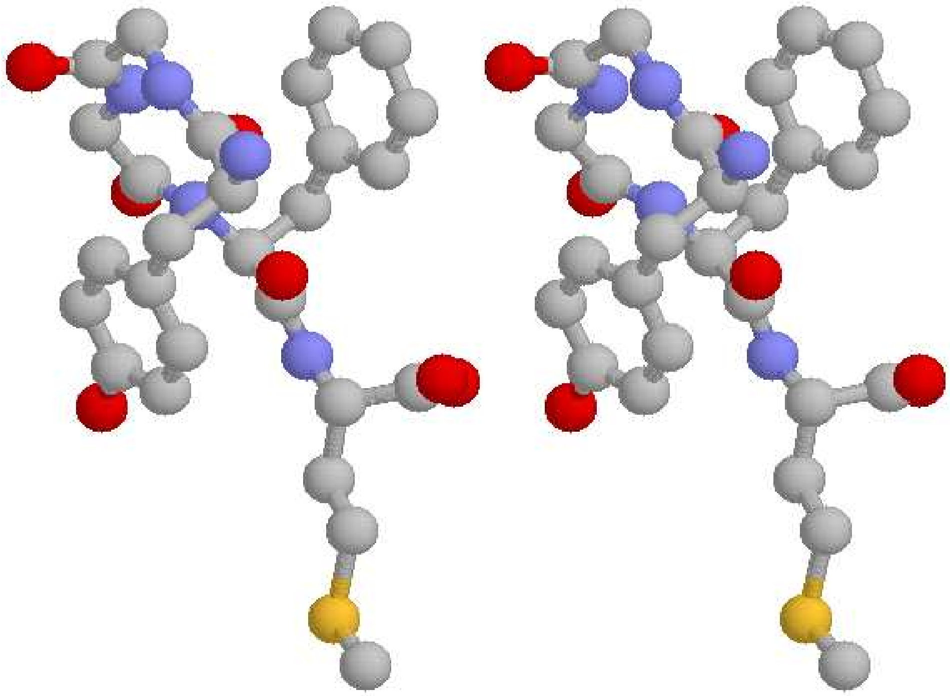


Figure S8:

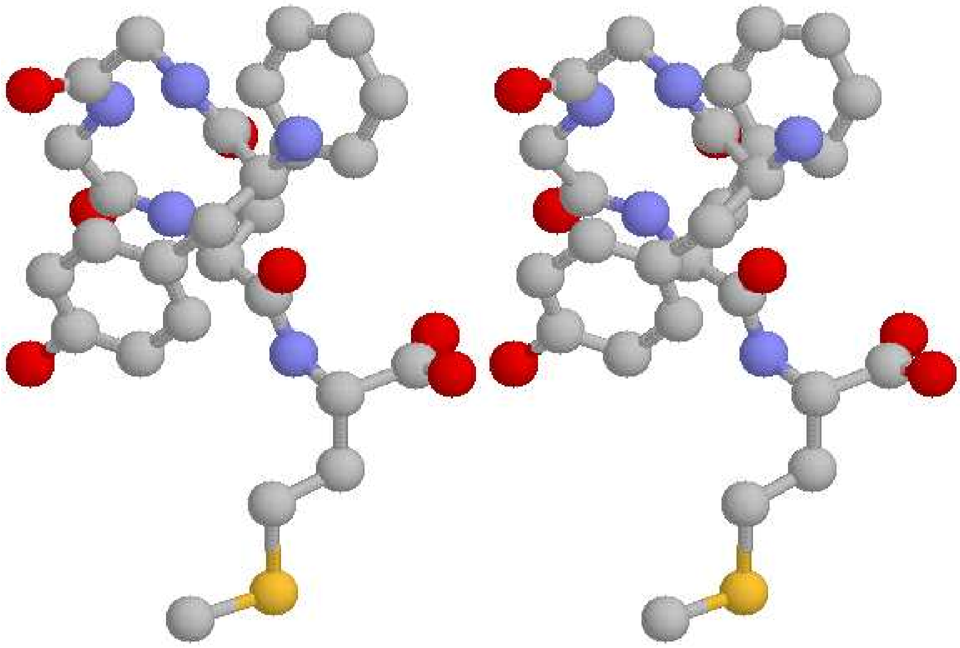


Figure S9:

