Theoretical Study on Hydrophobicity of Amino Acids by the Solvation Free Energy Density Model

Junhyoung Kim, Ky-Youb Nam, Kwang-Hwi Cho,[†,*] Seung-Hoon Choi,[‡] Jae Sung Noh,[§] and Kyoung Tai No[¶,#,*]

Department of Chemistry and CAMD Research Center, Soongsil University, Seoul 156-743, Korea
[†]Department of Bioinformatics and Life Science, Soongsil University, Seoul 156-743, Korea
[‡]In Silico Tech Inc., Sungnam-Shi, KyungKi-Do 463-805, Korea
[§]Korea Research Institute of Chemical Technology, Daejeon 305-343, Korea
[¶]Department of Biotechnology, Yonsei University, Seoul 120-749, Korea
Received June 23, 2003

In order to characterize the hydrophobic parameters of *N*-acetyl amino acid amides in 1-octanol/water, a theoretical calculation was carried out using a solvation free energy density model. The hydrophobicity parameters of the molecules are obtained with the consideration of the solvation free energy over the solvent volume surrounding the solute, using a grid model. Our method can account for the solvent accessible surface area of the molecules according to conformational variations. Through a comparison of the hydrophobicity of our calculation and that of other experimental/theoretical works, the solvation free energy density model is proven to be a useful tool for the evaluation of the hydrophobicity of amino acids and peptides. In order to evaluate the solvation free energy density model as a method of calculating the activity of drugs using the hydrophobicity of its building blocks, the contracture of Bradykinin potentiating pentapeptide was also predicted from the hydrophobicity of each residue. The solvation free energy density model can be used to employ descriptors for the prediction of peptide activities in drug discovery, as well as to calculate the hydrophobicity of amino acids.

Key Words : Hydrophobicity, Solvation free energy density model, Bradykinin potentiating pentapeptide, Quantitative structure-activity relationship

Introduction

The concept of hydrophobicity has become an invaluable tool in drug research and pharmaceutical sciences. First, it takes the role of a physicochemical descriptor that can be empirically correlated with an unending variety of pharmacodynamic and pharmacokinetic parameters. As such, hydrophobicity has much to contribute to drug design and to chemical interpretation of pharmacological processes. Second, as a molecular property, hydrophobicity allows unique insights into intramolecular effects and intermolecular recognition forces.[1]

There have been many studies to determine parameters which would allow one to describe nonpolar, polar, and ionic side chains.[2-7] These parameters were found experimentally from measurements of amino acid solubility in water and in organic solvents. The free energy differences obtained from transferring amino acid chains from an organic solvent to water, are called the partition coefficients, log*P*, or hydrophobicity scales.

Nozaki and Tanford[8] were the first to identify hydrophobicity scales for ten non-polar amino acids. This work relied upon the measurement of the hydrophobicity scales of free amino acids in water/dioxane and water/ethanol mixtures. Wolfenden *et al.*[9] used the partitioning between water and its vapor from the side chains of amino acids where the backbone residue is replaced by one hydrogen atom. Fauchère and Pliška[10] provided a complete list of hydrophobicity scales of 20 amino acids based on experiments (Table 1). These hydrophobicity scales became a new reference for comparison with previous data, and helped to explain discrepancies between the existing scales and to circumvent the better individual hydrophobic constants. The use of free amino acids in solubility measurements and partitioning experiments, suffers from a series of disadvantages. Fauchère and Pliška[10] described measurements of the hydrophobicity scales of amino-acid side chains performed on the new series of the *N*-acetyl-L-amino-acid amides. Recently, Shin *et al.* provided an experimental scale of hydrophobicity for nucleic acid bases.[11] The hydrophobicity scale has many usages in many areas. For instance, establishing a good set of hydrophobicity scales for amino acids is a valuable tool for the study of the three dimensional protein structures, and provides insights into processes such as protein folding and binding and for the study of quantitative structure-activity relationships (QSAR) in polypeptide hormones.

Peptides are very important molecules in biological systems. Many pharmaceutically useful peptide or peptidomimetic agents are known to serve as inhibitors, agonists, or antagonists.[12-13] Despite difficulties in the QSAR analysis of

[*]To whom correspondence should be addressed. Tel: +82-2-825-1785; Fax: +82-2-825-1795; E-mail: khcho1@ssu.ac.kr, kyn@bmdrc.org
[#]Member of Hyperstructured Organic Material Research Center, Korea

peptides and proteins, there have been several successful reports describing molecular structure in a quantitative way.[14-16] The study of Hellberg *et al*.[17] was the precursor to developing the descriptors for peptides. Their QSAR methodology for peptides is based on the parameterization of each amino acid occurring in a peptide chain with three *z*-values. Using a principle component analysis (PCA), the *z*-values were derived from a collection of experimental data on a number of peptides. This data included HPLC retention times, pK_as, NMR-derived properties, and other measurable variables related to hydrophobicity, size, and electronic features. Even though the method could be extended to unnatural amino acids,[18] it does have limitations. For instance, the method does not allow a straightforward interpretation of the results from a QSAR study in terms of physicochemical factors that are important for biological activity. Since the *z*-scales are derived from macroscopic properties,[18] the various conformations of the molecules are not explicitly considered in their derivation. This represents a very important defect in view of the fact that a molecule can have various conformations in different environments. One way to overcome the problem is to generate theoretical features that can take into consideration the three-dimensional structure of a molecule. To overcome the conformation problem, Cocchi *et al*. conducted another parameterization of amino acid side chains.[19] In this approach, scores defined as *t*-scores, were derived from a PCA of the interaction energies and calculated using a program called GRID.[19] The scores proved to be effective when applied to a QSAR study of a set of dipeptide ACE inhibitors. Collantes *et al*.[20] demonstrated that two computable 3D-descriptors, Isotropic Surface Area (ISA), and Electronic Charge Index (ECI), could be successfully used as side-chain descriptors for amino acides. While ISA correlates well with the first component of *z*-score values and with Fauchère and Pliska's hydrophobicity scale,[10] the ECI showed a good correlation with the amino acid free energy of vaporization. These results provide evidence to support the suggestion that calculated and structurally derived properties can be used to generate a robust description of residues in a peptide sequence. It has now become obvious that in order to develop meaningful quantitative models of structure-activity relationships, it is necessary to consider the three-dimensional structures of the active compounds.

In our previous paper, a solvation free energy density (SFED) model[21] was proposed. This model allows for the prediction of hydration free energy and a method for calculating logP values.[21-22] The success of this approach in obtaining logP for common organic compounds encouraged us to expand the SFED model for a physicochemical study of natural peptides and proteins. In this study, the hydrophobicity scales of amino acids and peptide side chains were calculated using the SFED model and the quality of the model was evaluated. The basis of verification was the calculated hydrophobicity scales. These scales were verified by comparing them to the experimental hydrophobicity scales of the amino acid side chains. Since the conformation of molecules has an effect on the solvent environment, this calculation was performed in the gas and solution phase in order to allow observation of the variation in hydrophobicity which resulted from the influence of the solvent environment. The theoretical hydrophobicity scales obtained from the application of the SFED model, agree favorably with the experimental hydrophobicity scales obtained for the *N*-acetyl amino acid amides.

The hydrophobicity scales, which are predicted using the SFED model, are the sum of the free energy of interaction and the free energy of cavitation, and take into account hydrophobicity scales which contain spatial distribution with structural information. In order to study the hydrophobicity variations in relation to the conformational changes in peptide side chains caused by the structural effect of neighboring side chains, the method was applied to small peptides. The calculated hydrophobicity scales containing spatial distribution as descriptors were verified using 30 Bradykinin potentiating pentapeptides as a testing molecule set for the QSAR study. The multiple linear regression (MLR) method was used to model the data in the verification procedure. The calculated hydrophobicity scales were obtained for both *N*-acetyl amino acid amide side chains and *N*-acetyl Bradykinin potentiating pentapeptide amide side chains using ECEPP/NKS[22-23] potential energy parameters. In fact, the hydrophobicity of a polypeptide cannot be represented by the simple sum of the hydrophobicity of single amino acid side chains, even though the molecular formulae for side chains are the same as those for peptides. Moreover, hydrophobicity scales which incorporated the conformational effect, functioned as better descriptors. Since proteins have large conformational changes, the SFED model, which can take into account conformational variations, proved to be more appropriate for the study of peptides and proteins.

Method

Our solvation free energy model is applied to the calculation of hydrophobicity scales of terminal blocked single amino acids. The activities of polypeptides are also predicted using a QSAR method based on the hydrophobicity of its residues as a descriptor with/without the consideration of conformational variations. The details of the computational procedure are shown in two stages for the sake of convenience of explanation. At the first stage, the formulation of the SFED model and the calculation scheme for partition coefficients of molecules are briefly explained. Next, the procedure of hydrophobicity calculation of *N*-acetyl amino acid amides and *N*-acetyl Bradykinin potentiating pentapeptide amides is described. To take into account the contributions of side chains to hydrophobicity, calculations are performed individually in the gas and solution phases.

Solvation Free Energy Density Model and the Calculation of Hydrophobicity Scale.

Solvation Free Energy Density Model: In the SFED model,[21] the solvation free energy was described as

$$\Delta G_{sol} = \int_Q \left[\sum_i^{N_A} f(\mathbf{r}_{ik}) + g_{cav} \right] d\mathbf{r}_k = \int_Q g(\mathbf{r}_k) d\mathbf{r}_k \quad (1)$$

Where N_A is the number of atoms in the solute, g_{cav} represents the cavitation energy per unit surface area, Q represents the solvent domain around the solute which is indicated by the shaded area in Figure 1, $\sum_i f(\mathbf{r}_{ik})$ contains the contributions from the solute-solvent and solvent-solvent stabilization energies at the point $\mathbf{r}_k$, and $g(\mathbf{r}_k)$ represents the free energy density at the grid point $\mathbf{r}_k$.

In the computational procedure, as an approximation, $\int_Q$ in equation (1) was replaced by a summation over a grid around the solute, $\sum_Q$. Then, equation (1) was approximated as

$$\Delta G_{sol} = a_l \sum_k^Q \sum_i^{N_A} f(r_{ik}) + \Delta G_{cav} \quad (2)$$

Where a_i is proportionality constant that depends on the size of the grid interval (Δl). As illustrated in Figure 1, the space between the Solvent Accessible Surface (SAS) and the outer surface corresponds to the solvation shell. The SAS of the solute is defined by the overlap of the atomic SAS. The SAS is the spherical surface whose radius is the sum of the atomic *van der Waals* radius, and the effective solvent (water or 1-octanol) shell thickness (R_w or R_{oct}). The optimum values of Δl, R_w, R_{oct} and R_{shell} were taken as a compromise considering both the computing time and the accuracy of the calculation. Details of the calculation were described in our previous papers.[21]

The Partition Coefficient (log*P*) of Amino Acid Chains Calculation: The free energy of transferring amino acid chains from an organic solvent to water is defined[8,10] as follows,

$$\Delta G_{trans} = \Delta G_{solv}^{water} - \Delta G_{solv}^{oct} \quad (3)$$

where ΔG_{solv}^{water} and ΔG_{solv}^{oct} are the free energy of the solute in water and in 1-octanol phase, respectively.

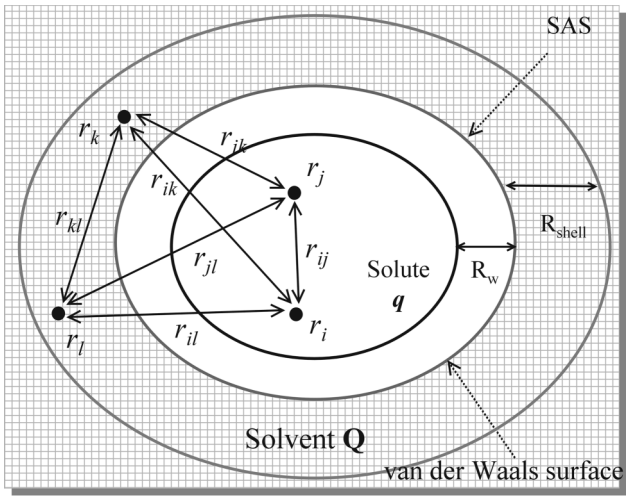

Figure 1. Schematic representation of the Solvation Free Energy Density (SFED) model.

In the case of water/1-octanol solution, the transfer free energy, ΔG_{trans}, is related to partition coefficients, log*P* and is defined and approximated as follows,

$$\log P = -\frac{1}{2.030RT}(\Delta G_{solv}^{water} - \Delta G_{solv}^{oct}) = \frac{1}{2.303RT}\Delta G_{trans} \quad (4)$$

Hydrophobicity Scales of Peptides.

***N*-Acetyl Amino Acid Amides:** As shown in Figure 2, the dihedral angles of backbone for the initial structures of *N*-acetyl amino acid amide are taken from the C_7^{eq} conformation corresponding to $\phi \cong 80.0^o$, $\psi \cong 70.0^o$ and $\omega \cong 180^o$ except in the case of N-acetyl proline amide. In the case of side chains, the default peptide library of Insight II/ BIOPOLYMER[21] is adopted as the initial conformations. Since the side chains of Arg, Asp, Glu, and Lys are ionized under these physiological conditions, the simulation is performed with the charged forms. Next, these structures are independently minimized in the gas phase using the ECEPP/ NKS potential energy function[28-35] and the SUMSL minimization algorithm.[26] In the solution phase, the procedure is conducted using the ECEPP/NKS with the SFED model and the SMSNO minimization algorithm.[26] In the case of the solution phase, the solvation free energy term is appended to the total energy as follows,

$$E_{tot}^{gas} = E_{tor} + E_{es} + E_{vdw} + E_{hb} \quad (5)$$

$$E_{tot}^{gas} = \sum V_n(1 + s\cos n\omega) + \sum \frac{q_i q_j}{\varepsilon_0 r} + \sum \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r^6} \right) + \sum \left(\frac{B_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r^6} \right) \quad (6)$$

$$\Delta G_{tot}^{sol} = \sum V_n(1 + s\cos n\omega) + \sum \frac{q_i q_j}{\varepsilon_0 r} + \sum \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r^6} \right) + \sum \left(\frac{B_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r^6} \right) + \Delta G_{sol} \quad (7)$$

The total conformational energy (E_{tot}) is calculated as the sum of the electrostatic energy (E_{es}), the nonbonded energy (E_{vdw}), and the *intrinsic* torsional energy (E_{tor}). The hydrogen-bond energy (E_{hb}) is included in the nonbonded energy components.[27] The forms of the electrostatic and nonbonded potentials were taken as those of the ECEPP/ NKS potential energy function.[27] The partial atomic charges for each molecule were calculated by the MPEOE method of No *et al.* [28-31] The effective dielectric constant was taken as

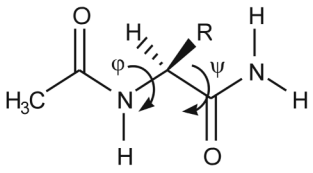

Figure 2. Schematic diagrams of *N*-acetyl amino acid amides (NAA).

unity[28-35] instead of using 2 (effectively 4) as in the ECEPP/NKS. The recently derived values of dispersion coefficients[32] were used for the attractive term of the nonbonded Lennard-Jones potential, E_{nb},[35] and the repulsive coefficients were obtained from crystal data on organic compounds and amino acids.[35] The hydrogen-bond energy (E_{hb}) was calculated by using 6-12 type potential functions.[34] The parameters were optimized using potential energy surfaces of several hydrogen-bonded molecular pairs. The energy surfaces were obtained using *ab initio* Molecular Orbital calculation at the HF/6-31G** level of theory. The functional form replaces the 10-12 type (E_{hb}) of the ECEPP/NKS version.

The hydrophobicity scales of these N-acetyl amino acid amides in the aqueous and 1-octanol phase are subsequently calculated using the SFED model. It is well established that the hydrophobicity scales of an amino acid are represented as the difference between glycine and the other amino acids. The values for the other amino acids are scaled to the values of glycine that is set at 0.0.

The hydrophobicity scale, $\Delta\Delta G_{trans}^{cal}$, is defined and approximated as follows,

$$\begin{aligned}\Delta\Delta G_{tans}^{cal} &= \Delta G_{trans}^{X} - \Delta G_{trans}^{glycine} \\ &= -2.303(\log P_{X}^{cal} - \log P_{glycine}^{cal}) \qquad (7)\end{aligned}$$

where ΔG_{trans}^{X} and $\Delta G_{trans}^{glycine}$ are the transfer free energy of any amino acid and the transfer free energy of glycine, respectively.

***N*-Acetyl Bradykinin Potentiating Pentapeptide Amides:** The charged states of side chains and the initial structures of Bradykinin potentiating pentapeptides terminally blocked by acetyl and amide groups, are built in the same manner as those described for the *N*-acetyl amino acid amide except in the following cases. The antiparallel β-sheet conformation, corresponding to $\phi \cong 120.0^{o}$, $\psi \cong 120.0^{\circ}$ and $\omega \cong 180.0^{o}$, is selected for the backbone torsion angles of the initial conformation based on previous experimental work.[36] A random conformational search of side chains was performed to acquire stable conformations of side chains. This was carried out using the default peptide library of Insight II/BIOPOLYMER as the initial conformations.

In the case of the gas phase simulation, the conformation of side chains was minimized using a conjugate gradient method. In order to avoid the deformation of backbone structure, the minimization for side chains was only carried out under the fixed conformation of backbone torsions. The entire structures of pentapeptides were subsequently minimized in the solution phase as described earlier by using the ECEPP/NKS and SFED models. At this stage, after minimization of the side chain conformation, deformation of the backbone does not occur under the unfrozen backbone torsion angles.

The solvation free energy density model of each of the side chains of the minimized structures was carried out using the SFED model. Then, the hydrophobicity scales for each of the residues were calculated using the SFED model, as shown in Figure 3.

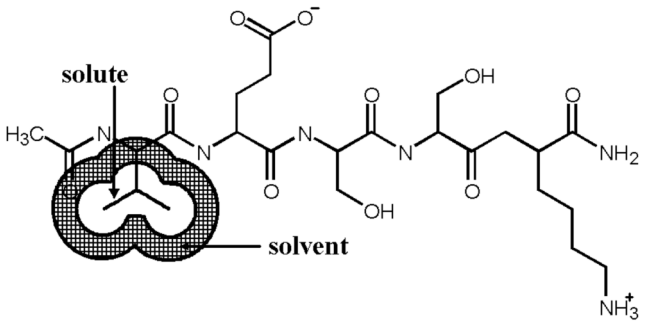

Figure 3. The schematic representation of local side chain hydrophobicity for Bradykinin potentiating pentapeptide.

Results and Discussion

The Hydrophobicity Scales of *N*-Acetyl Amino Acid Amides. The initial backbone structure of N-acetyl amino acid amide was set to the C_7^{eq} conformation because many other theoretical works suggest this conformation as the most stable structure in the gas phase.[37-42] The initial structure of each model amino acid was minimized for 1000 steps using the SUMSL method (for the gas phase) and the SMSNO method (for the solution phase by ECEPP/NKS with the SFED model). The hydrophobicity was calculated according to the equation (3).

Table 1 shows the experimental[10] and theoretical hydrophobicity scales for each amino acid, where RMSE stands for root mean square errors from the experimental hydrophobicity scales. The comparison of RMSE for non-charged residues in Table 1 shows that our prediction of hydrophobicity scales based on our SFED model is better than Rosemans theoretical work.[43] From the RMSE between experimental and calculated hydrophobicity scales in the gas phase using ECEPP/NKS and in the solution phase using ECEPP/NKS and SFED method are 1.893 and 1.912 (Table 1), respectively.

Figure 4 plots the experimental and calculated hydrophobicity scales. The gas phase and solution phase data are plotted separately. Some of the disagreement between experimental and calculated hydrophobicity scales is caused by ionizable side chains of amino acids. The carboxyl groups of Asp and Glu side chains ionize with the intrinsic pK_a values of 3.9 and 4.3, respectively. That means that these residues are ionized and polar under physiological conditions. The side chain of Lys is a hydrophobic chain of four methylene groups capped by an amino group that ionizes with an intrinsic pK_a value of 11.1 in the absence of perturbing factors. Therefore, it is also ionized under most physiological conditions. The Arg side chain consists of three non-polar methylene groups and the strongly basic δ-guanido group with a usual pK_a value of approximately 12.0. The δ-guanido group is ionized over the entire pH range in which proteins exist naturally. For these reasons, the hydrophobicity scales of Asp^{-}, Glu^{-}, Lys^{+} and Arg^{+} were calculated in the ionized states. The RMSE of the hydrophobicity scales at the gas and solutions phases, except for those of the charged *N*-acetyl amino acid amides (Arg^{+},

Table 1. The hydrophobicity scales of *N*-acetyl amino acid amides side chains

residue	$\log P^{\exp}$ [a]	$\log P_{ref}^{cal}$ [b]	$\log P_{gas}^{cal}$ [c]	$\log P_{sol}^{cal}$ [d]
Ala	-0.31	-0.390	-0.341	-0.350
Arg(+)	1.01	3.950	4.049	4.086
Asn	0.60	1.910	0.555	0.687
Asp(−)	0.77	3.810	5.999	6.067
Cys	-1.54	-0.250	-0.717	-0.688
Glu(−)	0.64	2.910	5.860	5.877
Gln	0.22	1.300	0.246	0.575
Gly	0.00	0.000	0.000	0.000
His	-0.13	0.640	0.040	-0.037
Ile	-1.80	-1.820	-1.383	-1.353
Leu	-1.70	-1.820	-1.362	-1.308
Lys(+)	0.99	2.770	2.576	2.594
Met	-1.23	-0.960	-1.419	-1.416
Phe	-1.79	-2.270	-1.808	-1.842
Pro	-0.72	-0.990	-1.145	-1.198
Ser	0.04	1.240	0.228	0.292
Thr	-0.26	1.000	-0.196	-0.030
Trp	-2.25	-2.130	-1.660	-1.684
Tyr	-0.96	-1.470	-1.276	-0.920
Val	-1.22	-1.300	-1.015	-0.964
RMSE[e]		1.358	1.893	1.912
RMSE[f]		0.768	0.341	0.364

[a] The experimental hydrophobicity values measured from the partition coefficient between water and octanol of the N-actyl amino acid amides by Fauchère and Pliška.[10] [b] reference 43. [c] The calculated hydrophobicity values obtained from the conformation of N-acetyl amino acid amides that minimized in the gas phase using ECEPP/NKS potential energy function. [d] The calculated hydrophobicity values obtained from the conformation of N-acetyl amino acid amides that minimized in solution using ECEPP/NKS potential and the SFED model. The solvation free energy term is added to the total energy. [e] The root mean square errors of 20 N-acetyl amino acid amides. [f] The root mean square errors of 16 N-acetyl amino acid amides except K, E, D and R.

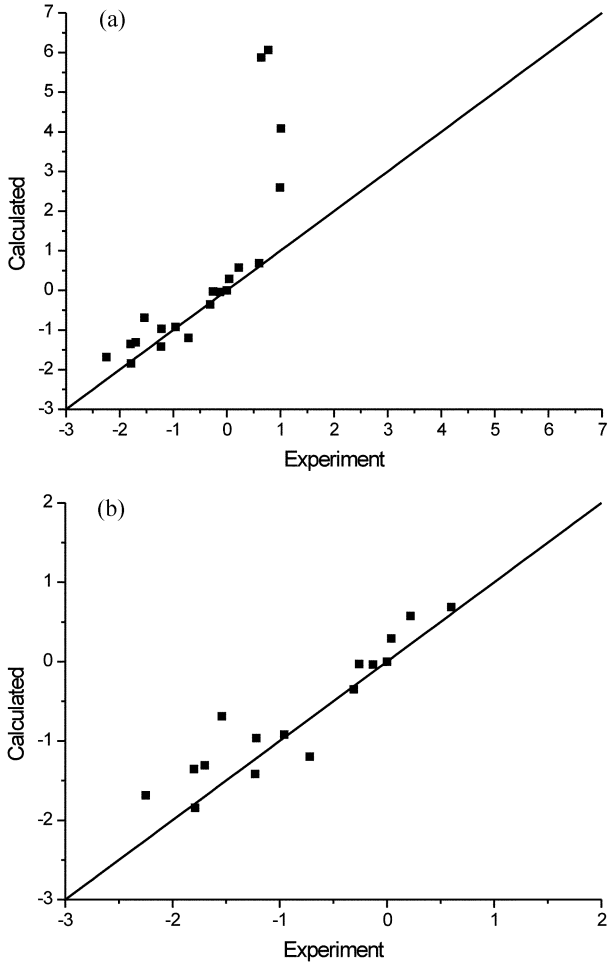

Figure 4. Experimental and calculated hydrophobicity plots of (a) 20*N*-acetyl amino acid amides are minimized at solution phase and (b) except charged form (Arg^+, Asp^-, Glu^- and Lys^+), 16*N*-acetyl amino acid amides are minimized in solution environment.

Asp^-, Glu^-, Lys^+), are 0.341 and 0.364, respectively.

The predictability of hydrophobicity scales of Asp^-, Glu^-, Lys^+ and Arg^+ amino acids decreased as shown in Table 1. The theoretical results of Roseman *et al.*[43] showed the same tendency. They explained that the loss of side-chain hydrophobicity can be attributed either to proximity effects, or to intra-molecular hydrogen bonding. In this calculation, the hydrophobicities of pentapeptide side chains are smaller than those for single amino acids (Table 1). The hydrophobicity scales of ionizable side chains showed a greater decrease (Table 1 and Table 2). It seems that the ionizable side chains interact with other ionizable side chains.

The experimental hydrophobicity scales of N-acetyl amino acids were obtained from equilibrium structures.[10] However, the calculated hydrophobicity scales were computed with a single conformation that is postulated for almost all stable structures. In spite of this fact, the relative order of calculated hydrophobicity scales is in good agreement with the experimental hydrophobicity order. The SFED model shows that the hydrophobicity scales of the anion charged amino acid (Asp^-, Glu^-) is less hydrophobic than the cation charged amino acid (Arg^+, Lys^+). Furthermore, these hydrophobicity values are larger than the experimental values. To compensate for this factor, the free energy term associated with change in the state of ionization of the ionizable groups, was added to the SFED model. The free energy term occurs due to the transfer of the molecule from the gas phase to the solvent phase, at a fixed pH value of 7.4.

The Local Hydrophobicity scales of Bradykinin Potentiating Pentapeptide Side Chains. The initial conformation of Bradykinin potentiating pentapeptides is set to the β-sheet conformation. Ferreira *et al.*[36] mentioned the conformation of the Bradykinin potentiating peptide F. They suggested two main conformers of the molecule.These are a favorable 'stretched' and less favorable 'folded' conformation. As a basis for these two conformations, they showed the Mass and NMR spectrometry data for the main configuration. The experimental results do not show the intra-molecular interaction due to the helix form. The initial structures of the *N*-acetyl Bradykinin potentiating pentapeptide amide backbone were set to the β-sheet conformation. Although the torsional angles of the backbone moved to β-sheet range in the Ramachandran map, the hydrophobicity of the side chain could be seen to show only a slight change.

Table 2. The variance of the local hydrophobicity scales of Bradykinin potentiating pentapeptide side chains

Residue	$\log P_{gos}^{cal\ a}$	$\log P_{sol}^{cal\ b}$
Ala	-0.540 ~ -0.550	-0.536 ~ -0.552
Arg(+)	2.522 ~2.533	2.383 ~ 2.700
Asn		
Asp(-)		
Cys		
Glu(-)	3.942 ~ 3.959	3.779 ~ 4.193
Gln		
Gly	0.000	0.000
His	-0.773	-0.811
Ile		
Leu	-1.936	-1.931 ~-1.948
Lys(+)	1.362 ~ 1.396	1.167 ~ 1.543
Met		
Phe	-2.682 ~ -2.700	-2.672 ~ -2.703
Pro	-0.702 ~ -0.714	-0.696 ~ -0.728
Ser	-0.188 ~ -0.192	-0.169 ~ -0.200
Thr	-0.638	-0.655
Trp	-2.834 ~ -2.858	-2.807 ~ -2.876
Tyr	-2.251 ~ -2.261	-2.257 ~ -2.280
Val	-1.518 ~ -1.527	-1.515 ~ -1.536

[a] The calculated hydrophobicity values obtained from the local side chain region conformation of Bradykinin potentiating pentapeptide that are minimized in the gas phase using ECEPP/NKS potential and SUMSL minimization algorithm. [b] The calculated hydrophobicity values obtained from the local side chain region conformation of Bradykinin potentiating pentapeptide that are minimized in solution using ECEPP/NKS potential and SMSNO minimization algorithm. The solvation free energy term is added to the total energy.

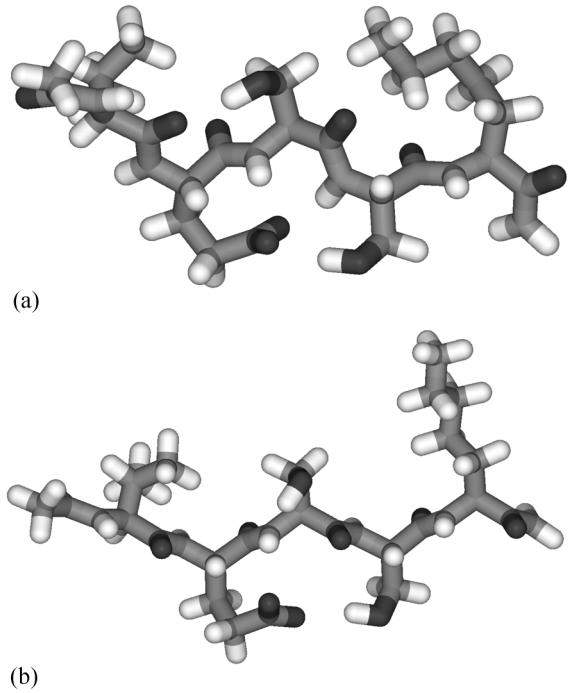

Figure 5. The minimized structures of Bradykinin potentiating pentapeptide (V-E-S-S-K) (a) in the gas phase and (b) in solution environment.

Figure 6 shows the minimized conformation of a peptide, with the sequence of VESSK, in the gas phase and in the solution phase. Table 3 shows the initial and minimized angels of the ϕ, ψ, and ω angles in both phases. The electrostatic energy of the VESSK peptide calculated using

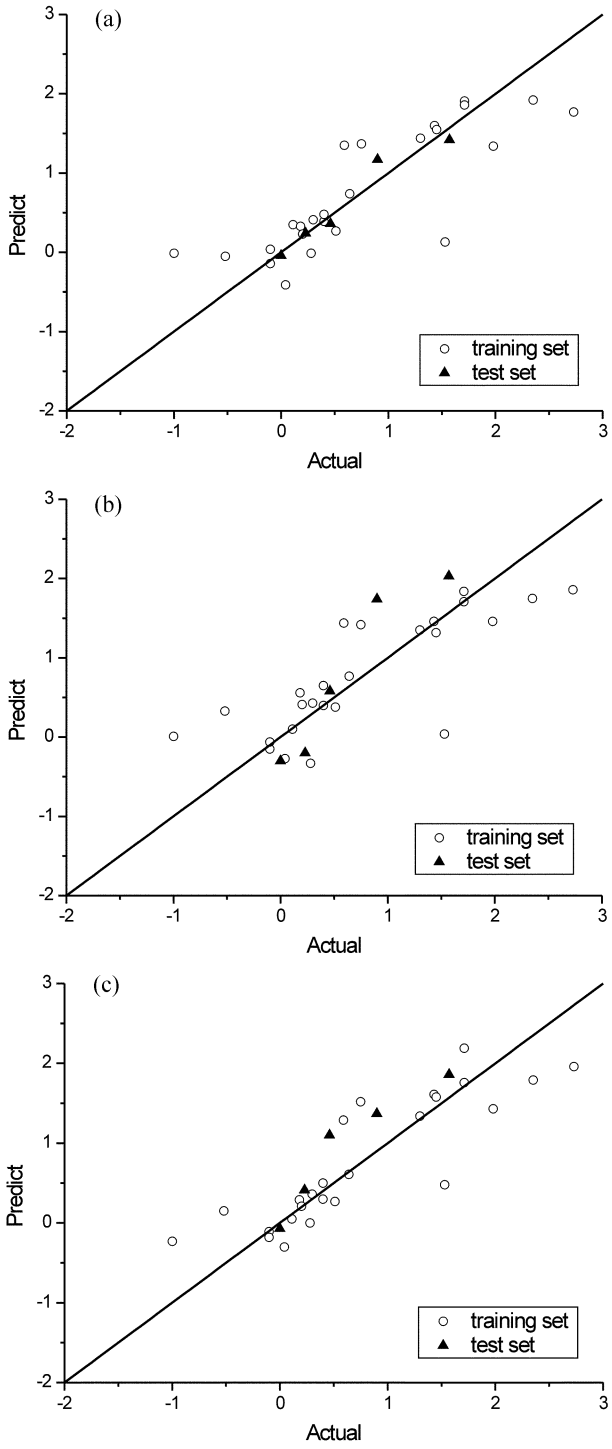

Figure 6. Actual vs. predictive activity plots of Bradykinin potentiating pentapeptide of 25 training set and 5 prediction set (a) with experimental hydrophobicities (b) with not considered the conformation of pentapeptides (c) with considered the conformation of pentapeptides, both in solution environment and MLR model.

Table 3. The torsional angles (in degree) of VESSK pentapeptides

sequence	initial dihedral angle			minimized dihedral angle (gas phase)			minimized dihedral angle (aqueous phase)		
	ϕ	ψ	ω	ϕ	ψ	ω	ϕ	ψ	ω
V	180.0	180.0	180.0	-159.9	147.1	177.4	-172.2	-169.9	179.9
E	-120.0	120.0	180.0	179.2	124.6	-179.4	-123.0	123.1	-176.4
S	-120.0	120.0	180.0	-163.6	62.0	172.1	-123.6	118.1	178.6
S	-120.0	120.0	180.0	-76.8	159.7	175.4	-123.4	121.2	177.8
K	-120.0	119.2	-174.0	-139.4	69.6	179.1	-121.1	119.7	-175.2

the ECEPP/NKS, in both the gas and solution phases was 100.52 kcal/mol and 48.86 kcal/mol, respectively. Table 2 shows a list of the hydrophobicity scales for local side chains of the Bradykinin potentiating pentapeptide. The omitted hydrophobicity values of the amino acid are not included in the Bradykinin potentiating pentapeptide data set. The variation of hydrophobicity values in the solution phase is larger than that of the calculated hydrophobicity values in the gas phase. The hydrophobicity scales of the charged amino acid side chains (Arg^+, Glu^-, Lys^+) increased compared to the calculated values of the side chain of *N*-acetyl amino acid amide side chains. Furthermore, all the hydrophobicity values shifted. Proline normally has a hydrophobic character, $\log P_{sol}^{cal}$ value is -1.198. However proline in the Bradykinin potentiating pentapeptide decreased a hydrophobic character. the local hydrophobicity scales is from -0.696 to -0.728. This is due to the fact that proline dependently varied the environment.

The hydrophobicity scale of a side chain would vary according to the conformational variation of a pentapeptide, since the hydrophobicity scales are derived from the solvation and cavitation free energy calculations. For the same reason, it is obvious that the hydrophobicity of a peptide could be related to the contributions made by each residue in the peptide, such as the surface area and environmental solvent. Roseman[43] shows that the chemical environment of each side chain depends on its neighbors. Wimley *et al.*[44] experimentally determined n-octanol/water solvation free energies of amino acid side chains and backbone and derived atomic solvation parameters (ASP) in pentapeptide models. They also examined the effects of conformational flexibility of the polypeptide chain and the neighboring side chains.

The QSAR Application for Peptide. A set of Bradykinin potentiating pentapeptide, previously analyzed using the *z*-scales and ISA-ECI descriptors, was reported.[45] The Bradykinin potentiating activity of the peptides was determined on the isolated guinea pig ileum and expressed as the relative activity index.[46] This index is the ratio of the molar concentration of VESSK and the molar concentration of the peptide under the test in relation to an equivalent Bradykinin potentiation. The amino acid sequences of the peptides and the Bradykinin potentiating activity are listed in Table 3. These Bradykinin potentiating pentapeptide data set were was employed in the QSAR study. The number of observations (rows) is larger than the number of independent variables able to be analyzed using multiple linear regressions. The QSAR results are listed in Table 3.

The hydrophobicity scale was the only descriptor for the QSAR study for peptide which represented the relative electrostatic and steric effect between each of the amino acids. The *z*-scale and ISA-ECI descriptors are described in these terms. For peptides encoded using ISA-ECI and Z_1-Z_2-Z_3 descriptors, PLS and cross-validation methods were used to construct QSAR equations. Table 5 shows the predictability of the QSAR study using the calculated experimental data,[10] and local side chain hydrophobicity descriptors. The hydrophobicity scales of *N*-acetyl amino acid amide side chains which are not considered representative of the side chainside chain interaction effects, are used in the Bradykinin potentiating pentapeptides training set. The 'local' hydrophobicity scales of pentapeptides are considered to represent the side chain – side chain interaction effects of the whole structure of Bradykinin potentiating pentapeptides (See Table 2). In order to obtain insight into the contribution made by the side chain – side chain interaction effect of peptide chains, local hydrophobicity was also applied to this training set. In this case, the hydrophobicity scales do not have a specific value, but instead, fall within a particular range of values. These values vary and the variance is caused by neighboring amino acids. Each variable consists of 15 descriptor values with *Z* descriptors, or 10 descriptor values with ISA-ECI descriptors [See references 11 and 14 for details]. However, in this method, only the hydrophobicity descriptor for only one amino acid is used. The good relationship between observed and predicted activity for the training set is shown in Table 4, Figure 8 and Figure 9. In Table 4, fixed hydrophobicity means the descriptor is obtained from the hydrophobicity scale of single conformation of amino acids as shown in Table 1 and variable hydrophobicity means the descriptors are obtained from the hydrophobicity scale of various conformations of amino acids as shown in Table 2. The test set, which is composed of 5 compounds, is used to prove the prediction ability of this QSAR model (Table 5). The test compounds are selected from every 5th set of activity data in order in 30 pentapeptide compounds. Results obtained from the QSAR calculation using our hydrophobicity descriptors are as good as those obtained using the ISA-ECI and *Z*-scale descriptors (See also Table 6). The results of the QSAR calculation using the hydrophobicity descriptors calculated from the gas and solution phases of the structure show similar results (See Table 6, Figure 8 and Figure 9). When the minimization was performed in the gas phase, the backbones of pentapeptides

Table 4. Observed and MLR predicted activity using Hydrophobicity descriptors for the training set

Peptide	$\log RAI^{exp}$ [a]	MLR prediction/Residual				
		fixed hydrophobicity[b]			variable hydrophobicity[c]	
		$\log P^{exp}$	$\log P^{cal}_{gas}$	$\log P^{cal}_{sol}$	$\log P^{cal}_{gas}$	$\log P^{cal}_{sol}$
VEGGK	-1.00	-0.01/-0.99	0.05/-1.05	0.01/-1.01	-0.25/-0.75	-0.23/-0.77
GEAAK	-0.52	-0.05/-0.47	0.32/-0.84	0.33/-0.85	0.15/-0.67	0.15/-0.67
VAAAK	-0.10	0.04/-0.14	-0.16/ 0.06	-0.15/ 0.05	-0.10/ 0.00	-0.11/ 0.01
AAAAA	-0.10	-0.14/ 0.04	-0.03/-0.07	-0.06/-0.04	-0.16/ 0.06	-0.18/ 0.08
VWAAK	0.04	-0.41/ 0.45	-0.27/ 0.31	-0.27/ 0.31	-0.30/ 0.34	-0.30/ 0.34
VKAAK	0.11	0.35/-0.24	0.09/ 0.02	0.10/ 0.01	0.06/ 0.05	0.05/ 0.06
VEAAP	0.18	0.33/-0.15	0.60/-0.42	0.56/-0.38	0.30/-0.12	0.29/-0.11
VEASK	0.20	0.23/-0.03	0.46/-0.26	0.41/-0.21	0.20/ 0.00	0.21/-0.01
VESAK	0.28	-0.01/ 0.29	-0.25/ 0.53	-0.33/ 0.61	0.02/ 0.26	0.00/ 0.28
FEAAK	0.30	0.41/-0.11	0.42/-0.12	0.43/-0.13	0.37/-0.07	0.36/-0.06
LEAAK	0.40	0.39/ 0.01	0.39/ 0.01	0.40/-0.00	0.31/ 0.09	0.30/ 0.10
RYLPT	0.40	0.48/-0.08	0.59/-0.19	0.65/-0.25	0.50/-0.10	0.50/-0.10
VEAAK	0.51	0.27/ 0.24	0.37/ 0.14	0.38/ 0.13	0.27/ 0.24	0.27/ 0.24
VELAK	0.59	1.35/-0.76	1.49/-0.90	1.44/-0.85	1.30/-0.71	1.29/-0.70
FSPFR	0.64	0.74/-0.10	0.84/-0.20	0.77/-0.13	0.61/ 0.03	0.61/ 0.03
AAWAA	0.75	1.37/-0.62	1.41/-0.66	1.42/-0.67	1.54/-0.79	1.52/-0.77
EKWAP	1.30	1.44/-0.14	1.38/-0.08	1.35/-0.05	1.32/-0.02	1.34/-0.04
VAWAA	1.43	1.60/-0.17	1.45/-0.02	1.46/-0.03	1.61/-0.18	1.61/-0.18
VAWAK	1.45	1.55/-0.10	1.28/ 0.17	1.32/ 0.13	1.58/-0.13	1.58/-0.13
VEHAK	1.53	0.13/ 1.40	-0.04/ 1.57	0.04/ 1.49	0.44/ 1.09	0.48/ 1.05
VKWAA	1.71	1.91/-0.20	1.70/ 0.01	1.71/ 0.00	1.77/-0.06	1.76/-0.05
VEWVK	1.71	1.86/-0.15	1.72/-0.01	1.84/-0.13	2.19/-0.48	2.19/-0.48
RKWAP	1.98	1.34/ 0.64	1.48/ 0.50	1.46/ 0.52	1.43/ 0.55	1.43/ 0.55
VKWAP	2.35	1.92/ 0.43	1.76/ 0.59	1.75/ 0.60	1.77/ 0.58	1.79/ 0.56
VEWAK	2.73	1.77/ 0.96	1.82/ 0.91	1.86/ 0.87	1.96/ 0.77	1.96/ 0.77

[a] Reported by Ufkes *et al.*[46] [b] The hydrophobicity values, $\log P^{cal}_{gos}$ and $\log P^{cal}_{sol}$ which are taken from Table 1 used for regression as descriptors which does not consider the conformational variation of pentapeptides. [c] The hydrophobicity values, $\log P^{cal}_{gos}$ and $\log P^{cal}_{sol}$ which are taken from Table 2 used for regression as descriptors which consider the conformational variation of pentapeptides.

Table 5. Observed and MLR predicted activity ($\log P^{exp}$) for the test set

Peptides	$\log P^{exp}$	MLR prediction / residual				
		Descriptor from Table 1 [b]			Descriptor from Table 2 [c]	
		$\log P^{exp}$	$\log P^{cal}_{gas}$	$\log P^{cal}_{sol}$	$\log P^{cal}_{gas}$	$\log P^{cal}_{sol}$
VESSK	0.00	-0.04/ 0.04	-0.16/ 0.16	-0.30/ 0.30	-0.06/ 0.06	-0.07/ 0.07
VAAWK	0.23	0.24/-0.01	-0.35/ 0.58	-0.20/ 0.43	0.41/-0.18	0.41/-0.18
AAYAA	0.46	0.36/ 0.10	0.99/-0.53	0.58/-0.12	1.10/-0.64	1.10/-0.64
PGFSP	0.90	1.17/-0.27	1.81/-0.91	1.74/-0.84	1.39/-0.49	1.37/-0.47
VEFAK	1.57	1.42/ 0.15	1.98/-0.41	2.03/-0.46	1.85/-0.28	1.86/-0.29

[a] Reported by Ufkes *et al.*[46] [b] The hydrophobicity values, $\log P^{cal}_{gos}$ and $\log P^{cal}_{sol}$ which are taken from Table 1 used for regression as descriptors which does not consider the conformational variation of pentapeptides. [c] The hydrophobicity values, $\log P^{cal}_{gos}$ and $\log P^{cal}_{sol}$ which are taken from Table 2 used for regression as descriptors which consider the conformational variation of pentapeptides.

were fixed at the β-sheet conformation. The conformations of the side chains were changeable for the pentapeptides in the gas phase. Therefore, the restriction of backbone conformation improves the QSAR results. In contrast, the QSAR results of the solution phase have an ascendancy over the gas phase results. The QSAR results obtained using 'local' hydrophobicity scales as a descriptor, are more agreeable than those using single peptide hydrophobicity scales as a descriptor, since this z-scale descriptor does not take into consideration the interactions occurring between side chains.

Conclusions

In theoretical works, it is important to show that the results obtained from the calculations agree well with experimental values. In this study, the calculated hydrophobicity of *N*-acetyl amino acid is compared with the experimental

Table 6. The result from the MLR model

Parameter	MLR prediction/residual				
	Descriptor from Table 1[a]			Descriptor from Table 2[b]	
	$\log P^{exp}$	$\log P^{cal}_{gas}$	$\log P^{cal}_{sol}$	$\log P^{cal}_{gas}$	$\log P^{cal}_{sol}$
R^2	0.690	0.616	0.635	0.747	0.750
F-test	8.471	6.091	6.618	11.201	11.404
PRESS	10.016	11.862	11.074	7.760	7.704
Q^2	0.502	0.411	0.450	0.614	0.617

[a] The hydrophobicity values, $\log P^{cal}_{gas}$ and $\log P^{cal}_{sol}$ which are taken from Table 1 used for regression as descriptors which does not consider the conformational variation of pentapeptides. [b] The hydrophobicity values, $\log P^{cal}_{gas}$ and $\log P^{cal}_{sol}$ which are taken from Table 2 used for regression as descriptors which consider the conformational variation of pentapeptides.

hydrophobicity scales. Hydrophobicity is derived from theoretical free energy of transferring using the SFED model. The hydrophobicity scales of the peptide side chains in both the gas and solution phase are then calculated. The theoretical hydrophobicity scales obtained with the SFED model agree well with the experimental values. Based on this agreement, this model can then be applied to the hydrophobicity scales of peptide side chains. The proxy value of the hydrophobicity scales of the peptide side chain using *N*-acetyl amino acid amides were compared with the hydrophobicity scales obtained from the actual peptide structure using a QSAR model. The QSAR results for hydrophobicity scales of *N*-acetyl pentapeptide amides are better than those for the hydrophobicity scales of *N*-acetyl amino acid amides. The side chain hydrophobicity scales effect the solution environment and side – chain side chain interaction, and function well as a descriptor for predicting the activities of Bradykinin potentiationg peptides. These results show that the SFED model is reliable for hydrophobicity scales of peptide molecules and the hydrophobicity scales of amino acids are related to variations in side chain conformation in a peptide. It is evident that the proper use of hydrophobicity scales is a good representation on the study of peptides.

Acknowledgements. This work was supported by the Korea Research Foundation Grant. (KRF-2001-004-D00014).

References

1. Van de Waterbeemd, H.; Testa, B. *Advances in Drug Research*; Academic Press: London, 1987; Vol. 16, pp 87-227.
2. Levitt, M. *J. Mol. Biol.* **1976**, *104*, 59.
3. Hopp, S.; Woods, K. S. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 3824.
4. Kyte, J.; Doolittle, R. *J. Mol. Biol.* **1982**, *157*, 105.
5. Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199.
6. Ooi, T.; Oobatake, M.; Nemethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086.
7. Yoshida, T. *J. Chromatogr. A* **1998**, *808*, 105.
8. Nozaki, Y.; Tanford, C. *J. Biol. Chem.* **1971**, *246*, 2211.
9. Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. *Biochemistry* **1981**, *20*, 849.
10. Fauchère, J. L.; Pliška, V. *Eur. J. Med. Chem.* **1983**, *4*, 369.
11. Shin, P.; Pedersen, L. G.; Gibbs, P. R.; Wolfenden, R. *J. Mol. Biol.* **1998**, *280*, 421.
12. Ward, D. J. *Peptide Pharmaceuticals*; Elsevier: New York, 1991; pp 1-17.
13. Tayor, M. D.; Amidon, G. *Peptide Based Drug Design; Controlling Transport and Metabolism*; American Chemical Society: Washington, DC, 1995.
14. Charton, M. *Prog. Phys. Org. Chem.* **1990**, *18*, 163.
15. DePriest, S. A.; Mayer, D.; Naylor, C. D.; Marshall, G. R. *J. Am. Chem. Soc.* **1993**, *115*, 5372.
16. Waller, C. L.; Oprea, T. I.; Giolitti, A.; Marshall, G. R. *J. Med. Chem.* **1993**, *36*, 4152.
17. Hellberg, S.; Sjöström, M.; Skagerberg, B.; Wold, S. *J. Med. Chem.* **1987**, *30*, 1126.
18. Sandberg, M.; Eriksson, L.; Jonsson, J.; Sjöström, M.; Wold, S. *J. Med. Chem.* **1998**, *41*, 2481.
19. Cocchi, M.; Johansson, E. *Quant. Struct.-Act. Relat.* **1993**, *12*, 1.
20. Collantes, E. R.; Dunn, W. J., III. *J. Med. Chem.* **1995**, *38*, 2705.
21. No, K. T.; Kim, S. G.; Cho, K. H.; Scheraga, H. A. *Biophysical Chemistry* **1999**, *78*, 127.
22. In, Y. Y.; No, K. T. to be published
23. Kang, Y. K.; No, K. T.; Scheraga, H. A. *J. Phys. Chem.* **1996**, *100*, 15588.
24. Kwon, O. Y.; Kim, S. Y.; No, K. T. *J. Phys. Chem.* **1996**, *100*, 17670.
25. Holm, L.; Sander, C. *J. Mol. Biol.* **1991**, *218*, 183.
26. Dennis, J. E.; Gay, D. M.; Welsch, R. E. *ACM Trans. Math. SOFTWARE* **1981**, *7*, 369.
27. Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361.
28. No, K. T.; Grant, J. A.; Scheraga, H. A. *J. Phys. Chem.* **1990**, *94*, 4732.
29. No, K. T.; Grant, J. A.; Jhon, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1990**, *94*, 4740.
30. Park, J. M.; No, K. T.; Jhon, M. S.; Scheraga, H. A. *J. Comput. Chem.* **1993**, *14*, 1482.
31. Park, J. M.; Kwon, O. Y.; No, K. T.; Jhon, M. S.; Scheraga, H. A. *J. Comput. Chem.* **1995**, *16*, 1011.
32. No, K. T.; Cho, K. H.; Jhon, M. S.; Scheraga, H. A. *J. Am. Chem. Soc.* **1993**, *115*, 2005.
33. No, K. T.; Cho, K. H.; Kwon, O. Y.; Jhon, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1994**, *98*, 10742.
34. No, K. T.; Kwon, O. Y.; Kim, S. Y.; Jhon, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1995**, *99*, 3478.
35. No, K. T.; Kwon, O. Y.; Kim, S. Y.; Cho, K. H.; Yoon, C. N.; Kang, Y. K.; Gibson, K. D.; Jhon, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1995**, *99*, 13019.
36. Ferreira, L. A. F.; Auer, H.; Haslinger, E.; Fedele, Chr.; Habermehl, G. G. *Toxicon* **1999**, *37*, 661.
37. Chipot, C.; Pohorille, A. *J. Phys. Chem. B* **1998**, *102*, 281.
38. Vásques, M.; Némethy, G.; Scheraga, H. A. *Macromolecules* **1983**, *16*, 1043.
39. Grant, J. A.; Williams, R. L.; Scheraga, H. A. *Biopolymers* **1990**, *30*, 929.
40. Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III.; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989.
41. Böhm, H. J.; Brode, S. *J. Am. Chem. Soc.* **1991**, *113*, 7129.
42. Gould, I. R.; Cornell, W. D.; Hillier, I. H. *J. Am. Chem. Soc.* **1994**, *116*, 9250.
43. Roseman, M. A. *J. Mol. Biol.* **1988**, *200*, 513.
44. Wimley, W. C.; Creamer, T. P.; White, S. H. *Biochemistry* **1996**, *35*, 5109.
45. Cho, S. J.; Zheng, W.; Tropsha, A. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 259.
46. Ufkes, J. G. R.; Visser, B. J.; Heuver, G.; Wynne, H. J.; Van der Meer, C. *European J. Pharmacol.* **1982**, *79*, 155.