Long Time Dynamics of Met-Enkephalin: Comparison of Explicit and Implicit Solvent Models

Min-yi Shen and Karl F. Freed

James Franck Institute and Department of Chemistry, The University of Chicago, Chicago, Illinois 60637 USA

ABSTRACT   Met-enkephalin is one of the smallest opiate peptides. Yet, its dynamical structure and receptor docking mechanism are still not well understood. The conformational dynamics of this neuron peptide in liquid water are studied here by using all-atom molecular dynamics (MD) and implicit water Langevin dynamics (LD) simulations with AMBER potential functions and the three-site transferable intermolecular potential (TIP3P) model for water. To achieve the same simulation length in physical time, the full MD simulations require 200 times as much CPU time as the implicit water LD simulations. The solvent hydrophobicity and dielectric behavior are treated in the implicit solvent LD simulations by using a macroscopic solvation potential, a single dielectric constant, and atomic friction coefficients computed using the accessible surface area method with the TIP3P model water viscosity as determined here from MD simulations for pure TIP3P water. Both the local and the global dynamics obtained from the implicit solvent LD simulations agree very well with those from the explicit solvent MD simulations. The simulations provide insights into the conformational restrictions that are associated with the bioactivity of the opiate peptide dermorphin for the δ-receptor.

INTRODUCTION

Biological "processes" are per force dynamical in nature. For example, the dynamical motions of certain sections of RNA molecules and enzymes determine their key biological functions. However, rather little theory is available at the atomistic level concerning the long time dynamical behavior of flexible natural and synthetic macromolecules. In principle, the long time dynamics of biomolecules are governed by the complete equations of classical mechanics that involve realistic atomistic potentials and that include the explicit presence of solvent molecules. Although it is impossible to solve these equations with realistic atomistic potential functions for a macroscopic system with 10^{23} degrees of freedom, the molecular dynamics (MD) method provides a computationally feasible way to investigate the N-body system on a miniature scale. However, even with increasing computing power, the evaluation of the long time ($>$10 ns) behavior of large scale biological systems becomes prohibitive. As a measure of the current state of the art, the use of the most powerful Cray T3E supercomputer has pushed the simulation time for the 36-residue villin headpiece (HP-36) up to a remarkable one microsecond trajectory (Duan and Kollman, 1998). In contrast, most current MD simulations for larger realistic systems are usually restricted to a few nanoseconds or picoseconds. Moreover, these atomistic MD simulations generate coordinates and velocities for all solvent and solute atoms and, therefore, produce overwhelmingly more information than is hopefully needed (Kostov and Freed, 1999). Thus, it is

important to develop physically significantly more efficient methods for determining the long time dynamical properties of biosystems. In the present paper, we develop an implicit solvent method that greatly reduces the required computational cost, and we test this method against full explicit solvent MD dynamics for the small flexible peptide Met-enkephalin. This development facilitates important investigations of the link between the peptide structure, dynamics, and biological function.

Greater efficiency in peptide dynamics computations is necessary to resolve many fundamental questions associated with understanding how a ligand that lacks a fixed shape recognizes several different molecular receptors and exerts multiple significant biological functions. Several structurally flexible proteins exhibit physiological significant roles. One example is provided by the β-amyloid protein, which causes the neurobiological disorder, Alzheimer's disease. The 42-residue peptide displays no dominant single structure, but its structure varies with temperature, solvent, and solution acidity (Barrow and Zagorski, 1991; Kirshenbaum and Daggett, 1995). In the appropriate physical environment, the β-amyloid aggregates into a β-sheet conformation and forms the plaque that is one consequence of the disease process. A recent investigation indicates that even the well-studied globular protein myoglobin can aggregate and form plaque under the appropriate conditions of temperature and pH (Fändrich et al., 2001). A first step of understanding the physical process behind these phenomena lies in developing theories to describe the dynamics of a single flexible peptide in aqueous solution.

Our implicit water dynamics method is tested for the dynamics of Met-enkephalin (Tyr-Gly-Gly-Phe-Met), which is one of the smallest neurotransmitter peptides. Since the first isolation and identification of this molecule in 1975 from pig brains (Hughes et al., 1975), it has been found to be plentiful in nerve terminals. Met-enkephalin has

---

Address reprint requests to Min-yi Shen, The James Franck Institute, 5640 S. Ellis Ave., Room 304, Chicago, IL 60637. Tel.: 1-773-702-7223; Fax: 1-773-702-5863; E-mail: smy@kffl.uchicago.edu.

been investigated extensively with many different techniques, including x-ray crystallography (Smith and Griffin, 1978), nuclear magnetic resonance (Roques et al., 1976; Deber and Behnam, 1984), and molecular simulations (Hruby et al., 1988; Wang and Kuczera, 1996). Enkephalins and endorphins all originate from the larger protein, β-lipotropin (Stryer, 1988), which produces Met-enkephalin, β-endorphin, and γ-endorphin. The five-amino acid sequence of Met-enkephalin is shared by all larger endorphins as their headpiece. The enkephalins and endorphins play diverse roles in biological systems, such as pain inhibition, aiding immune response, and producing tolerance and dependence (Plotnikoff et al., 1986). Inside the human body, there are μ, κ, and δ opiate receptors, to which the enkephalins and endorphins bind as ligands. Among these three types of receptors, the δ-receptor is more sensitive to the enkephalins. The enkephalins and endorphins are usually produced while the organism is under mental or physical stress. The ligand–receptor binding modifies the neural signal by changing the potassium ion conduction (Pasternak, 1988).

In accord with the general trend that short peptide chains do not adopt a single native structure, experimental studies show that Met-enkephalin does not adopt a single conformation in aqueous solution (Graham et al., 1992). The first computation of an energetically minimized structure for Met-enkephalin in vacuo using a Monte Carlo searching algorithm is based on the empirical conformational energy program for peptides model potentials (Isogai et al., 1977; Paine and Scheraga, 1985). Later computations using ECEPP/2 potential function and a simple implicit water model do not predict a single dominant structure either (Li and Scheraga, 1987). A combined multidimensional NMR and molecular modeling study (Graham et al., 1992) likewise indicates the absence of a distinguishable secondary structure for Met-enkephalin in aqueous solution. When dissolved in 50 mM of sodium dodecyl sulfate micelles, Met-enkephalin adopts a β-turn structure that is stabilized by a phenyl–phenyl hydrophobic interaction. Despite these findings, many aspects of both the structure and dynamics of Met-enkephalin in aqueous solution remain unresolved.

The next section describes the MD and Langevin dynamics (LD) methods that are applied to treat Met-enkaphalin dynamics with explicit and implicit solvent models, respectively, using the AMBER95 (Cornell et al., 1995) force field. A variety of implicit solvent models is tested by comparing the explicit and implicit solvent dynamics without the use of adjustable parameters. MD simulations are performed to estimate the viscosity of the three-site transferable intermolecular potential (TIP3P) (Jorgensen et al., 1983) water model that is used in the explicit solvent MD simulations because this viscosity is a necessary input parameter for the implicit solvent treatment of friction coefficients. In Results and Discussions, the various equilibrium structural properties generated by the explicit water MD

simulations are compared with those obtained from a set of different implicit water LD simulations. The local and global time-correlation functions obtained from these two types of simulations are compared. The only variables involved in the comparison are the choices of the solvation potential, the solvent viscosity, and dielectric constant, but these are prescribed by various methods from experiment or simulations. A subsequent paper will apply our mode-coupling theory of long time peptide dynamics (Kostov and Freed, 1999) to Met-enkephalin for comparison with both the MD and LD simulations. (M.-y. Shen and K. F. Freed, in preparation).

COMPUTATIONAL DETAILS

Explicit water simulations

The explicit water MD simulation is based on a total system energy that is the sum of different types of individual contributions,

$$U_{\text{tot-exp}} = U_{\text{b}} + U_{\text{bend}} + U_{\text{imp-tors}}$$
$$+ U_{\text{tors}} + U_{\text{ch}}(\epsilon = 1) + U_{\text{vdW}}, \qquad (1)$$

and that includes both solvent and solute molecules. The explicit water simulations are performed by using the standard AMBER95 force field and the TIP3P model for water. The simulation is carried out using a modified version of the TINKER 3.7 molecular design package (Ponder et al., 1999). The bonding (U_{b}) and bond–bond bending (U_{bend}) interactions are modeled by harmonic potentials, and the regular and improper torsional energies (U_{tors}, $U_{\text{imp-tors}}$) are described with the standard periodic functions. The charge–charge interaction is expressed in terms of atomic partial charges q_{i},

$$U_{\text{ch}}(\epsilon) = \sum_{\text{i}>\text{j}} \frac{q_{\text{i}}q_{\text{j}}}{\epsilon r_{\text{ij}}}, \qquad (2)$$

where ϵ is a fixed dielectric constant that represents the screening of Coulomb interactions in liquid water, and r_{ij} is the interparticle distance between atomic charges i and j. In the explicit water MD calculations, ϵ is set to unity for all charge–charge interactions. The van der Waals energy (U_{vdW}) is evaluated by using Lennard–Jones 6-12 potentials.

The MD simulations are generated by integrating the atom positions and velocities by using the standard velocity Verlet algorithm (Allen and Tildesley, 1987),

$$\mathbf{r}_{\text{i}}(t + \Delta t) = \mathbf{r}_{\text{i}}(t) + \mathbf{v}_{\text{i}}(t)\Delta t + \tfrac{1}{2}\mathbf{a}_{\text{i}}(t)(\Delta t)^2, \qquad (3)$$

$$\mathbf{v}_{\text{i}}(t + \Delta t) = \mathbf{v}_{\text{i}}(t) + \mathbf{a}_{\text{i}}(t)\Delta t, \qquad (4)$$

where $\mathbf{r}_{\text{i}}(t)$, $\mathbf{v}_{\text{i}}(t)$, and $\mathbf{a}_{\text{i}}(t)$ denote the particle positions, velocities, and accelerations, respectively, at time t. The integration time step is chosen as $\Delta t = 2$ fs. The simulated

system includes 186 solvating TIP3P water molecules in a cubic, periodic boundary box with 18.6-Å sides. Four 5-ns and two 10-ns MD trajectories (total duration 40 ns) are used to compute the time correlation functions and equilibrium averages. An 8-Å nonbonding force cutoff is applied, and the temperature is controlled at 300 K using a Berendsen-type thermal bath coupling (Berdensen et al., 1984).

Implicit water simulations

The total energy expression for the implicit solvent model differs slightly from that of the explicit water model because the explicit solvent–solute and solvent–solvent interactions in Eq. 1 are replaced by dielectric screening ($U_{ch}(\epsilon)$) and solvation terms ($U_{solv}(\sigma)$),

$$U_{tot-imp} = U_b + U_{bend} + U_{imp-tors}$$
$$+ U_{tors} + U_{ch}(\epsilon) + U_{vdW} + U_{solv}(\sigma). \quad (5)$$

Unlike the explicit water simulation, $U_{tot-imp}$ for the implicit water treatment only depends on the solute coordinates. The U_b, U_{bend}, $U_{imp-tors}$, U_{tors} and the solute–solute portion of U_{vdW} are identical for both the explicit water and the implicit water models. Two separate dielectric screening constants are tested for the solute–solute interactions described by U_{ch}, namely, $\epsilon = 78.5$, the experimental dielectric constant of water at 300 K, and $\epsilon = 53.1$, the lowest theoretical value for simple point charge (SPC) water at 300 K (Smith and van Gunsteren, 1994). The choice of $\epsilon = 53.1$ provides an estimated lower bound to the TIP3P dielectric constant because of the similarity between the TIP3P and SPC models for water. The AMBER95 parameter set is also used for all solute–solute interactions.

A significant distinction between the explicit and implicit water models is the presence of the macroscopic solvation potential $U_{solv}(\sigma)$. Two solvation potential functions are utilized, the atomic solvation parameter (ASP) (Eisenberg and McLachlan, 1986; Wesson and Eisenberg, 1992) and the Ooi–Scheraga solvent-accessible surface area (SASA) (Ooi et al., 1987) potentials, although a third solvation potential, the generalized Born-surface area (GB/SA) potential (Qiu et al., 1997), is found to be wholly unsatisfactory in performance and calculated results. In both solvation potential treatments, a contact free energy is evaluated in terms of the accessible surface areas (ASA) σ_i of all atoms i in the peptide. Thus, the solvation free energy is a sum of individual atom contributions,

$$U_{solv} = \sum_{i=1}^{N} g_i \sigma_i, \quad (6)$$

where the empirical atom solvation energy parameters g_i have been determined by fitting experimental aqueous solvation free energies of amino acids and selected organic compounds to Eq. 6. Because our simulation uses the TIP3P model for water, the empirical constants g_i determined for real water may not be completely adequate for the TIP3P model water, producing one unavoidable inherent limitation to our tests of implicit water simulation models. A more thorough test would require the fitting of the g_i parameters to simulations performed for many amino acid–TIP3P water systems, but the amount of work involved would be enormous and the value of this effort would only be in providing a more stringent test of our theory. Obviously, of much greater importance are improved solvation potentials that provide superior fits to experiment. In this regard, another proposed atomic solvation parameter optimization procedure (Baysal and Meirovitch, 1997, 1998) uses experimental native structures in solution as a motif and adjusts the solvation potential parameters to the global energy minimum. This method is a promising alternative for future reparameterizations of experimental data. Except for the hydrogen atoms, whose presence is ignored in computing the σ_i (Eisenberg and McLachlan, 1986), the solvent accessible surface area σ_i is calculated from the exposed surface area of intersecting spheres centered on each atom, using a probe sphere of radius 1.4 Å, which corresponds to the effective radius of a water molecule. Thus, when a pair of peptide atoms are too close for a solvent molecule to fit between them, no solvation free energy is applied to their contact energies; only the van der Waals and electrostatic interactions remain. The combination of these two features contributes to the description of the hydrophobic effect. Of the two sets of ASP parameters available in literature, we use the set determined by Wesson and Eisenberg (1992). The SASA method of Ooi et al. (1987) uses solvation free energy determined by a similar strategy but applied to a slightly larger classification of atom types, and, hence, the SASA model produces a larger number of parameters in the set $\{g_i\}$.

In contrast to the explicit water model where the dynamics of both solvent and solute atoms must be considered and where the solute constitutes roughly 10% of the overall atoms, the implicit water model only describes the motion of the solute atoms. Thus, the evaluation of $U_{tot-imp}$ is much faster than that of $U_{tot-exp}$. Our tests are designed to determine whether the much faster implicit solvent simulations accurately reproduce the far more expensive explicit water MD simulation.

Oliva et al. (2000) have also used the SASA solvation potential model to compute forces within an implicit solvent generalized Langevin equation (GLE) model. Although this feature coincides with one aspect tested by us, their approach describes the frictional effects of the solvent using memory functions whose computation slows down the GLE computations with respect to those with an explicit SPC water simulation by a factor of seven. When the memory function is taken as fixed (from a prior computation), their implicit GLE treatment is only a factor of three faster than

the explicit water simulation. In contrast, we use simple friction coefficients in our implicit solvent approach, enabling us to achieve orders of magnitude speed ups of the implicit solvent LD simulations compared to explicit solvent (TIP3P water) MD simulations. Another difference between our work and that of Oliva et al. (2000) arises because they focus on simulations for an equilibrium property, the native structure of a folded protein, whereas we compare the explicit and implicit solvent models for both equilibrium and dynamical properties of a highly flexible peptide. A future work will describe the application of our methods to study the stability of the native structure of the villin headpiece, where, again, both equilibrium and dynamical properties can be computed and the speed enhancement over explicit water simulations exceeds a couple of orders of magnitude (M.-y. Shen and K. F. Freed, submitted for publication).

Our main modifications to the TINKER package are associated with the computation of friction coefficients and the solvation potential. The friction coefficients are either set as constants throughout the LD trajectory or can be updated periodically with the procedure suggested by Pastor and Karplus (1988). The velocity Verlet algorithm (Allen and Tildesley, 1987) is also used in the LD calculations. The LD algorithm is similar to the algorithm given by Eqs. 3 and 4 for MD simulations but includes factors c_0, c_1, and c_2, arising from the frictional forces due to the implicit water molecules, and the random term necessary by virtue of the fluctuation-dissipation theorem to assure the proper equilibrium limit at long times,

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + c_1\mathbf{v}_i(t)\Delta t + \tfrac{1}{2}c_2\mathbf{a}_i(t)(\Delta t)^2 + \mathbf{r}_{gi}, \quad (7)$$

$$\mathbf{v}_i(t + \Delta t) = c_0\mathbf{v}_i(t) + c_1\mathbf{a}_i(t)\Delta t + \mathbf{v}_{gi}, \quad (8)$$

with

$$c_{0i} = e^{-\zeta_i\Delta t/m_i}, \quad (9)$$

$$c_{1i} = (-\zeta_i\Delta t/m_i)^{-1}(1 - c_{0i}), \quad (10)$$

$$c_{2i} = (-\zeta_i\Delta t/m_i)^{-1}(1 - c_{1i}), \quad (11)$$

where m_i is the mass of atom i. The $\mathbf{r}_{gi}$ and $\mathbf{v}_{gi}$ are Gaussian random variables with variances depending on the friction coefficients in the usual fashion (Allen and Tildesley, 1987). The friction coefficients ζ_i used in the Langevin dynamics simulation are determined from the solvent-accessible surface area σ'_i with zero probe radius. Thus, we calculate the friction coefficient ζ_i using stick boundary conditions,

$$\zeta_i = 6\pi\eta r_{eff;i}, \quad (12)$$

where η is the solvent viscosity and $r_{eff;i}$ is the effective hydrodynamic radius of atom i that is computed from σ'_i, using $r_{eff;i} = (\sigma'_i/4\pi)^{1/2}$.

We contrast two types of LD simulations, called the dynamical and static LD simulations. In the dynamical LD simulations, the accessible surface area, and, hence, the atomic friction coefficients, are updated every 100 dynamical steps, whereas, in the static LD simulations, the friction coefficients remain constant throughout the simulation. The solvation potential is updated every 100 dynamical steps (0.25 ps in physical time), which speeds up the LD simulations by a factor of 4 over LD simulations with the solvation potentials recomputed at each step. To test the adequacy of updating the solvation potential every 100 steps, we have generated a 30-ns test trajectory for a case in which the solvation potential has been calculated at every step. These two procedures lead to nearly identical distributions for the radius of gyration and sufficiently similar end-to-end vector dynamics. Appreciable enough changes appear in the atomic solvent-accessible surface areas and, hence, in the solvation-free energy only when major structural deformations occur in solute. The 0.25-ps interval between updates of the solvation potential is quite short compared to the time scales for the occurrence of large structural changes during the Met-enkephalin dynamics. No tests have been performed to determine the maximum acceptable update interval, but larger intervals (and faster simulations) should be possible.

The constant friction coefficients for the static LD simulations are computed by averaging ASA friction coefficients over a 6-ns sample trajectory obtained after equilibration. The integration time step Δt for the LD simulation is chosen as $\Delta t = 2.5$ fs. Eight LD trajectories of 60 ns each (total duration of trajectories is 480 ns) are used for the evaluation of time-correlation functions (TCF) and equilibrium averages. No nonbonding force cutoff is applied in the LD simulations. The LD algorithm is used instead of a pure van Gunsteren–Berendsen Brownian dynamics (BD) algorithm (van Gunsteren and Berendsen, 1982) because the BD algorithm applies only for very high atomic friction coefficients, whereas some of the hydrogen atomic ASA friction coefficients are small. Therefore, using the BD algorithm in this case would severely shorten the feasible integration time step to less than 0.1 fs in our simulations where hydrogen atoms are explicitly treated.

Determination of TIP3P water viscosity

The viscosity of the surrounding medium is an essential parameter for calculating the atomic friction coefficients (see Eq. 12). In pure BD simulations, the viscosity only affects the overall time scale for all of the dynamics. Therefore, the equilibrium time correlation functions for different choices of the solvent viscosity may readily be interconverted simply by rescaling the time. However, the dependence on the solvent viscosity is more complicated for the LD simulations in which inertial forces are retained in the equations of motion. Below, we investigate the extent to

which LD dynamics with different viscosities may be inter-converted by the time-rescaling procedure available for the BD dynamics. To compare the LD and full MD simulations, the shear viscosity for the TIP3P water is required. Because this viscosity is unavailable for the conditions used in our MD simulations, the viscosity is calculated from a MD simulation using a system containing 216 TIP3P water molecules in a cubic box with a side of 18.6 Å and periodic boundary conditions. This system has the same temperature and water density (1.003 g/cm^3) as in the Met-enkephalin MD simulations.

The standard simulation methods for bulk systems represent the shear viscosity in terms of the off-diagonal components of the pressure tensor,

$$P_{\alpha\beta}(t) = \frac{1}{V} \sum_{i=1}^{N} \left(\frac{p_{i\alpha}p_{i\beta}}{m_i} + r_{i\alpha}(t)f_{i\beta}(t) \right), \quad (13)$$

where V is the total volume, and m_i, $r_{i\alpha}$, $p_{i\alpha}$, and $f_{i\alpha}$ are the masses, the components ($\alpha = x, y, z$) of the particle positions, momenta, and forces acting on atom i, respectively. We determine the shear viscosity of TIP3P water from the Einstein–Helfand relation (Helfand, 1960),

$$\eta = \frac{V}{2k_BT} \lim_{t\to\infty} \frac{d}{dt} \left\langle \left(\int_0^t P_{\alpha\beta}(t') \, dt' \right)^2 \right\rangle$$

$$= \frac{V}{2k_BT} \lim_{t\to\infty} \frac{d}{dt} \Delta G(t), \quad (14)$$

where the angular bracket $\langle \cdots \rangle$ denotes the equilibrium average

$$\langle \cdots \rangle = \frac{\int d\mathbf{r} e^{(-\beta U(r))}(\cdots)}{\int d\mathbf{r} e^{(-\beta U(r))}}. \quad (15)$$

Evaluating the viscosity from Eq. 14 is superior to that using the Green–Kubo expression (Zwanzig, 1965)

$$\eta = \frac{V}{k_BT} \int_0^\infty \langle P_{\alpha\beta}(0)P_{\alpha\beta}(t') \rangle \, dt' \quad (16)$$

because the numerical derivative of the square of the time integral of the equilibrium-averaged pressure-tensor element in Eq. 14 involves a function that is much better behaved than the integrand in the Green–Kubo expression. The presence of the infinite limit in Eq. 14 does not cause numerical difficulties because the derivative $d\Delta G(t)/dt$ converges quickly as time grows. In contrast, the Green–Kubo formulation requires the time integration of a fluctuating function in Eq. 16 when the upper limit of integral is large. Calculations of the time-correlation function in Eq. 16 have large uncertainties as t grows, making the Green–Kubo-type integral difficult to converge numerically. Use of a short time simulation deteriorates the situation due to the large

statistical errors in the time-correlation function. These errors are proportional to $\sqrt{\tau/T}$, where the τ is the correlation time for the time-correlation function in the integrand of Eq. 16 and T is the total duration of the simulations. Thus, when T is too short, a large error appears in the pressure time-autocorrelation function of Eq. 16 and renders the integral difficult to evaluate accurately. Because all-atom MD simulations require considerable computer time, the total simulation duration T is rather small.

The energy of the TIP3P water system is first energy minimized and then pre-equilibrated for 100 ps before running a 600-ps water MD simulation with a 2-fs step size. The three off-diagonal elements of the pressure tensor $P_{\alpha\beta}(t)$ are recorded and averaged at every step. The function $\Delta G(t)$ is computed by numerical integration and the viscosity is obtained from the slope of the integrated function in Eq. 14.

Description of different simulations

To assess the accuracy of various approximations inherent in the implicit water simulations, we have performed a large number of "mix and match" LD simulations. For brevity, the following refers to each set of simulations with the acronym ASPS to indicate simulations with LD dynamics/static friction/ASP solvation potential/experimental dielectric constant for water; ASPD to indicate LD dynamics/dynamic friction/ASP solvation potential/experimental dielectric constant for water; ASPE to indicate LD dynamics/static friction/ASP solvation potential/theoretical dielectric constant (53.1); SASA to indicate LD dynamics/static friction/SASA solvation potential/experimental dielectric constant for water; ASPL to indicate LD dynamics/static friction/ASP solvation energy/experimental dielectric constant for water/TIP3P water viscosity; NASP to indicate LD dynamics/static friction/no solvation potential/experimental dielectric constant for water; and MD to indicate the all atom molecular dynamics simulations.

RESULTS AND DISCUSSIONS

Parameter determination and computation efficiency

We begin by describing the simulations for the TIP3P water viscosity because this provides a central parameter for computing the friction coefficients. Figure 1 presents the equilibrium average of $\Delta G(t)$ from Eq. 13. The short time ($t < 0.3$ ps) behavior of this function is highly nonlinear, but the linearity of $\Delta G(t)$ emerges for $t > 0.3$ ps and leads to a nearly constant first derivative $d\Delta G(t)/dt$ over this time range. The average TIP3P water viscosity η evaluated from this derivative (for $t = 4$–10 ps) is 0.506 ± 0.043 cp, which is ~60.2% of the experimental value ($\eta_{exp} = 0.84$ cp) at 300 K and 1 atm. A similar discrepancy between the cal-

(a)

(b)

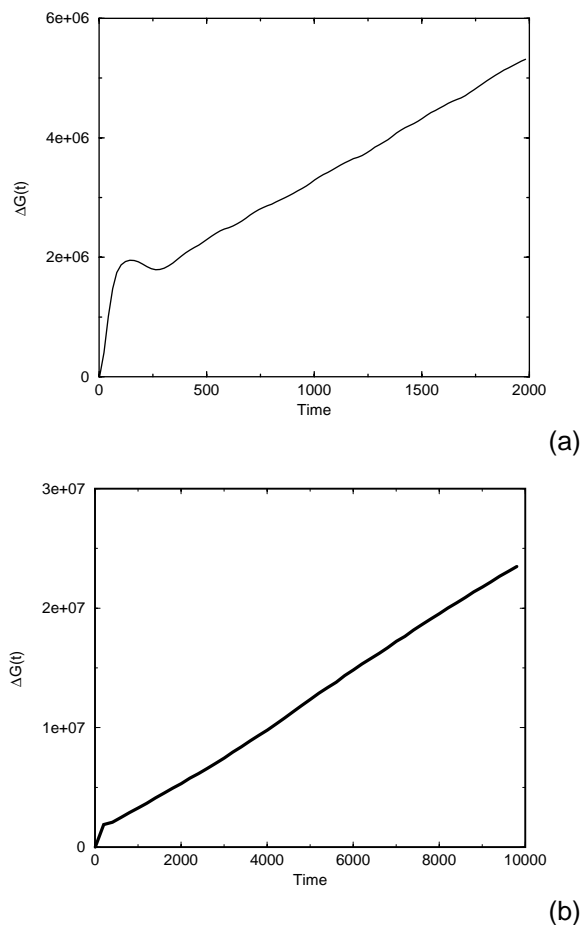FIGURE 1 $\Delta G(t)$ (in $g^2/(amu - ps - Å)^2$) evolution with time (in fs). (*A*) Short time behavior (0–2 ps). (*B*) Long time behavior (0–10 ps).

TABLE 1   The average atomic friction coefficients (divided by mass) (in ps^{-1}) and standard deviations for selected heavy atoms

Atom Number	Atom Type	Friction Coefficient	Standard Deviation
1	terminal N	73.29	0.80
2	C_α	32.31	1.54
3	backbone C	39.09	1.40
4	O	61.07	0.75
9	C_β	25.30	1.56
10	C_γ	35.03	1.36
12	C_δ	44.43	0.72
24	backbone N	51.15	1.80
25	C_α	35.61	1.17
26	backbone C	41.68	1.00
27	O	61.01	0.74
31	backbone N	50.74	1.45
32	C_α	36.03	1.06
33	backbone C	41.80	1.29
34	O	61.08	0.73
38	backbone N	47.13	1.32
39	C_α	24.08	1.51
40	backbone C	40.14	1.22
41	O	60.76	0.93
44	C_β	24.99	1.47
45	C_γ	34.94	1.29
46	C_δ	44.29	1.07
58	backbone N	48.06	1.59
59	C_α	23.30	1.57
60	terminal C	43.44	1.23
64	C_β	25.61	1.57
66	S	70.71	0.98

culated TIP3P water viscosity and experimental measurements is reported by Feller et al. (1996), who have estimated $\eta = 0.62 \pm 0.08$ cp, at $T = 293$ K whereas $\eta_{exp} = 1.0$ cp. Therefore, the TIP3P water viscosity of Feller et al. is also only 62% of the experimental value.

Table 1 summarizes the calculated average atomic friction coefficients, along with their standard deviations, that are evaluated for atoms in Met-enkephalin from a short 6-ns trajectory by using the LD simulation with ASP solvation potential. Figure 2 depicts the numbering used in Table 1 and descriptions below for heavy atoms in Met-enkephalin. The hydrogen atoms bonded to the nitrogen and oxygen atoms have vanishing solvent-accessible areas and, therefore, zero friction coefficients because they are embedded within the van der Waals radii of the central atoms. This is one reason for using the LD rather than BD simulations, because vanishing atomic friction coefficients are not permitted in BD simulations. The friction coefficients of most heavy atoms fluctuate by less than 3% during the 6-ns simulation trajectory, and only few of them exhibit larger variations (about 6.5%), especially for those atoms bonded to a pair of heavy groups, such as the C_α atoms of tyrosine,

phenylalanine, and methionine. The larger variation arises because an atom bonded to a pair of heavier groups experiences greater fluctuating forces that produce larger distortions in the local geometry and thereby affect the accessible surface area.

The implicit water LD simulations provide a huge saving in computational time over explicit solvent MD simulations on a single CPU machine. Most of the MD and LD calculations are performed with Intel Pentium III 750 MHz or AMD Athlon (K7) 800 MHz Linux systems. On the Intel machines, the typical elapsed times for a 60-ns LD simulation with implicit solvent is ~20 h, whereas the full MD
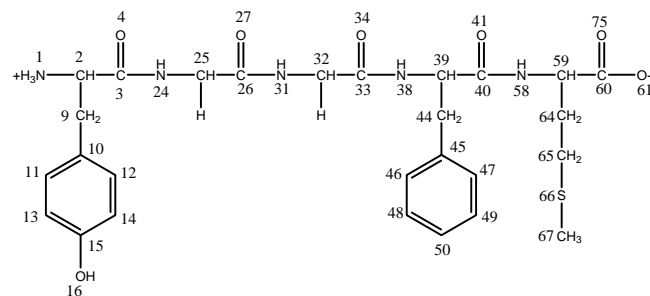

FIGURE 2   Heavy atoms numbering of Met-Enkephalin.

takes more than 660 h for just a 10-ns run. The AMD processors yield slightly better performance for the LD simulations and are expected to be identical to the Pentium III systems for the MD simulations. The total duration T of the trajectories is a crucial parameter determining the accuracy of computed time correlation functions $C(t)$, with the error in $C(t)$ decreasing as $T^{-1/2}$. Thus, to achieve comparable qualities of computed time correlation functions, a MD simulation of 60 ns would require 200 times the CPU time needed for a 60-ns LD simulation. Because the LD simulations are performed without a cut-off, whereas the MD simulations introduce a time-saving cut-off on nonbonding interactions, the ratio would grow substantially if a cut-off is not applied in the MD simulations.

The most time-consuming step in the dynamics simulations involves the evaluation of atomic forces. Combining the use of a cut-off on the nonbonded interactions with the use of analytical derivatives and Ewald summation methods implies that the computational time t for the explicit solvent MD calculations typically scales with the number of atoms N from $t \sim N$ to $t \sim N^2$. The implementation, for instance, of high performance MD codes, such as GROMACS (Berendsen, 2001), has reduced the computational time to scale nearly linearly with N. In a MD simulation of a fully solvated system, the simulation box normally has many more solvent molecules than solute atoms, and the number of solvent molecules increases as the cube of the box size. The minimum box size for an explicit solvent simulation should be at least as large as the average radius of gyration of the polymer molecule. Thus, when the polymerization index of a given polymer is increased by a factor of P, the minimum explicit water box size for the MD simulation should at least be multiplied roughly by $P^{1/2}$. Hence, the total number of solvent and solute atoms grows by a factor of roughly $P^{3/2}$. Even assuming an efficient linear scaling of t with N, the resulting scaling $t \sim P^{3/2}$ for the MD simulations unfavorably contrasts with the scaling $t \sim P$ for the LD simulations. Therefore, the obvious advantages of the implicit solvent model in treating the long time dynamics should be explored further for larger systems.

Distributions for structural properties

In this section, we compare the distributions for several structural properties that are calculated by averaging along the trajectories obtained from different simulation methods and solvation models. These equilibrium structural averages for the peptides are dictated by the model potential functions. For the implicit water simulations, we test two different dielectric constants and two macroscopic solvation potentials to mimic the influence of the solvent on the solute dynamics in the explicit water MD simulations. A comparison of the equilibrium structural averages calculated from the implicit water LD simulation and from the all-atom MD simulations provide the first portion of an optimal test (with

no extra adjustable parameters) for the adequacy of using the implicit solvation potentials to model the influence of the TIP3P water on Met-enkephalin dynamics.

One useful statistical measure for comparing different simulation methods is provided by the probability distribution for the radius of gyration (R_g) of Met-enkephalin,

$$R_g^2 = \frac{1}{N} \sum_{i=1}^{N} (r_i - r_g)^2, \qquad (17)$$

where r_g is the position of the center of gravity of Met-enkephalin, and the sum runs over all atoms in the solute. The square of the radius of gyration (R_g^2) is computed for each recorded time point for all dynamical trajectories, and histograms of R_g^2 are constructed by binning the values into 100 discrete intervals that range between 0 and 60 Å^2 for R_g^2. Figure 3 A–D, present the R_g^2 probability distributions computed from the MD, ASPS, SASA, and NASP simulations, and Fig. 3 E compares all four R_g^2 distributions. The maximum R_g^2 found for Met-enkephalin is ~60 Å^2, and certain aspects of the distributions can be understood from a more detailed examination of the Met-enkephalin trajectories. In particular, there are three classes of dominant configurations: the "extended" state in which R_g^2 lies in the range of 50–60 Å^2, the "semi-packed" state with $R_g^2 \sim 20$–40 Å^2, and the "packed" state where $R_g^2 \approx 15$–20 Å^2.

We now compare the range of variation of R_g^2 between the different simulations. Globally, R_g^2 oscillates between 15 and 55 Å^2 during the MD simulation, and the ASPS trajectories display a similar range for R_g^2. Inspection of the time evolution of R_g^2 obtained from the MD simulation indicates that Met-enkephalin stays in the packed and semi-packed states most of the time and only occasionally jumps to the extended state. The distribution for R_g^2 suggests the existence of a rapid interconversion between the packed and semi-packed states and a slower transition process between them and the extended state, a process that is verified by viewing the time course of LD trajectories and R_g^2. This physical picture is supported by the presence of two peaks at 21 and 27 Å^2 in Fig. 3 A and from the shoulder in the extended region near 40–50 Å^2. The latter feature implies a higher probability for being in the transition state leading to the more stretched configurations. The R_g^2 distribution extracted from the ASPS trajectories (see Fig. 3 B) differs from that illustrated in Fig. 3 A for the MD simulations by having a slightly lower occurrence frequency for the semi-packed state. The R_g^2 distribution from ASPS simulations is single-peaked as opposed to the double-peaked structure found for the MD simulations. The SASA simulations display a similar behavior but with broader structure and a shifted maximum in the R_g^2 distribution of Fig. 3 C. One peak is present in the SASA distribution with a maximum at 25 Å^2, which lies between the MD maxima of 21 and 27 Å^2. The SASA R_g^2 distribution is wider than that for the ASPS
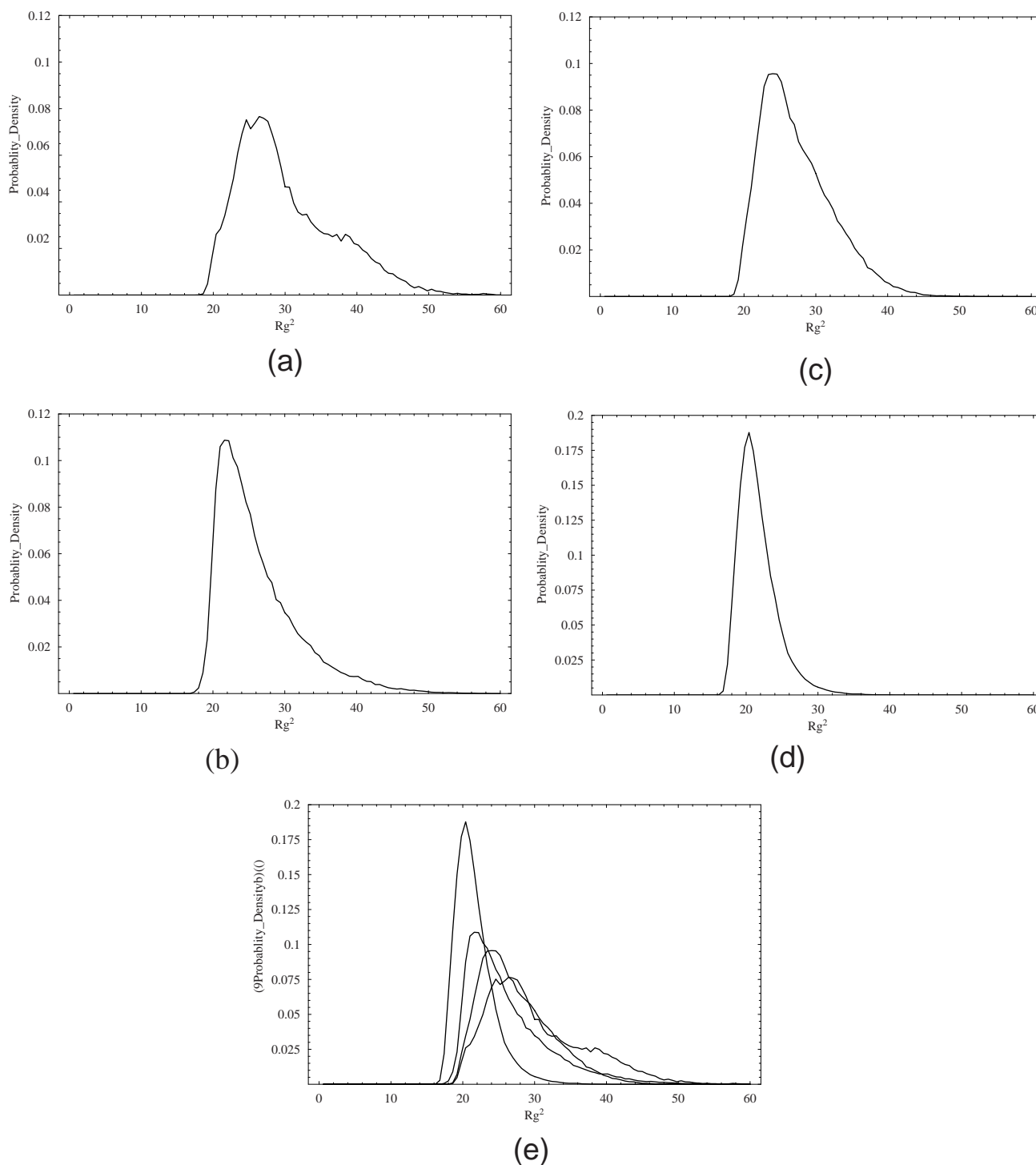
FIGURE 3 R_g^2 (in Å^2) distribution from different simulation methods. (*A*) Full atom MD. (*B*) ASP solvation potential. (*C*) SASA solvation potential. (*D*) No solvation potential. (*E*) A combined plot.

trajectory and has greater population for the extended state, in better agreement with the MD distribution. The SASA model uses a more detailed classification for types of units, and this feature apparently is responsible for the improved agreement of the LD simulation using the SASA solvation

parameter set with the MD simulations. The R_g^2 probability distribution calculated from a single 60-ns trajectory by using SASA model solvation potential also exhibits a double-peaked structure (which converts to a single peak after averaging over eight trajectories) that is similar to that of the
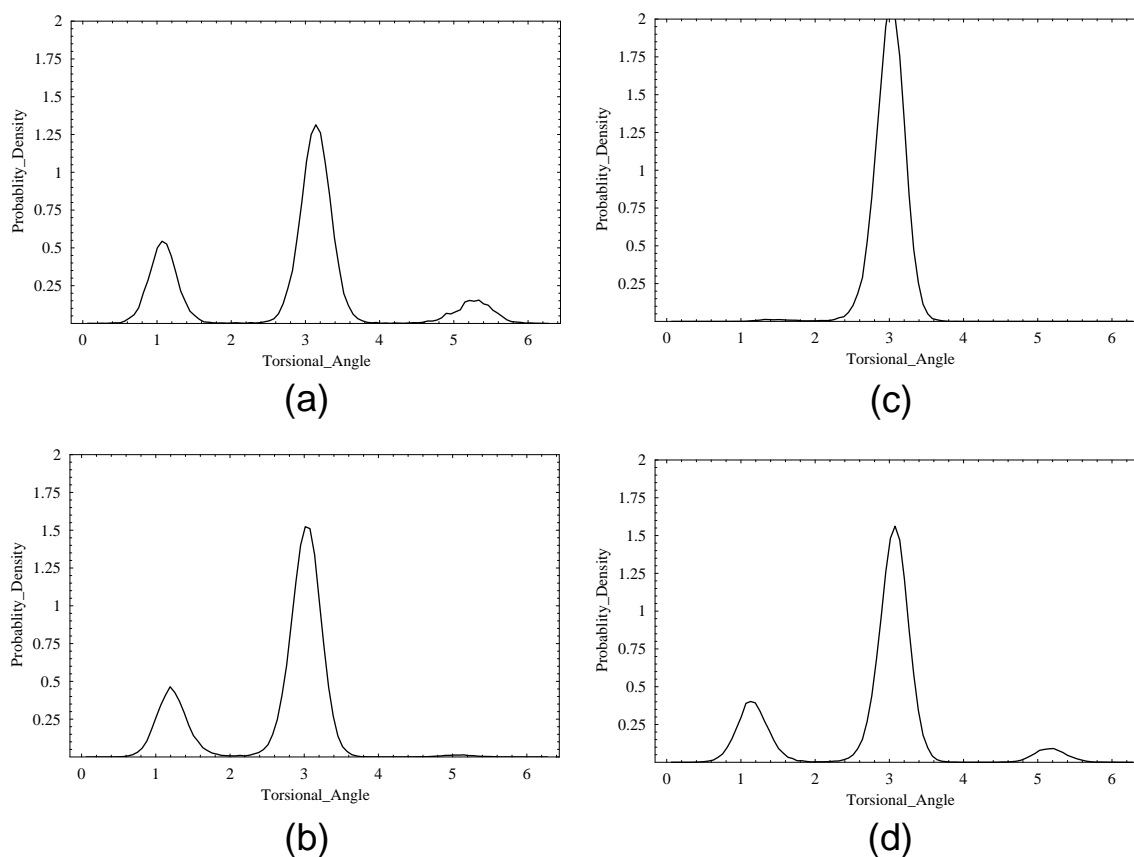
FIGURE 4   Tyr C_α–C_β torsional angle (in radians) distribution from different simulation methods. (*A*) Full atom MD. (*B*) ASP solvation potential. (*C*) SASA solvation potential. (*D*) No solvation potential.

MD simulations. This indicates that the double-peaked structure from the MD simulation may disappear for averages over longer trajectories.

The R_g^2 distribution from the NASP simulations is presented in Fig. 3 *D* and displays a narrower single-peaked structure with a restricted R_g^2 fluctuation range between 15 and 40 Å^2 that provides a poor representation of the MD distribution. The comparison of all the R_g^2 distributions in Fig. 3 *E* demonstrates the necessity for using a solvation potential so the LD simulations adequately approximate the R_g^2 distribution from the MD simulation. Because no extended state conformations appear in the NASP simulations, the addition of the solvation potentials helps stabilize the extended state configuration, which has a maximal charged-atom surface area that is stabilized by the solvation potentials. Moreover, the lack of appearance of extended conformations in the NASP simulations severely degrades its description of the long time dynamics, especially the global dynamics that are discussed in the next subsection.

Tourwé et al. (1996) have studied the conformational restrictions in the opioid peptide dermorphin that are necessary for the bioactivity of its binding to the δ receptor. The peptide must be in the *trans* conformation for the C_α–C_β bond of both the Tyr and Phe residues for maximum binding strength. The *gauche*($-$) conformation of Tyr greatly weakens the receptor docking affinity, decreasing the δ-receptor binding affinity by a factor of 300. The *gauche*($+$) conformers for both Tyr and Phe are likewise not favorable for δ-receptor binding. Fixing either of the Tyr or Phe residues in the *trans* conformation elevates the bio-activity for docking with the δ-receptor. Earlier computational studies conclude that the favored Tyr C_α–C_β conformer is *trans* in aqueous solution, whereas that of Phe is *gauche*($-$). Other experimental evidence supports these configurations for the opioid peptide in aqueous solution (Tourwé et al., 1996; Harms et al., 1998). Hence, Tourwé's work suggests that the solution structure of opioid peptides might differ from its optimal receptor bound form.

Figures 4 and 5 display the probability distribution for the torsional angles of the Tyr and Phe residues, respectively, as obtained from averaging over the trajectories with different simulation models. Every distribution in Figs. 4 and 5 exhibits a similar three-peaked structure, in which the peaks at $\pi/3$, π, and $5\pi/3$ correspond to the *gauche*($-$), *trans*, and *gauche*($+$) conformations, respectively. All the simulations yield qualitatively similar descriptions for the Tyr residue conformation: the dominant population is *trans*. The *gauche*($-$) conformer has only 23%, 26%, and 25% occur-
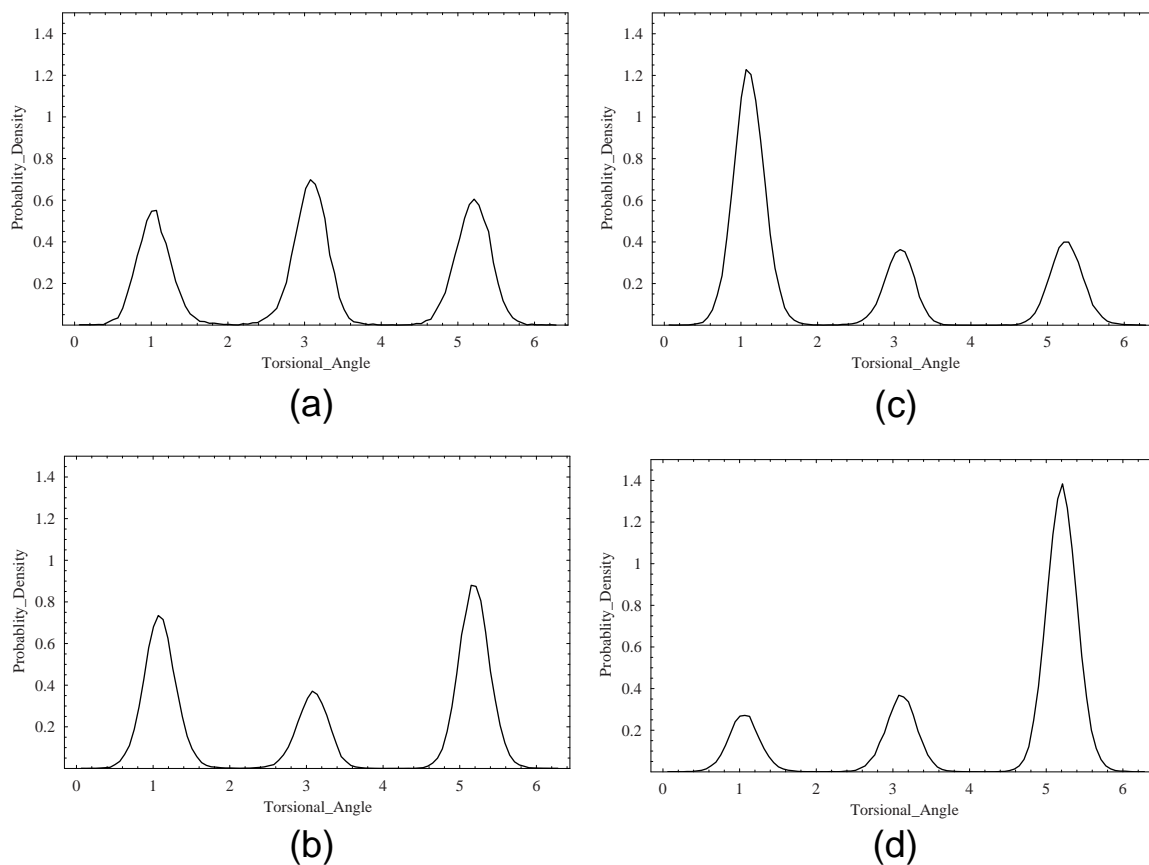
FIGURE 5 Phe C_α–C_β torsional angle (in radians) distribution from different simulation methods. (*A*) Full atom MD. (*B*) ASP solvation potential. (*C*) SASA solvation potential. (*D*) No solvation potential.

rences for the ASPS, NASP, and MD simulations, respectively, whereas the *gauche*(+) conformer only has a trace probability, usually less than 10%. The SASA solvation potential simulations predict the *trans* configuration as dominant with 99.8% of the total occurrence probabilities. An earlier simulation study by Hruby et al. (1988) suggests a dominant *gauche*(−) rotamer for Tyr, which disagrees with later computations obtained from a longer trajectory (1 ns) (Wang and Kuczera, 1996) and with our even longer simulations.

As shown in Fig. 5, all the LD simulations produce very different distributions for the Phe C_α–C_β conformation, but all of them favor the *gauche* conformations. The *gauche*(+)/*gauche*(−) ratio changes drastically from one model to another. Without a solvation potential (the NASP simulations), the rotamer count strongly favors *gauche*(+) (73%), with only smaller contribution of *trans* (18%) conformers. Turning on the ASP solvation potential raises the population of the *gauche*(−) conformer from 9% in the NASP simulation to 32% in the ASP case, but the basic structure of the distributions is similar for the NASP and ASPS simulations. Unlike in the ASPS and NASP simulations, *gauche*(−) is strongly favored in the SASA solvation potential simulation, and its *trans* probability is the highest

(21%) among the LD simulations. In contrast, the explicit water MD simulation presents a totally different picture with an almost even distribution between *gauche*(+) and *gauche*(−) and with a slightly higher *trans* ratio. The *trans* occurrence frequency in the MD simulation is 37%, which is higher than the *gauche*(−) (29%) and *gauche*(+) (34%) frequencies. The *trans* probability from MD simulation is slightly higher than those predicted with various LD methods. The MD simulations prediction of a greater preference for the *trans* conformer agrees quite well with the bioactivity assay.

The comparisons in Fig. 5 indicate a subtle inadequacy of the currently available solvation model potentials. A more detailed inspection of the molecular movie for the Met-enkephalin trajectories provides some insight into features of the potentials that produce this overestimation of the *gauche* conformations by the LD simulations. The close encounter between the two phenyl rings is generally accompanied by transitions out of the *trans* conformation. The ASPS simulations trajectories display the pair of phenyl rings as tending to stay close together throughout most of the simulation, and, consequently, the configuration for the Tyr and Phe are predominantly *trans* and *gauche*(+), respectively. Therefore, the underestimation of the population

in the *trans* state by the ASP solvation potential simulations may be due to an inaccuracy in the description of the nonbonding aromatic–aromatic interactions. Such an inaccuracy would also partly explain why the ASP solvation potential yields a different distribution for the radius of gyration than the distributions from the SASA and MD simulations. The ASP model lumps all carbon atoms into only one classification for which the effective solvation free energy per unit of accessible surface area is slightly positive, i.e., slightly hydrophobic. The SASA solvation model assigns aromatic carbon atoms as very slightly hydrophilic with a small negative solvation free energy-per-unit-accessible surface area. When two phenyl rings come in close contact, part of the solvent-accessible surface for the aromatic carbons is lost. The positive solvation free energy for all carbon atoms in the ASP parameter set implies that this loss of surface area creates an effective attractive force between the phenyl rings, so frequent contact between the aromatic groups is expected and is, indeed, observed in the animation of the ASP model trajectories. This effective solvation-induced attractive force between the phenyl groups also restricts Met-enkephalin in the ASP model from accessing the extended state to the extent found for the MD simulations or the SASA model.

Segmental time-correlation functions

Time-correlation functions provide a graphic measure for the rate of change of dynamic variables. The TCF used here to compare the dynamics predicted by the different models are the P_1 dipole autocorrelation functions of interatomic position vectors, which are defined as

$$C_{ij}(t) = \frac{\langle \vec{l}_{ij}(0) \cdot \vec{l}_{ij}(t) \rangle}{l_{ij}^2}, \tag{18}$$

where the interatom vectors l_{ij} are

$$\vec{l}_{ij} = \vec{r}_i - \vec{r}_j. \tag{19}$$

The angular brackets $\langle \cdots \rangle$ in Eq. 18 denote an equilibrium average. The P_1 correlation function is totally determined in the MD simulation by the amino acid sequence, the AMBER atomic interaction potentials, and the frictional forces exerted by surrounding solvent. By choosing atoms i and j to be proximate or distant, respectively, the correlation function $C_{ij}(t)$ reflects the local and global flexibility of the peptide. Because this flexibility generally varies along the peptide sequence, we analyze a number of these local P_1 correlation functions. A measure of the local segmental rigidity is taken as the decay rate of the TCF for the local motion. An analogous P_2 TCF can be obtained from several experimental quantities, such as the relaxation times T_1 and T_2, the S^2 order parameter, and the nuclear Overhauser effect in NMR relaxation studies and also from fluorescence anisotropy measurements. Because the relaxation time for

the P_2 TCF of a given $l_{ij}(t)$ is roughly half that for P_1, we provide computations for P_1 that are more easily evaluated from our theory for the long time dynamics that will be applied elsewhere to the present set of MD and LD simulations (M.-y. Shen and K. F. Freed, in preparation).

The local and global time-correlation functions considered here are evaluated by averaging the scalar product $\vec{l}_{ij}(0) \cdot \vec{l}_{ij}(t)$ over a trajectory of finite duration time. The statistical error due to the use of finite trajectories is estimated by the Zwanzig–Ailawadi formula (Zwanzig and Ailawadi, 1969),

$$\sigma = \sqrt{\frac{2\tau'}{T}} [1 - C(t)], \tag{20}$$

for $T > \tau'$, where τ' is the correlation time defined by,

$$\tau' = \int_0^\infty [C(t')]^2 \, dt', \tag{21}$$

and T is total duration of trajectories. Clearly, when the correlation function $C(t)$ approaches zero or when τ'/T is large, the statistical uncertainty in $C(t)$ is the greatest. Thus, the largest error arises for slow dynamical variables and in the long time tail of $C(t)$, especially for the much shorter (smaller T) explicit water MD simulations. The plots of the calculated correlation functions contain error bars for the $C(t)$ in each of the three curves depicted, with the values $C_\pm(t) = C(t) \pm \sqrt{2\tau/T} [1 - C(t)]$ providing estimates for the uncertainty in the computed TCF by use of finite duration trajectories.

The global dynamics of a polymer or peptide is usually associated with the slowest motions of the molecular system, such as tumbling, folding, and extending. These motions usually involve cooperative displacements of large portions of the solute, and the global motions may also affect the long time dynamics of more local molecular movements. The slowly varying variables we illustrate here are the backbone end-to-end vector (the $C_{\alpha1}$–$C_{\alpha5}$ separation vector) and the inter-phenyl ring vector ($C_{\gamma1}$–$C_{\gamma4}$). The inter-phenyl vector is considered here because of its relevance to the bioactivity of opioid peptides, as described in the previous subsection. The time scale for the conformational dynamics of the separation vector between the two phenyl groups provides a different perspective for the factors affecting the δ-receptor binding process. Figure 6 displays the simulated end-to-end TCF with the estimated statistical error bars placed at 500, 1000, and 2000 ps. The end-to-end vector correlation functions $C(t)$, whose relaxation time τ' is ~480 ps, exhibit similar behaviors for the ASPS, ASPD, and SASA simulations. The TCF obtained from the ASPS simulation is nearly identical to that from the ASPD computation with only very small deviations in the long time tail as shown in Fig. 6 A, but both depart significantly from the MD correlation functions. The MD
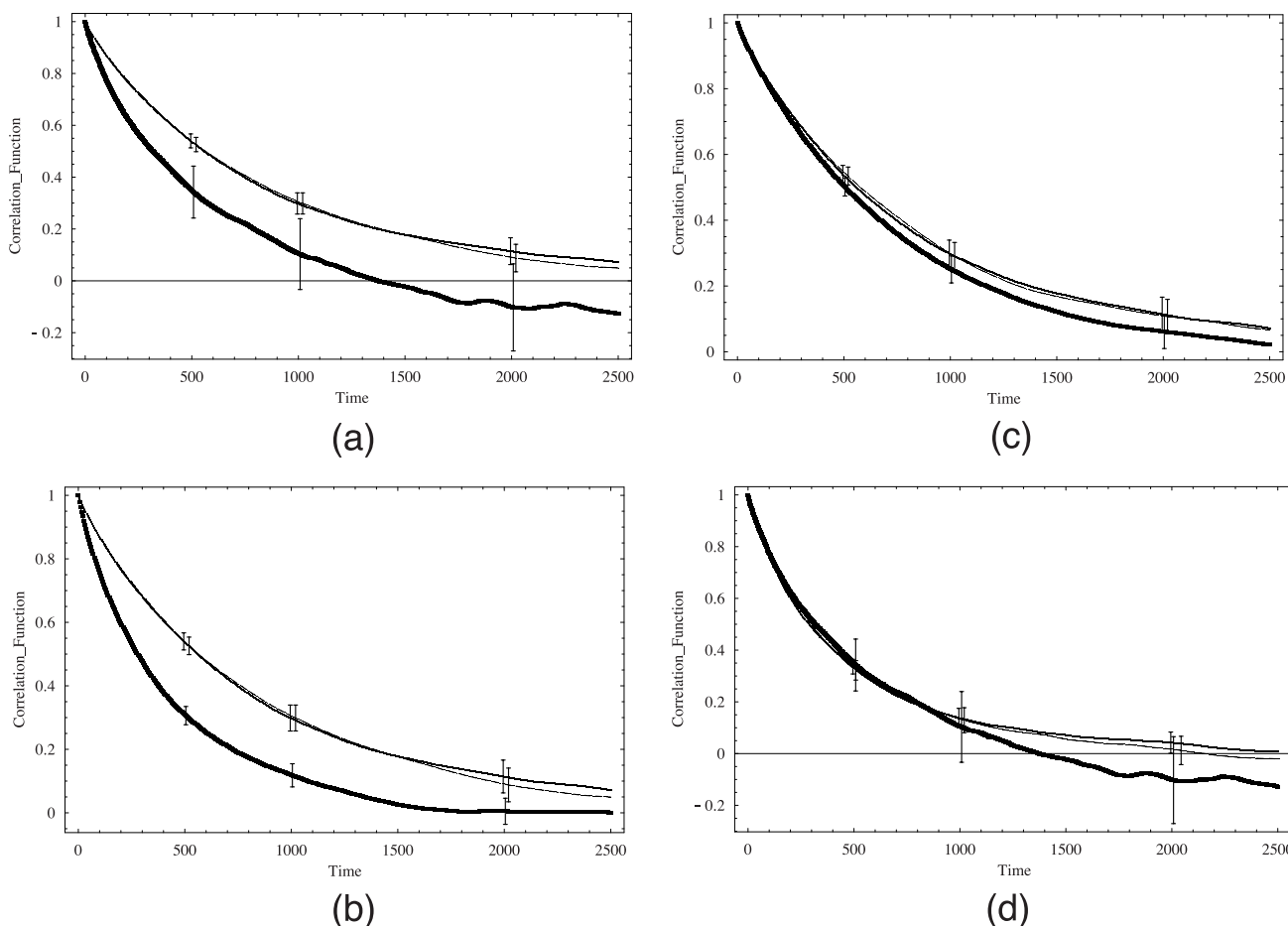
FIGURE 6   The end-to-end TCF from different simulation methods. (*A*) Comparison between explicit and implicit water model with experimental water viscosity. *Dark solid*, Full atom MD; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*B*) Comparison between different solvation potentials. *Dark solid*, the simulation without solvation potential; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*C*) Comparison between dynamical, static friction models and different dielectric constant. *Dark solid*, ASPS simulation with $\epsilon = 53.1$; *solid*, ASPS; *dashed*, ASPD simulations with experimental water viscosity. (*D*) Comparison between time-scaled dynamics, low-friction LD simulation and MD simulation. *Dark solid*, MD simulation; *solid*, ASPS simulation with time scaled by a factor 0.603; *dashed*, ASPS simulations with TIP3P water viscosity.

correlation function in Fig. 6 *A* appears closer to the NASP correlation function (see Fig. 6 *B*), which is computed without a solvation potential. This similarity between the NASP and MD simulations, however, is accidental, because the ASPS, ASPD, and SASA simulations in Fig. 6 have been performed using the experimental water viscosity of 0.84 cp at a temperature of 300 K rather than the TIP3P water viscosity of 0.50 cp that should be used in the LD simulations to mimic the frictional forces of the TIP3P water in the solvated MD simulations. The atoms in the MD simulation experience lower drag forces and, hence, exhibit faster global dynamics than the atoms in the LD simulation performed with the experimental water viscosity. Consequently, to compare the MD and LD simulations on an equal footing, two strategies have been pursued. First, another set of LD simulations has been generated using the TIP3P water viscosity (which leads to lower atomic frictions coefficients), and, second, we simply rescale the time with the

ratio between the simulated TIP3P and experimental water viscosities. The rescaling of time would render these two simulations identical if the dynamical influence of the inertia term of the Langevin equation is negligible for the long time dynamics (despite the fact that it represents the dominant force on some hydrogen atoms).

Figure 6 *D* compares the time rescaled correlation function from the ASPS (to correct for the use of the experimental viscosity) with the ASPL simulations that use the TIP3P viscosity obtained from MD calculations of pure TIP3P water. The TCF from the time-scaled ASPS and the TIP3P water viscosity ASPL simulations agree very well with each other and with the MD simulation TCF for the shorter time range where the error bars overlap substantially, and the agreement persists within error bars over the full range. This imply(s), as expected, that the inertial terms are fairly negligible for the long time behavior of the slower global motions. Inertial forces may be more important for
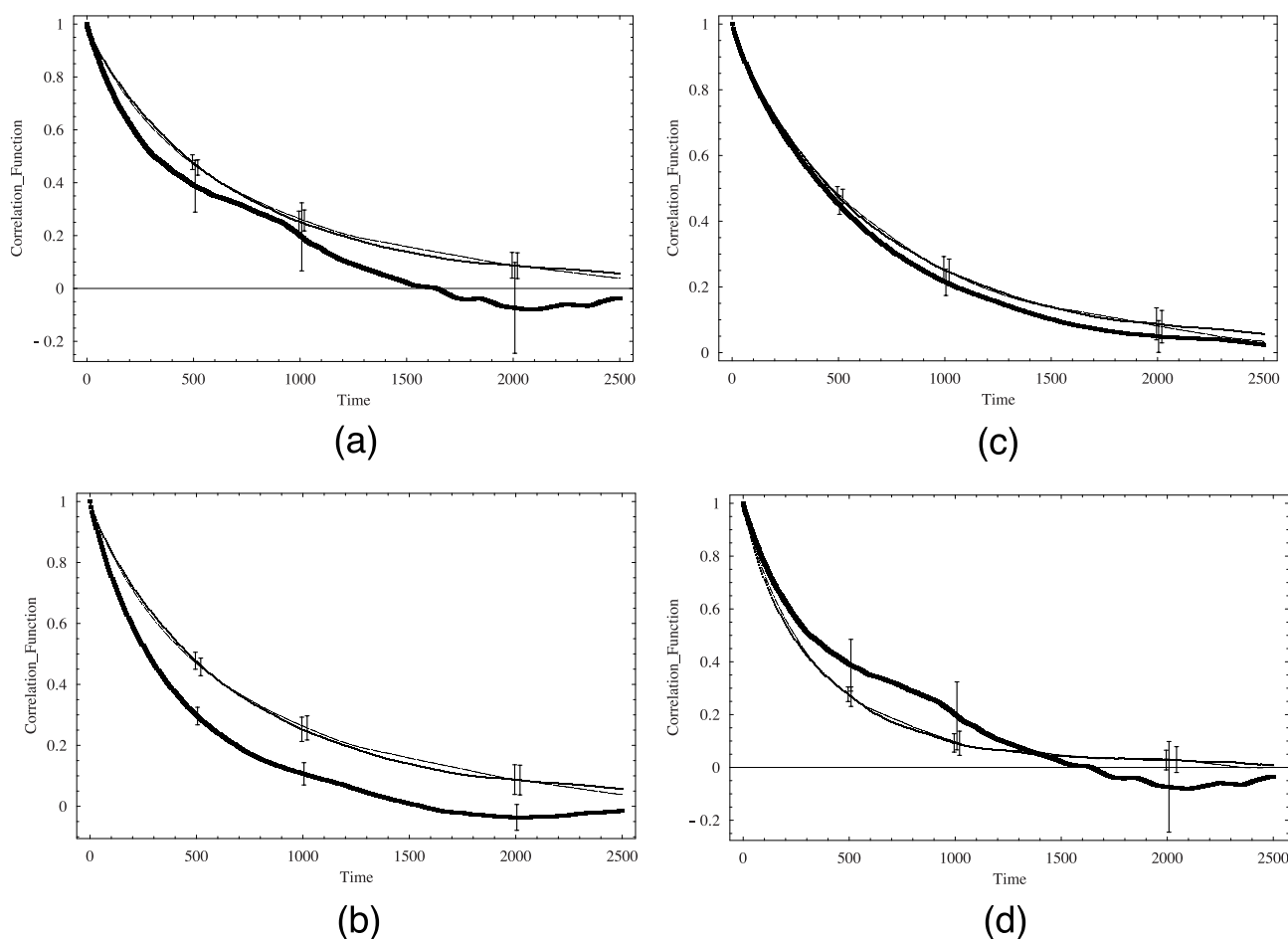
FIGURE 7   The phenyl–phenyl vector (atom number 10–45) time correlation functions from different simulation methods. (*A*) Comparison between explicit and implicit water model with experimental water viscosity. *Dark solid*, Full atom MD; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*C*) Comparison between different solvation potentials. *Dark solid*, the simulation without solvation potential; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*C*) Comparison between dynamical, static friction models and different dielectric constant. *Dark solid*, ASPS simulation with $\epsilon = 53.1$; *solid*, ASPS; *dashed*, ASPD simulations with experimental water viscosity. (*D*) Comparison between time-scaled dynamics, low-friction LD simulation and MD simulation. *Dark solid*, MD simulation; *solid*, ASPS simulation with time scaled by a factor 0.603; *dashed*, ASPS simulations with TIP3P water viscosity.

the dynamics of some buried hydrogen atoms whose ASA atomic friction coefficients are very small or zero. Time-scaled NASP correlation functions for global motions appear to decay too fast but behave relatively well for some local correlation functions. The origin of this differing behavior of the NASP model may be a direct consequence of the lack of appearance of extended configurations in the NASP simulations. The slower global motions couple more strongly with the overall rotation and are therefore affected by the presence or absence of the extended state. Due to the absence of a solvation potential in the NASP simulations, the more compact peptide tumbles at a faster rate because of a smaller diffusion tensor.

The inter-phenyl vector TCF in Fig. 7 exhibits similar trends among the different simulation methods as found for the end-to-end vector correlation function. The correlation times of the inter-phenyl vector are only slightly smaller

than the times for the end-to-end vector, indicating that the long time dynamics of the inter-phenyl separation vector is mainly governed by global tumbling. This slow relaxation of the inter-phenyl motion ensures a slower relative motion between phenyl groups in Met-enkephalin, and, therefore, affects its receptor affinity. Slower inter-phenyl dynamics conserve the relative orientations of Tyr and Phe residues along the main chain, producing the bioactivity of Met-enkephalin.

The local positional TCF involve more complicated intramolecular motions and are usually much faster than the correlation functions for global motions. The P_1 correlation time τ is defined by

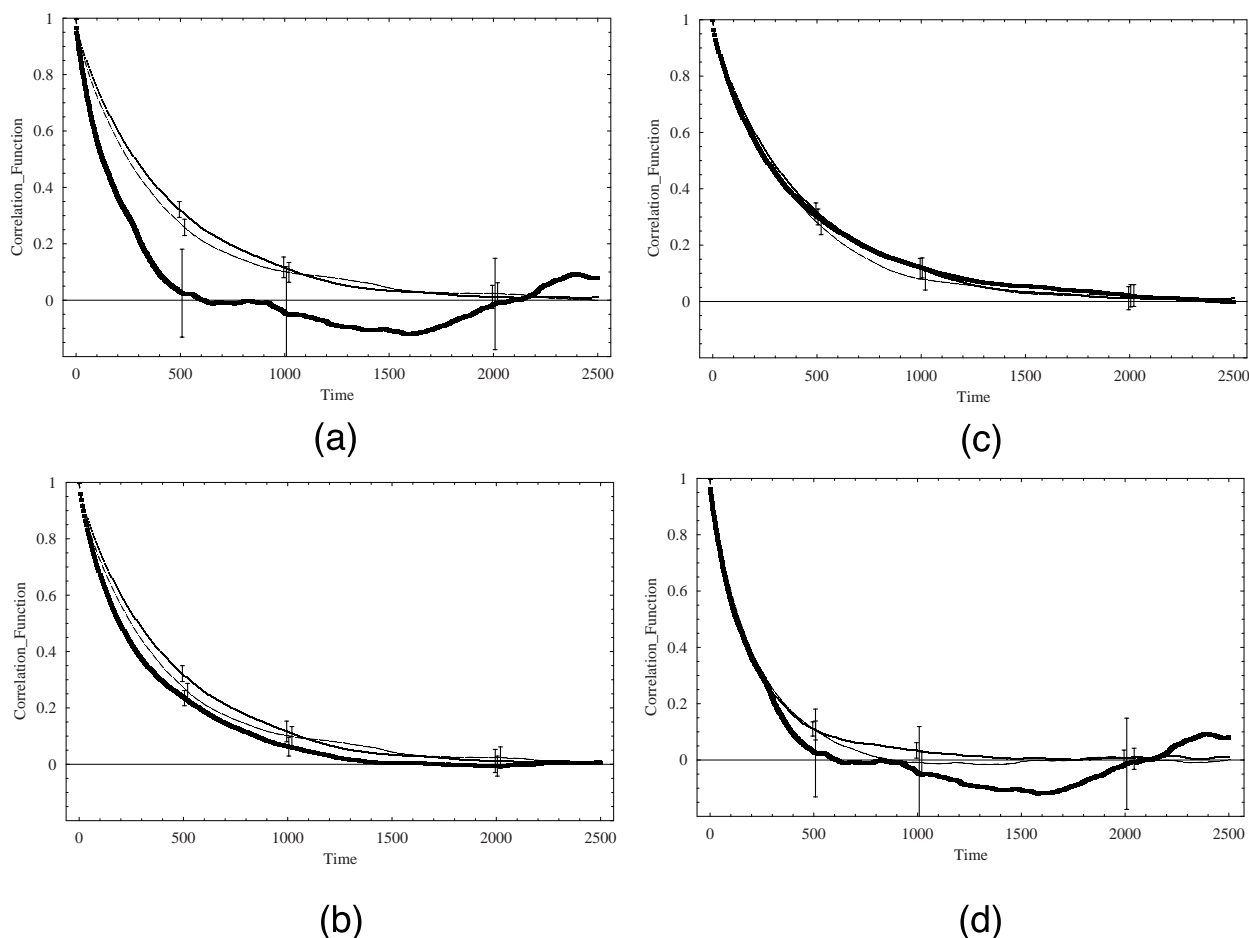$$\tau = \int_0^\infty C(t') \, \mathrm{d}t'. \tag{22}$$

(a)

(c)

(b)

(d)

FIGURE 8   The Phe-4 C_α–H correlation functions from different simulation methods. (*A*) Comparison between explicit and implicit water model with experimental water viscosity. *Dark solid*, Full atom MD; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*B*) Comparison between different solvation potentials. *Dark solid*, the simulation without solvation potential; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*C*) Comparison between dynamical, static friction models and different dielectric constant. *Dark solid*, ASPS simulation with $\epsilon = 53.1$; *solid*, ASPS; *dashed*, ASPD simulations with experimental water viscosity. (*D*) Comparison between time-scaled dynamics, low-friction LD simulation and MD simulation. *Dark solid*, MD simulation; *solid*, ASPS simulation with time scaled by a factor 0.603; *dashed*, ASPS simulations with TIP3P water viscosity.

The τ for local motions ranges from less than 20 ps, such as for the terminal C–H bond, to 100–200 ps for the central backbone bonds. We illustrate the general trends in Fig. 8 by considering the C_α–H bond of Phe-4, a central C–C bond to characterize the backbone dynamics, and one C–C bond on the aromatic side chains for describing the phenyl flip-

TABLE 2   The P_1 correlation times (in ps) for the C_α–H motion from Langevin dynamics simulations using various implicit solvent models (water viscosity is $\eta = 0.84$ cp)

Residue Name	ASPS	SASA	NASP
Tyr	138	154	103
Gly	124	101	113
Gly	153	119	130
Phe	207	183	150
Met	169	163	129

ping dynamics. The P_1 correlation times τ of five C–H correlation functions are listed in Table 2. These five correlation functions generated from the ASPS solvation simulation exhibit fairly similar features: all have approximately a 150-ps correlation time with a faster decay at short times. The Phe-4 C–H bond has the slowest decay (in 0–400 ps range) among the five. The Met-5 C–H bond has the slowest tail for $t > 400$ ps, but these correlation functions are still very close together. The nearly identical dynamics for all C–H bonds from ASPS simulations can be explained by the following structural feature discussed in the previous subsection: the positive free energy parameter between carbon atoms in the ASP model leads to an effective attractive force between phenyl groups, so the dynamics produced by this model describe Met-enkephalin as very prone to be in the semi-packed state. Because the semi-
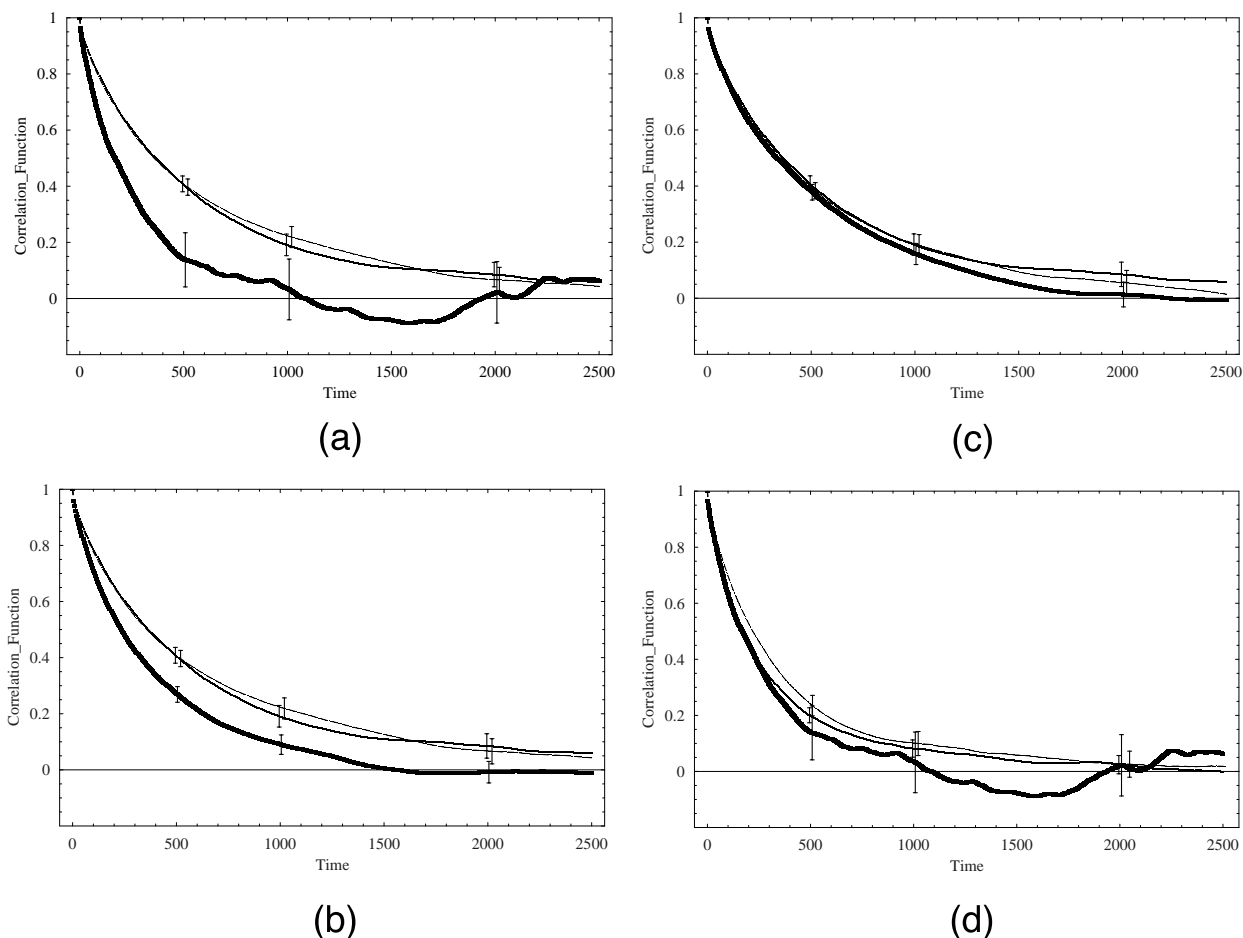
FIGURE 9 The central backbone C–C bond (atom number 25–26) time correlation functions from different simulation methods. (*A*) Comparison between explicit and implicit water model with experimental water viscosity. *Dark solid*, Full atom MD; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*B*) Comparison between different solvation potentials. *Dark solid*, the simulation without solvation potential; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*C*) Comparison between dynamical, static friction models and different dielectric constant. *Dark solid*, ASPS simulation with $\epsilon = 53.1$; *solid*, ASPS; *dashed*, ASPD simulations with experimental water viscosity. (*D*) Comparison between time-scaled dynamics, low-friction LD simulation and MD simulation. *Dark solid*, MD simulation; *solid*, ASPS simulation with time scaled by a factor 0.603; *dashed*, ASPS simulations with TIP3P water viscosity.

packed state has less backbone flexibility, the long time tail in all the C–H dynamics is mainly governed by rigid tumbling, leading to only a single decay rate for all the C–H bonds in the ASP model.

The time correlation functions computed with the SASA solvation parameters yield two distinguishable groups of C–H relaxation times, faster dynamics for Gly-2 and Gly-3 and slower for Tyr-1, Phe-4, and Met-5. The Phe-4 C–H bond correlation function decays slightly more slowly, and the Tyr-1 C–H relaxation time is the fastest among the three slow C–H bonds. A similar behavior emerges from the MD-calculated TCF, although the statistical errors are moderately large. All the simulation models predict the Gly-2 C–H relaxation as the fastest, and this prediction arises because the light hydrogen side chain and its position at the end of the backbone yield the expected faster motion. The internal backbone deformation motion, such as that of the

C–C bond shown in Fig. 9, has a relaxation time (τ) scale of ~400 ps, much faster than the time scale of global tumbling (900 ps). These results facilitate constructing a qualitative physical picture for the dynamical evolution of the Met-enkephalin dynamics. The faster moving leading portion (Tyr) swings around the heavier tail (Phe-Met) through a conformational transition occurring in the flexible connection between the two glycines. Met-enkephalin can undergo rapid transitions to the preferred δ-receptor docking position because the flexible Gly-Gly portion allows the fast structural readjustments that lead to higher binding strength. This conclusion partly explains why the observed and calculated solution conformation populations of opiate peptides differ from the one with the highest bioactivity. The local chain flexibility enables transitions to occur quickly between these two local conformations.
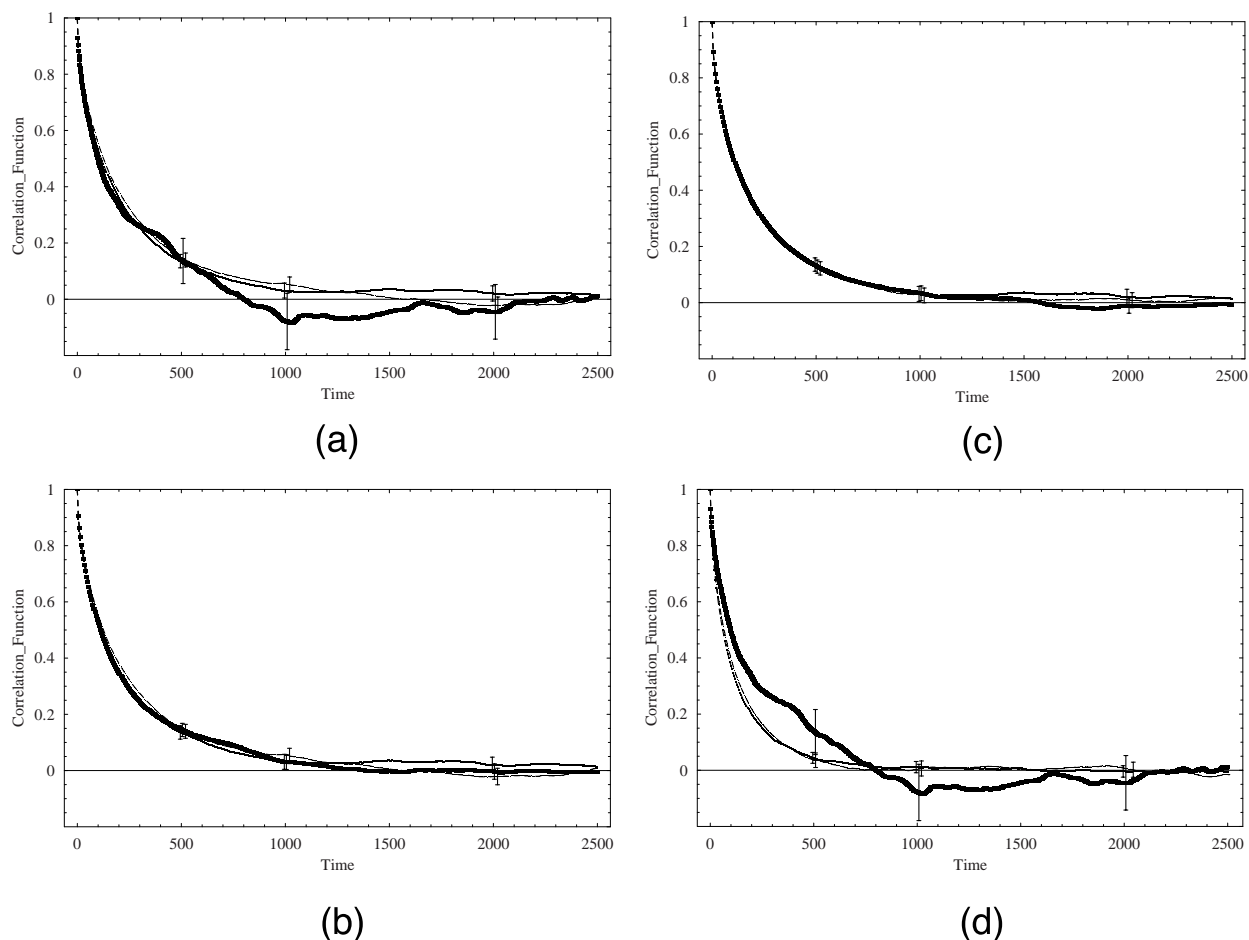
FIGURE 10 The Tyr-1 phenyl ring TCF from different simulation methods. (*A*) Comparison between explicit and implicit water model with experimental water viscosity. *Dark solid*, Full atom MD; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*B*) Comparison between different solvation potentials. *Dark solid*, the simulation without solvation potential; *solid*, ASPS; *dashed*, SASA simulations with experimental water viscosity. (*C*) Comparison between dynamical, static friction models and different dielectric constant. *Dark solid*, ASPS simulation with $\epsilon = 53.1$; *solid*, ASPS; *dashed*, ASPD simulations with experimental water viscosity. (*D*) Comparison between time-scaled dynamics, low-friction LD simulation and MD simulation. *Dark solid*, MD simulation; *solid*, ASPS simulation with time scaled by a factor 0.603; *dashed*, ASPS simulations with TIP3P water viscosity.

The phenyl ring flipping dynamics occurs on an even faster timescale. The two phenyl bond vectors display very different decay rates, with the Phe residue decay rate always slower than that for Tyr, a feature consistently predicted by MD and LD simulations. The ratios of the correlation times for the phenyl ring flipping range between 1.5 and 1.8. The slower dynamics of the Phe residue can be simply explained by steric hindrances imposed by the fast-moving Met side chain in its vicinity. Figure 10 displays the Tyr phenyl flipping TCF calculated from various methods.

Some of the aromatic group flipping motion relaxation rates can be measured by fluorescence anisotropy techniques, which may be applied for Tyr and Trp residues. The experiments yield the P_2-type correlation function,

$$P_2(t) = \frac{3}{2} \langle \cos^2 \theta_{ij}(t) \rangle - \frac{1}{2} = \frac{3}{2} \left\langle \frac{(\vec{r}_{ij}(0) \cdot \vec{r}_{ij}(t))^2}{r_{ij}^2(t) r_{ij}^2(0)} \right\rangle - \frac{1}{2},$$

(23)

whose correlation time is approximately one half as large as the time for the corresponding P_1 correlation function. Although we are unaware of measurements for Met-enkephalin, the tyrosyl rotational correlation time of Leu-enkephalin is measured as 84 ps from a time-resolved fluorescence study by Harms et al. (1998). The correlation time τ given by Eq. 21 for the Tyr residue is determined from the ASPS and SASA solvation potential simulations using the experimental water viscosity as 86 and 99 ps, respectively, agreeing quite well with the experimental value for Leu-enkephalin. However, the anisotropy decay parameters also shows a long time tail that has a time constant of 233 ps. This experimental long time structure is absent in our phenyl ring bond vector correlation functions. The long time tail is assigned as the global rotational time constant, which is shorter than the end-to-end vector relaxation time of 402 and 406 ps obtained from the ASP and SASA simulations for Met-enkephalin, respectively. The end-to-end relaxation

time calculated from the MD simulation is 232 ps. However, this excellent agreement should be examined carefully because of the difference between the TIP3P and real water viscosities. In addition, there are two reasons for the differences between the Leu- and Met-enkephalin dynamics. First, as discussed above, the local phenyl motion in the LD simulation does not strongly couple to the global tumbling motion, in contrast to the experimental observation for Leu-enkephalin. Coupling between these two types of motion could affect the decay rate. The other reason may arise from the fact that Leu-enkephalin has a branched side chain at the terminal amino acid as opposed to the more linear one in Met-enkephalin. The more compact side-chain structure in the Leu-enkephalin may be responsible for its faster global dynamics.

Adequacy of solvation model

Our comparisons of the various calculated properties between all-atom explicit solvent and reduced implicit solvent models use an approximation in which the dielectric constant is represented by a single constant. This simple approximation works well for describing the Met-enkephalin dynamics and is widely used in other studies for single small flexible peptides (Smith and Pettitt, 1994). We have also tested simulations performed with a lower dielectric constant 53.1 that may more closely reflect the dielectric constant of the TIP3P water model that is used in the MD simulation. However, this lower dielectric constant does not qualitatively change the equilibrium and dynamical properties, and the quantitative changes are quite small. This lack of sensitivity to varying the dielectric constant may arise because most residues in this system are hydrophobic. Garemyr and Elofsson (1999) have tested the sufficiency of the single dielectric model for a 108-amino acid system *Escherichia coli* thioredoxin (Trx) by using different dielectric models with various cut-offs. They find that the use of models with a distance-dependent dielectric function does not improve the accuracy over that obtained with a single dielectric constant, and neither does the use of a large distance cut-off (>8 Å) for all nonbonding interactions. Therefore, this electrostatic treatment is sufficient for a short, uncharged peptide, and it is difficult to improve upon this approach without severely degrading the computational efficiency. We have also benchmarked the GB/SA (Qiu et al., 1997) solvation potential model. This GB/SA model involves a costly pairwise summation and is consequently computationally nearly as expensive as the explicit water MD simulations. Moreover, the approach yields rather poor comparisons with MD simulations.

The macroscopic solvation potential is demonstrated as considerably affecting the relative conformational populations. The different solvation potential parameters for aromatic carbon–carbon interactions either create an effective attraction (the ASP) or repulsion (the SASA), but the computed Phe conformational distributions imply that the interactions in the MD simulation lie somewhat in between those of the ASP and SASA models. The effective carbon–carbon interactions in the solvation potential models arises from the overlap of the solvent-accessible atomic radius plus 1.4 Å (the solvent radius) (Richards, 1977) for the pair of interacting carbon atoms. This solvent-accessible radius is a universal single value for either hydrophilic or hydrophobic atoms, assuming the solvent molecules can always remain in contact with the exposed atomic surfaces. Realistically, solute atoms might push or pull the solvent in or out, so the solvent-accessible radius might be larger or smaller.

The basic assumption of a surface-based solvation potential may be questioned because investigations of the solvation free energy of a linear polymer system (Ben-Naim and Mazo, 1993) conclude that the solvation Gibbs free energy scales with the solute volume, not with the surface area. However, the volume occupied by Met-enkephalin changes little during its dynamical motion, so the surface portion may be most pertinent to modeling the solvation forces that are most relevant for the dynamics. Another problem arises because the solvation parameters are all fit to experimental solvation free energies. Just as for the water viscosity, the solvation parameters for the TIP3P water may differ from that of real water. Finally, as mentioned by Smith and Pettitt (1994), an atom-additive solvation free energy is not uniquely determined by rigorous statistical mechanics, but the beginnings of a statistical mechanical derivation are emerging (K. F. Freed, submitted for publication). Nevertheless, the excellent performance of the SASA solvation potential provides impetus for further tests and applications, for improvements of the solvation potential model, and for rigorous studies of its statistical mechanical underpinning.

REFERENCES

Allen, M. P., and D. J. Tildesley. 1987. Computer simulation of liquids. Oxford University Press, Oxford, U.K.

Barrow, C. J., and M. G. Zagorski. 1991. Solution structure of β-peptide and its constituent structure. *Science*. 253:179–182.

Baysal, C., and H. Meirovitch. 1997. Novel procedure for developing implicit solvation models for peptides and proteins. *J. Phys. Chem. B.* 101:7368–7370.

Baysal, C., and H. Meirovitch. 1998. Determination of the stable microstates of a peptide from NOE distance constrain and optimization of atomic solvation parameters. *J. Am. Chem. Soc.* 120:800–812.

Berdensen, H. J. C. 2001. GROMACS. Groningen University, Groningen, The Netherlands.

Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

Berendsen, H. J. C. 2001. GROMACS. Groningen University, Groningen, The Netherlands.

Ben-Naim, A., and R. M. Mazo. 1993. Size dependence of the solvation free energies of large solutes. *J. Phys. Chem.* 97:10829–10834.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117: 5179–5197.

Deber, C. M., and B. A. Behnam. 1984. Role of membrane-lipids in peptide-hormone function—binding of enkepahlins to micelles. *Proc. Natl. Acad. Sci. U.S.A.* 81:61–65.

Duan, Y., and P. A. Kollman. 1998. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science.* 282:740–744.

Eisenberg, D., and A. D. McLachlan. 1986. Solvation energy in protein folding and binding. *Nature.* 319:199–203.

Fändrich, M., M. A. Fletcher, and C. M. Dobson. 2001. Amyloid fibrils from muscle myoglobin. *Nature.* 410:165–166.

Feller, S. E., R. W. Pastor, A. Rojauckarin, S. Bogusz, and B. R. Brooks. 1996. Effect of electrostatic force truncation on interfacial and transport properties of water. *J. Phys. Chem.* 100:17011–17020.

Garemyr, R., and A. Elofsson. 1999. Study of the electrostatic treatment in molecular dynamics simulations. *Proteins Struct. Funct. Genet.* 37: 417–428.

Graham, W. H., E. S. Carter, II, and R. P. Hicks. 1992. Conformational analysis of Met-enkephalin in both aqueous solution and in the presence of sodium dodecyl sulfate micelles using multidimensional NMR and molecular modeling. *Biopolymers.* 32:1755–1764.

Harms, G. S., W. L. Freund, and C. K. Johnson. 1998. Time-resolved fluorescence study of conformation dynamics of opioid peptides. *J. Phys. Chem. B.* 102:5004–5010.

Helfand, E. 1960. Transport coefficients from dissipation in a canonical ensemble. *Phys. Rev.* 119:1–9.

Hruby, V. J., L. F. Kao, B. M. Pettitt, and M. Karplus. 1988. The conformational properties of the delta-opioid peptide [D-PEN2, D-PEN5] enkaphlain in aqueous-solution determined by NMR and energy minimization calculations. *J. Am. Chem. Soc.* 110:3351–3359.

Hughes, J., T. W. Smith, H. W. Kosterlitz, L. A. Fothergill, B. A. Morgan, and H. R. Morris. 1975. Identification of 2 related pentapeptides from brain with potent opiate. *Nature.* 258:577–579.

Isogai, Y., G. Némethy, and H. A. Scheraga. 1977. Enkephalin: conformational analysis by means of empirical energy calculations. *Proc. Natl. Acad. Sci. U.S.A.* 74:414–418.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Kirshenbaum, K., and V. Daggett. 1995. pH-Dependent conformation of amyloid β(1–28) peptide fragment explored using molecular dynamics. *Biochemistry.* 34:7629–7639.

Kostov, K. S., and K. F. Freed. 1999. Long-time dynamics of Met-enkephalin: comparison of theory with Brownian dynamics simulations. *Biophys. J.* 76:149–163.

Li, Z., and H. A. Scheraga. 1987. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 84:6611–6615.

Oliva, B., X. Daura, E. Querol, F. X. Avilés, and O. Tapia. 2000. A generalized Langevin dynamics approach to model solvent dynamics effects on proteins via a solvent-accessible surface. The carboxypeptidase A inhibitor protein as a model. *Theo. Chem. Accounts.* 105: 101–109.

Ooi, T., M. Oobatake, G. Némethy, and H. A. Scheraga. 1987. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proc. Natl. Acad. Sci. U.S.A.* 84:3086–3090.

Paine, G. H., and H. A. Scheraga. 1985. Prediction of the native conformation of a polypeptide by a statistical-mechanical procedure. I. Backbone structure of enkephalin. *Biopolymers.* 24:1391–1436.

Pasternak, G. W. 1988. Multiple morphine and enkephalin receptors and the relief of pain. *JAMA.* 259:1362–1367.

Pastor, R. W., and M. Karplus. 1988. Parametrization of the friction constant for stochastic simulations of polymers. *J. Phys. Chem.* 92: 2636–2641.

Plotnikoff, N. P., R. E. Faith, A. J. Murgo, and R. A. Good. 1986. Enkephalins and Endorphins: Stress and the Immune System. Plenum, New York.

Ponder, J. W., S. Rubenstein, C. Kundrot, S. Huston, M. Dudek, Y. Kong, R. Hart, M. Hodsdon, R. Pappu, W. Mooij, and G. Loeffler. 1999. TINKER: Software Tools for Molecular Design, Version 3.7. Washington University, St. Louis, MO.

Qiu, D., P. S. Shenkin, F. P. Hollinger, and W. C. Still. 1997. The GB/SA model for solvation. A fast analytical method for the calculation of approximate Born radii. *J. Phys. Chem. A.* 101:3005–3014.

Richards, F. M. 1977. Areas, volumes, packing, and protein-structure. *Ann. Rev. Biophys. Bioeng.* 6:151–176.

Roques, B. P., C. Garbay-Jaureguiberry, R. Oberlin, M. Anteunis, and A. K. Lala. 1976. Conformation of Met-enkephalin determined by high field PMR spectroscopy. *Nature.* 262:778–779.

Smith, G. D., and J. F. Griffin. 1978. Conformation of [*Leu$_5$*]enkaphalin from x-ray diffraction—features important for recognition at opiate receptor. *Science.* 199:1214–1216.

Smith, P. E., and B. M. Pettitt. 1994. Modeling solvent in biomolecular systems. *J. Phys. Chem.* 98:9700–9711.

Smith, P. E., and W. F. van Gunsteren. 1993. The viscosity of SPC and SPC/E water at 277 and 300 K. *Chem. Phys. Lett.* 215:315–318.

Stryer, L. 1988. Biochemistry. W. H. Freeman and Company, New York.

Tourwé, D., K. Verschueren, A. Frycia, P. Davis, F. Porreca, V. J. Hruby, G. Toth, H. Jaspers, P. Verheyden, and G. V. Binst. 1996. Conformational restriction of Tyr and Phe side chains in opioid peptides: information about preferred and bioactive side-chain topology. *Biopolymers.* 38:1–12.

van Gunsteren, W. F., and H. J. C. Berendsen. 1982. Algorithms for Brownian dynamics. *Mol. Phys.* 45:637–647.

Wang, Y., and K. Kuczera. 1996. Molecular dynamics simulations of cyclic and linear DPDPE: influence of the disulfide bond on peptide flexibility. *J. Phys. Chem.* 100:2555–2563.

Wesson, L., and D. Eisenberg. 1992. Atomic solvation parameters applied to molecular dynamics of proteins in solution. *Protein Science.* 1:227–235.

Zwanzig, R. 1965. Time-correlation functions and the transport coefficients in statistical mechanics. *Ann. Rev. Phys. Chem.* 16:67–102.

Zwanzig, R., and N. K. Ailawadi. 1969. Statistical error due to finite time averaging in computer experiments. *Phys. Rev.* 182:280–283.