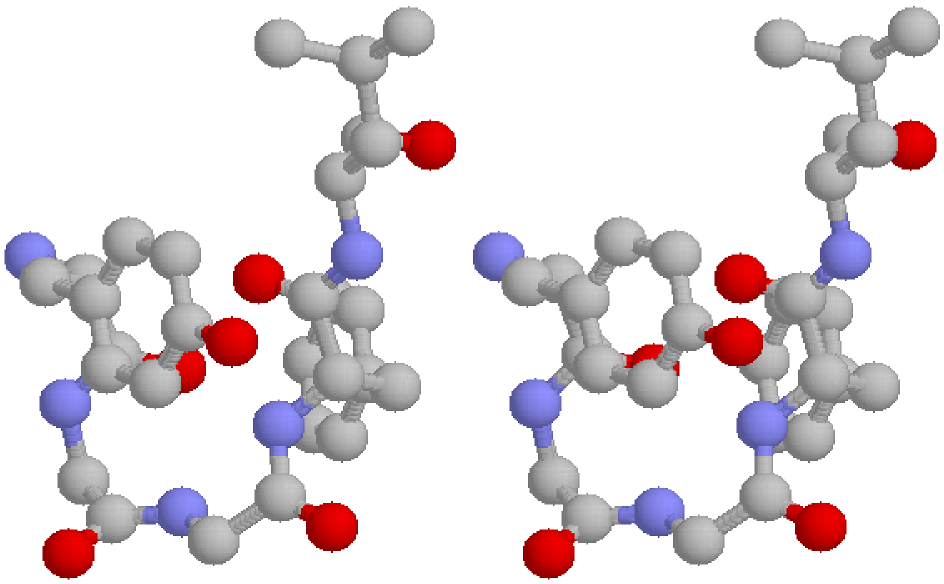


Figure S10:

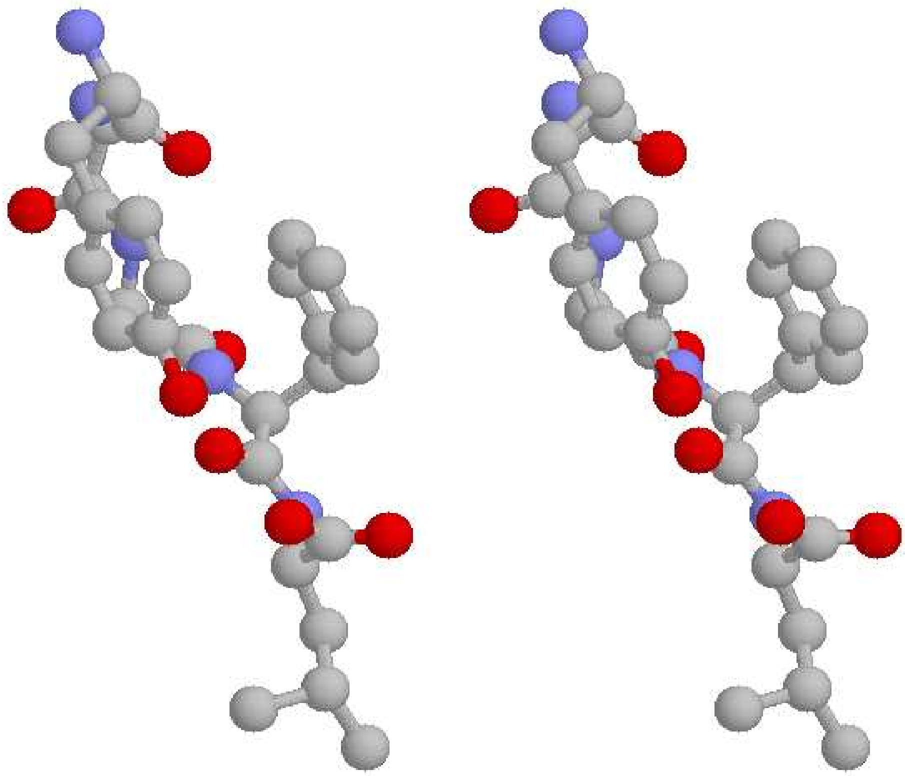


Figure S11:

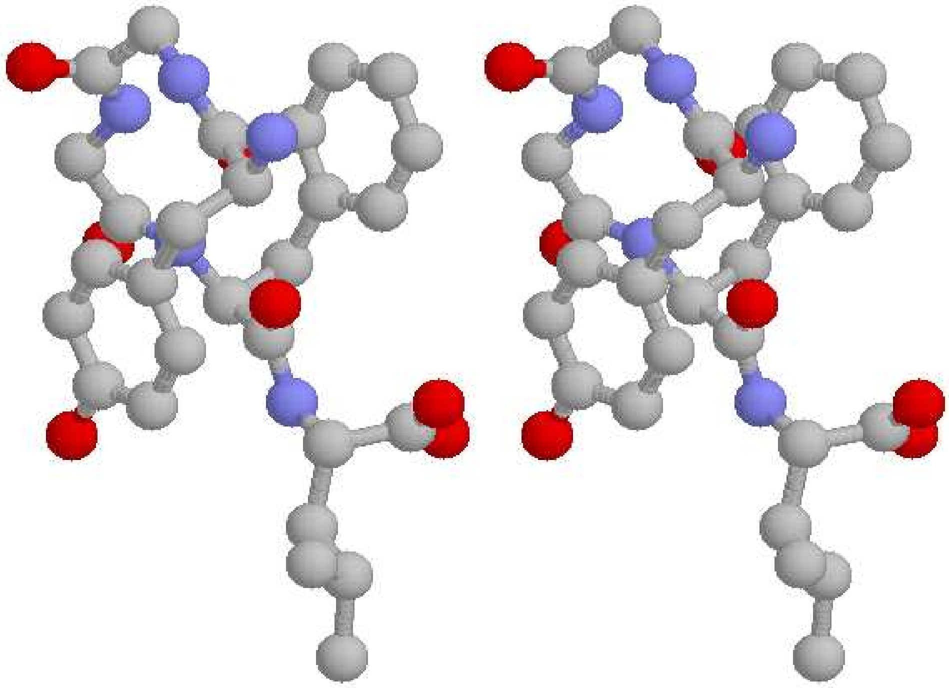


Figure S12:

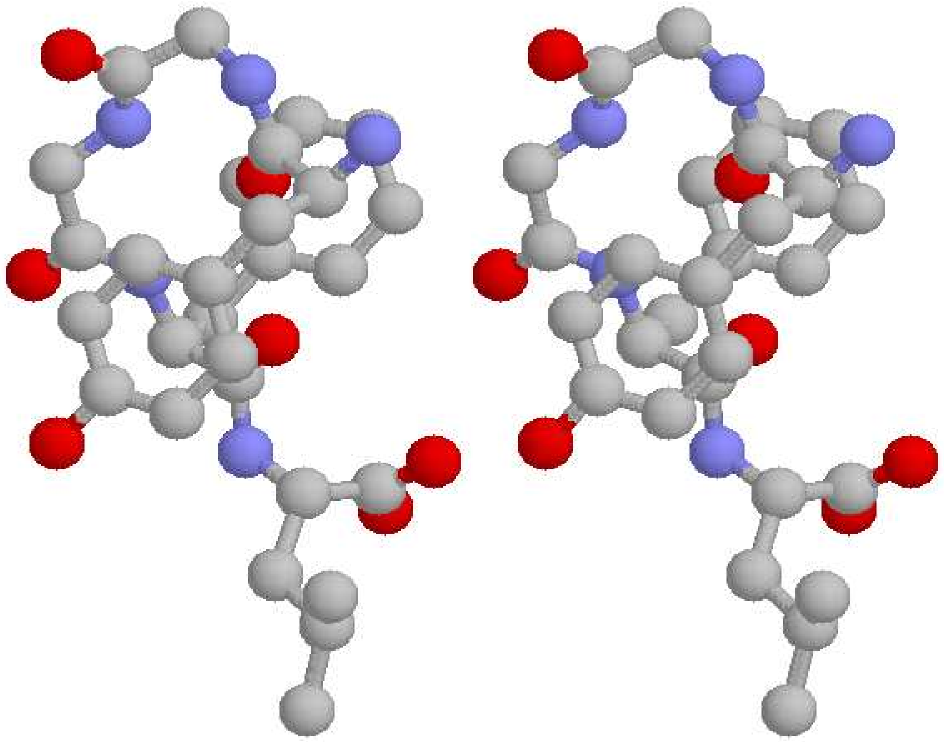


Figure S13:

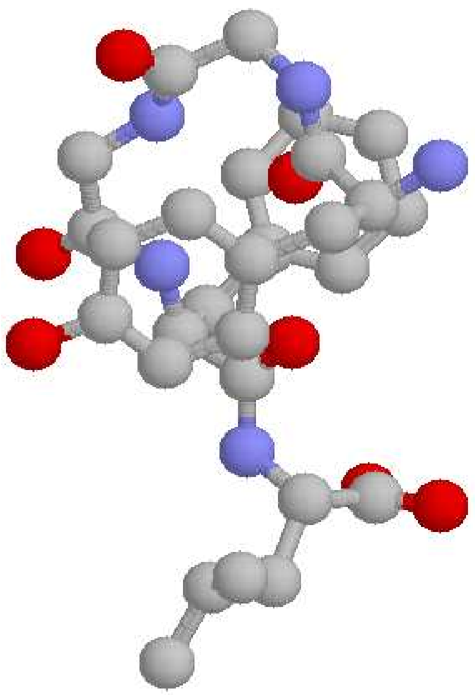


Figure S14:

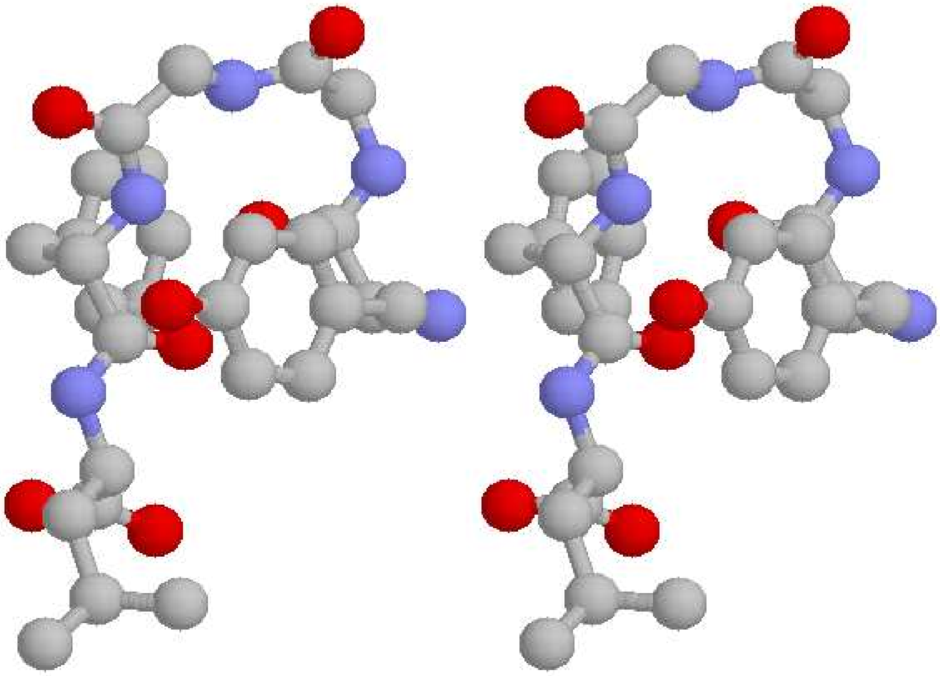


Figure S15:

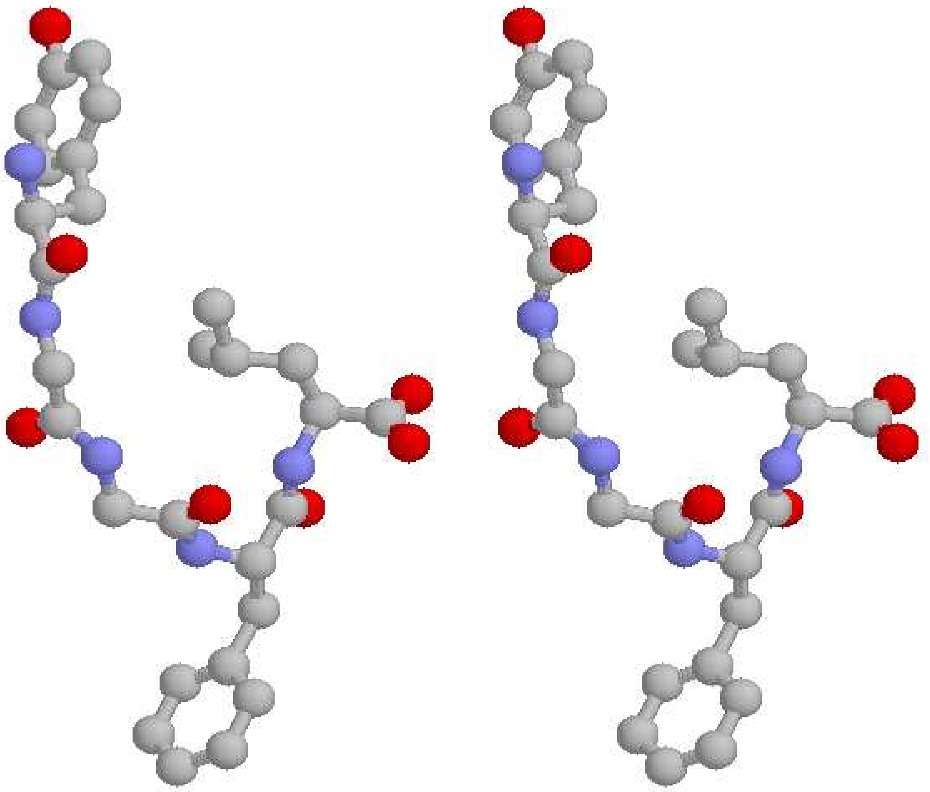


Figure S16: