Mass-weighted molecular dynamics simulation and conformational analysis of polypeptide

Boryeu Mao
Upjohn Research Laboratories, Kalamazoo, Michigan 49001 USA

ABSTRACT Atomic motions in protein molecules have been studied by molecular dynamics (MD) simulations; dynamics simulation methods have also been employed in conformational studies of polypeptide molecules. It was found that when atomic masses are weighted, the molecular dynamics method can significantly increase the sampling of dihedral conformation space in such studies, compared to a conventional MD simulation of the same total simulation time length. Herein the theoretical study of molecular conformation sampling by the molecular dynamics-based simulation method in which atomic masses are weighted is reported in detail; moreover, a numerical scheme for analyzing the extensive conformational sampling in the simulation of a tetrapeptide amide molecule is presented. From numerical analyses of the mass-weighted molecular dynamics trajectories of backbone dihedral angles, low-resolution structures covering the entire backbone dihedral conformation space of the molecule were determined, and the distribution of rotationally stable conformations in this space were analyzed quantitatively. The theoretical analyses based on the computer simulation and numerical analytical methods suggest that distinctive regimes in the conformational space of the peptide molecule can be identified.

INTRODUCTION

Intramolecular and intermolecular interactions in biological macromolecules such as proteins and nucleic acids can be described by an empirical potential energy function of the molecular mechanics type and studied by numerical methods such as energy minimization and molecular dynamics simulation techniques (McCammon and Harvey, 1987; Brooks et al., 1988). Characteristics of localized atomic motions such as the anisotropy and the anharmonicity of the atomic potential of mean force (Northrup et al., 1981; Mao et al., 1982; Ichiye and Karplus, 1987), and collective motions in proteins (Brooks and Karplus, 1983; Levy et al., 1985; Post et al., 1989) have been studied. Due to the large number of interacting atomic particles and the resulting high dimensionality of biomolecular systems however, it is usually necessary to devise alternative methods for studying large scale structural changes (Mao and McCammon, 1983; Harvey and McCammon, 1981; Thaisrivongs et al., 1990).

For short polypeptide chains, the sampling of molecular conformational space by molecular dynamics methods is expected to be more efficient in general, but may become limited in special cases (Kitson and Hagler, 1988; Al-Obeidi et al., 1989; Mao and Friedman, 1990). For the tetrapeptide Phe-Met-Arg-Phe-amide (FMRF-amide), which may function as a modulator of neurotransmitter release and ion channel activity in sensory neurons (Man-Son-Hing et al., 1989; Sweatt et al., 1989), two of the backbone dihedral angles remain in their respective dihedral potential wells during molecular dynamics simulations, even at elevated temperatures, but make transitions to other potential wells when atomic masses in the system are weighted (Mao and Friedman, 1990). Whereas others have employed mass scaling for efficient sampling of local potential energy surface in molecular dynamics (Bennett, 1975) and for damping high angular velocities of water molecules in computer simulations (Wood, 1979; Pomès and McCammon, 1990), we demonstrated that weighting atomic masses in a polypeptide molecular system increases the effective momentum of dihedral rotational motion in barrier crossing and therefore increases the efficiency of conformational search in molecular dynamics simulations (Mao and Friedman, 1990).

Here we report the details of the theoretical study in which the mass-weighted molecular dynamics (MWMD) method was applied to the conformational study of the tetrapeptide FMRF-amide; moreover, due to the extensive conformational sampling of the backbone dihedral angles, it is appropriate to devise a numerical analytical scheme for studying molecular structures generated from the computer simulation. In this numerical analytical scheme, distributions of the values of individual backbone dihedral angles of FMRF-amide in their respective full 360°-range were obtained from the MWMD trajectory; on the basis of these distributions, a set of nonoverlapping low-resolution structures covering all of the backbone dihedral conformational space of the molecule were determined uniquely without predetermined parameters. The results suggest that the confor-

mational space of the molecule is divided into regimes which are populated differentially and separated from one another by potential energy barriers; in contrast, for tetraglycine amide, all molecular conformations are nearly equally populated. A comprehensive analysis of peptide conformations such as that described in this report will facilitate the characterization of the structure of free and receptor-bound forms of this important class of biological molecules and their function.

METHODS

Structure generation, energy minimization, and molecular dynamics calculations are performed in the macromolecular program system CHARMM (Brooks et al., 1983), using standard procedures and options of the program package. Supplementary FORTRAN programs have been written for the analyses of the trajectories generated from the simulation. Parameters and characteristics of the MWMD simulation are first described in the Simulations subsection. Trajectories, or time series, of individual backbone dihedral angles are then computed from the coordinate sets saved during dynamics simulation.

For conformational analyses, the noise in individual dihedral trajectories from MWMD was first filtered numerically. Time steps at which a dihedral angle is in rotational transitional state are then identified on the basis of numerically determined first and second derivatives of the filtered trajectory; the number of potential wells and their location in the full 360° range are then determined from all time steps at which the dihedral angle is not in rotational transitional state. Numeric labels identifying dihedral potential wells are then assigned to individual dihedral angle values in the time series, and the backbone conformation of a molecular structure in the simulation is specified by the collection of such labels for backbone dihedral angles. These procedures are described in the Analyses subsection below.

Simulations

For the initial conformation of Phe-Met-Arg-Phe-amide, the internal coordinates (ie., covalent bonds, bond angles, and dihedral angles) of each amino acid are set to values in an average structure of the corresponding amino acid residue type which was computed from globular proteins in the Brookhaven Protein DataBank (Mao, B., unpublished data). The amino-terminal of the molecule is protonated, and for the carboxy-terminal amide, the internal coordinates are initially set to equilibrium values in the parameter set (Brooks et al., 1983) and hydrogen atoms are set in a planar arrangement. The molecule is then energy-minimized for relieving strains in internal coordinates and intramolecular close contacts. For energy minimization the dielectric constant is set at a constant value of 80 in place of explicit solvent molecules. For nonbonded interactions (van der Waals and electrostatic terms), atom pairs with interatomic distance greater than a cutoff of 8.0 Å are excluded (due to their negligible contribution to the total energy), and for hydrogen-bond interactions, donor-acceptor pairs with interatomic distance greater than a cutoff of 4.5 Å are excluded. In energy minimization of the initial structure, the nonbonded and hydrogen-bond lists are regenerated at every 25 steps in the adopted-basis Newton-Raphson procedure. Backbone dihedral angles ϕ and ψ of the energy-minimized structure are listed in Fig. 6.

Molecular dynamics of FMRF-amide was started from the energy-minimized structure, using the same set of nonbonded and hydrogen-bond cutoffs for intramolecular interactions as that in energy minimization. The numerical integration of Newton's equation during dynamics follows the scheme of Verlet (Brooks et al., 1983; Verlet, 1967), with an integration time step of 0.0009766 ps. The first 20 ps of a simulation was devoted to heating and equilibration of the system. During the heating phase, the temperature of the system was increased at 30°K increments and at 0.5 ps intervals until it reaches nominal system temperature T_s; for any temperature T during the heating phase, atomic velocities are assigned from a Gaussian distribution at T. In the equilibration phase, the temperature of the system was repetitively assigned at T_s at 0.5 ps intervals. At the conclusion of equilibriation phase, the molecular dynamics was then continued without any intervention.

For mass-weighted molecular dynamics (MWMD), atomic masses, and the system temperature were weighted with a scaling factor ω: $m_i' = \omega \cdot m_i$ for atoms i, and $T_s' = \omega \cdot T_s$, where m_i and T_s are respectively atomic masses and the system temperature of a corresponding MD simulation without mass weighting. The above relationships ensure that, for comparison purpose, root-mean-squared velocities of the two simulations are the same according to the equipartition relationship, $kT_s/2 = (\Sigma m_i \cdot V_i^2/2)/(3 \cdot N - 6)$, where N is the number of atoms in the molecular dynamics system, and k is the Boltzmann constant. In this study, ω is set to 10.0, and T_s is 600°K unless otherwise specified. For the mass-weighted MD simulation, harmonic constraints were placed on all internal coordinates except proper dihedral angles: bond lengths and bond angles are constrained with harmonic potentials of 10,000.0 kcal/mol/Å^2 and 4,000.0 kcal/mol/rad^2, respectively, and improper dihedral angles for planar and tetrahedral configurations (i.e., amide plane and tetrahedral Cα atoms, respectively) are constrained with harmonic potentials of 600.0 kcal/mol/rad^2. Coordinate sets were saved at every time step of the molecular dynamics integration, and for any structural quantity (such as a dihedral angle of the molecule), a time series can be constructed each member of which is computed from the corresponding coordinate set and indexed by an integer specifying the number of integration steps in 0.0009766 ps units.

Analyses

For conformational analyses, MWMD trajectories (i.e., time series) of backbone dihedral angles ϕ and ψ of all residues are calculated from the coordinate sets saved during dynamics simulations. Due to the mass weighting, fluctuations in dihedral angle trajectories are considerably (though not completely) damped. Moreover, these trajectories cover the entire range of values for dihedral space; thus, it is appropriate to devise a numerical analytical scheme for determining how the entire 360° range is divided into potential wells separated by rotational potential energy barriers, and whether at each time step in the simulation a dihedral angle is either in a potential well, or in a "rotational transitional state" when it is located in the region of the dihedral space where there is a local barrier for rotation. In this numerical analysis, $\varphi(t)$, the value of dihedral angle $\boldsymbol{\varphi}$ at the time step t in the time series, is first transformed to $\bar{\varphi}(t)$ according to the following formula:

$$\begin{aligned}\bar{\varphi}(t) = [&[\varphi(t-69) + \varphi(t-68)] + 2 \cdot [\varphi(t-67) + \varphi(t-66)] \\ &+ \ldots + 33 \cdot [\varphi(t-5) + \varphi(t-4)] + 34 \cdot [\varphi(t-3) + \varphi(t-2)] \\ &+ 35 \cdot [\varphi(t-1) + \varphi(t) + \varphi(t+1)] + 34 \cdot [\varphi(t+2) \\ &+ \varphi(t+3)] + 33 \cdot [\varphi(t+4) + \varphi(t+5)] + \ldots \\ &+ 2 \cdot [\varphi(t+66) + \varphi(t+67)] \\ &+ [\varphi(t+68) + \varphi(t+69)]]/A,\end{aligned}$$

where, for time step t, $\bar{\varphi}$ is a weighted average for the dihedral angle calculated from 139 values of φ in the period from $t - 69$ to $t + 69$, for which high frequency fluctuations within $\sim t \pm 0.070$ ps are filtered

out. The normalization factor A in the above expression is numerically equal to 2,485, which is the total number of φ terms weighted by the coefficients for the brackets.

From $\bar{\varphi}(t)$, the first derivative at time step t, $\varphi'(t)$, is first approximated by the average of two differences:

$$\varphi'(t) = |[\bar{\varphi}(t) - \bar{\varphi}(t-1)] + [\bar{\varphi}(t+1) - \varphi(t)]|/2$$
$$= [\bar{\varphi}(t+1) - \bar{\varphi}(t-1)]/2|.$$

A weighted average for the derivative, $\bar{\varphi}'(t)$, is then computed from $\varphi'(\tau)$ for all τ in the period from $t-5$ to $t+5$, using a formula analogous to that for $\bar{\varphi}(t)$ above but for the smaller range here which is sufficient for filtering out fluctuations in $\varphi'(t)$. The second derivative at time step t, $\varphi''(t)$, and the weighted average, $\bar{\varphi}''(t)$, are then determined from $\bar{\varphi}'(t)$ in the same manner as $\varphi'(t)$ and $\bar{\varphi}'(t)$ were determined from $\bar{\varphi}(t)$. Whether a dihedral angle is in a rotational transitional state (simply referred to as "transitional state" henceforth) at a given time step is then determined on the basis of the values of the derivatives, as follows. In the classical mechanics of a particle oscillating in a one-dimensional double-well harmonic potential with a total energy greater than the height of the central potential energy barrier, the speed of the particle is the largest when it is located at the two minima of the potential energy and the smallest when it is located at the central potential energy barrier maximum. Thus, dihedral angle $\boldsymbol{\varphi}$ is considered to be in a potential well at time step t if the rotational speed $\bar{\varphi}'(t)$ is large; i.e., when $\bar{\varphi}'(t)$ is near a local maximum of the first derivative curve $\bar{\varphi}'$. At time steps when $\boldsymbol{\varphi}$ is not in a local potential well, it is in a rotational transitional state (i.e., at the height of a local barrier). This scheme will be illustrated in the Results section.

In the time series of each dihedral angle, time steps at which the dihedral angle is in transitional state are identified and labeled as T. A histogram is then made from values of $\boldsymbol{\varphi}$ at all time steps that are not labeled T; from such a histogram, the number of potential wells in the 360° range of the dihedral angle is determined and each potential well identified with a numeric label. The dihedral angle at each time step in the entire time series is then assigned one of the numeric labels for its potential wells if $\varphi(t)$ is nontransitional (i.e., is not already assigned a T label) and falls within the range of the specified potential well.

To specify the backbone conformation of FMRF-amide molecule, a set of eight such labels are required, one for each of the eight ϕ and ψ dihedral angles of the four residues in the molecule: $\boldsymbol{\varphi}_j$, $j = 1, 2, \ldots, 8$, where $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_2$ are respectively the backbone dihedral angles ϕ and ψ of residue Phe_1 at the amino-terminus, and $\boldsymbol{\varphi}_3$ and $\boldsymbol{\varphi}_4$ are the pair of ϕ and ψ angles of residue Met_2, and so on. Amide bond dihedral angles have been maintained at near 180° by constraints. A molecular conformation is thus completely identified by an 8-tuple, $\mathscr{L}_1\mathscr{L}_2\mathscr{L}_3\mathscr{L}_4\mathscr{L}_5\mathscr{L}_6\mathscr{L}_7\mathscr{L}_8$, where $\mathscr{L}_j$ is either a numeric label specifying the potential well in which the dihedral angle is located, or a T specifying a rotational transitional state. As will be clear from results to be presented in the next section, the dihedral potential wells for FMRF-amide so determined are quite broad, some covering as wide as a 240° range of the dihedral space; molecular conformations specified by these labels are therefore low-resolution structures in that all dihedral angle values within a potential well are identically labeled. Moreover, all backbone dihedral angles were found to have a finite (and small) number of potential wells such that the permutation of numeric labels in an 8-tuple generates a relatively small number of nonoverlapping conformations that cover the entire backbone conformation space. The set of molecular structures saved from MWMD simulation is converted to a time series of these 8-tuples, one for each time step; this time series is then analyzed for the distribution of rotationally stable molecular structures in the molecular conformational space; rotationally stable molecular structures will be referred to simply as "stable molecular structures" or "stable molecular conformations" henceforth.

RESULTS

Mass-weighted molecular dynamics simulations

In the mass-weighted molecular system, the potential energy function terms which describe nonbonded interatomic interactions within the molecule remain unchanged whereas atomic masses are weighted by ω. The molecular dynamics integration of this system nonetheless should, as in molecular dynamics integration of normal systems, generate a microcanonical ensemble, in spite of mass weighting and the extra constraints placed on covalent interactions. As expected for a microcanonical ensemble, (*a*) the total energy (kinetic energy plus potential energy) of the system remains essentially constant during MWMD, with a mean total energy of ~1,880.0 kcal/mol, and (*b*) the total energy is approximately equally divided between the kinetic energy and the potential energy. The large potential energy is almost entirely due to the harmonic constraint terms; the aggregate of dihedral angle terms and that of nonbonded interactions contribute ~1% each to the total potential energy. The RMS fluctuation in the total energy is ~17.5 kcal/mol (or ~0.9% of the mean energy); this fluctuation is larger than that in normal MD, resulting from the larger kinetic energy in the system due to larger atomic masses and the large harmonic constraints placed on internal coordinates.

Fig. 1 shows the comparison of typical trajectories from MD and MWMD simulations in the first 100 ps for two of the dihedral angles. For MWMD, the fluctuations in dihedral angles are considerably but not completely damped and the slowly fluctuating and drifting values of $\boldsymbol{\varphi}$ cover the entire 360° range during the simulation. To determine the number and the range of potential wells in the 360° space for dihedral angle $\boldsymbol{\varphi}$, fluctuations in the time series for $\boldsymbol{\varphi}$ from MWMD trajectory, $\varphi(t)$, are filtered and time steps at which the speed of dihedral liberation, $\bar{\varphi}'(t)$, is at local maximum are determined; at those time steps for which $\bar{\varphi}'(t)$ is close to such maximal velocities the dihedral angle $\boldsymbol{\varphi}$ is considered to be located in a potential well, and at all other time steps, in transitional state. Fig. 2*a* shows the time series for $\boldsymbol{\varphi}_4$(i.e., the backbone dihedral angle ψ of Met_2), with the filtered values shown in the superimposed curve. The scheme shown in Fig. 2*b* then is used to determine whether dihedral angle $\boldsymbol{\varphi}$ is in transitional state at a given time step. When original values $\varphi(t)$ for dihedral angle ϕ of Phe_1 in the MWMD trajectory are displayed in a histogram (Fig. 3*a*, curve 1), the distribution is approximately even in the full 360° range apparently suggesting a flat potential energy curve for this dihedral angle. The histogram of all values of ϕ_1 identified to be

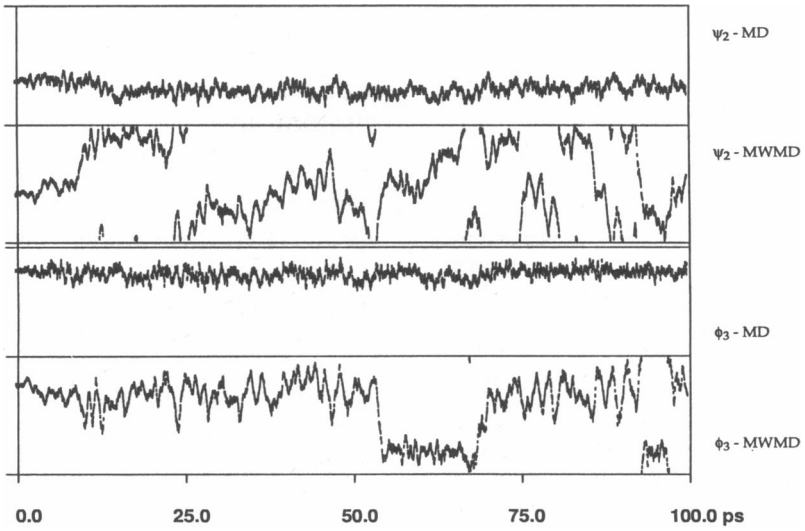

FIGURE 1 Comparison of characteristic time series of dihedral angle fluctuations during MD and MWMD simulations. The trajectories for ψ_2 are representative of those for backbone ψ dihedral angles, and the trajectories for ϕ_3 for backbone ϕ dihedrals. The vertical scale for each time series is the full 360° range. Higher frequency fluctuations are damped (though not completely), and low-frequency librations and drifting in mean values are introduced by the mass weighting.

in *transitional state* however shows three potential energy barriers (Fig. 3*a*, curve 2). Moreover, the histogram for all values of ϕ_1 identified to be nontransitional should indicate location of the potential wells and curve 3 shown in Fig. 3*a* is indeed an inverse of the potential energy barrier curve (Fig. 3*a*, curve 2). Our analysis of the time series is thus consistent with the classical mechanics description of the dihedral libration as an approximate multiwell harmonic oscillator described in the Methods section. Moreover, the potential energy barriers shown in curve 2 of Fig. 3*a* coincide with the threefold symmetry of the dihedral potential for ϕ_1 in the potential energy function, indicating that the rotational freedom of the amino-terminal amine group is essentially not hindered by intramolecular interactions in the intact molecule. Whereas other algorithms have been used for filtering noise in molecular dynamics trajectories (Sessions et al., 1988), the more straightforward analyses described here nonetheless present a complete description of the dihedral libration in a short peptide molecule such as FMRF-amide; Fig. 3*a* shows that the amino-terminal amine is a freely rotating group, not significantly constrained by intramolecular interactions, whereas the distributions of values for interior ϕ and ψ dihedral angles during the 400 ps simulation are bimodal (Fig. 3*b*), and the ϕ–ψ plot (Fig. 3*c*) correlates with contours in the potential energy surface of peptide stereochemistry (Ramachandran et al., 1966). Results in Fig. 3 show that the conformational space of ϕ and ψ dihedral angles of individual amino acids is well sampled in the MWMD simulation.

Conformational analyses

With the exception of ϕ_1 for the amino-terminal amine group, which has three potential wells (three peaks in Fig. 3*a*), all other backbone dihedral rotations have bimodal distributions (e.g., two peaks in Fig. 3*b*). Dihedral angle values that define the range of these potential wells are shown in Table 1. Given these definitions, the backbone conformation of any structure of FMRF-amide in the MWMD trajectory can be specified by an 8-tuple $\mathcal{L}_1\mathcal{L}_2\mathcal{L}_3\mathcal{L}_4\mathcal{L}_5\mathcal{L}_6\mathcal{L}_7\mathcal{L}_8$, where $\mathcal{L}_1 \in \{1, 2, 3, T\}$ and $\mathcal{L}_j \in \{1, 2, T\}$ for $j = 2, \ldots, 8$; these 8-tuples specify "low-resolution" conformations of the backbone of FMRF-amide. According to this discrete labeling scheme, there are 384 possible and distinct low-resolution, stable conformations covering the entire backbone dihedral space of the molecule, each specified by a *proper* 8-tuple in which all labels are numeric. In addition to proper 8-tuples, there are 8-tuples in which one or more of the labels is T; these are *improper* 8-tuples. In the following discussion of molecular conformation sampling, several additional categories of state labels are defined. An improper 8-tuple which contains a T label in position $\mathcal{L}_j$ is converted to a *pseudoproper* 8-tuple at $\mathcal{L}_j$ (abbreviated to pseudo-$\mathcal{T}_j$ 8-tuple) when $\mathcal{L}_j$ is reassigned a numeric label based on the transitional value of dihedral angle φ_j. On the other hand, due to the threefold symmetry of ϕ_1, shown in Fig. 3*a*, the 384 proper 8-tuples can be reduced to 128 distinct proper 7_1-tuples by allowing any numeric label for $\mathcal{L}_1$ (i.e., $\mathcal{L}_2\mathcal{L}_3\mathcal{L}_4\mathcal{L}_5\mathcal{L}_6\mathcal{L}_7\mathcal{L}_8$, from which $\mathcal{L}_1$ representing any of the three possible numeric labels for the

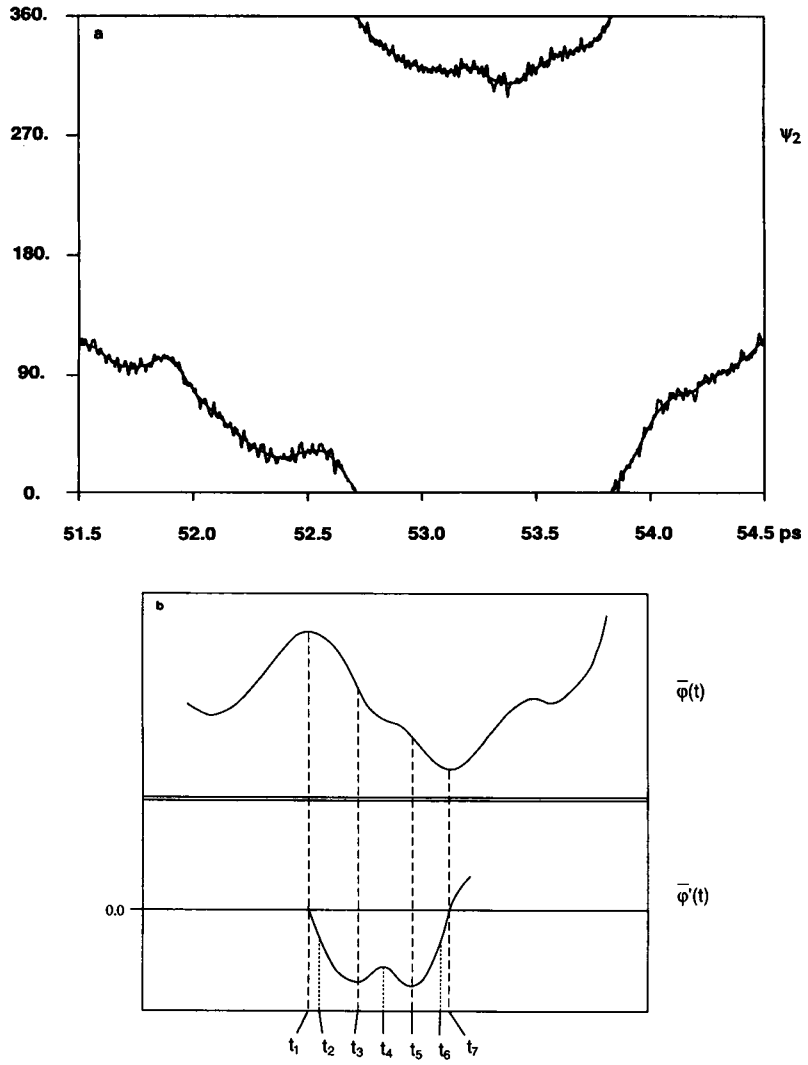

FIGURE 2 Numerical analyses of dihedral angle trajectories. (*a*) Filtering of high frequency fluctuations in dihedral angle trajectory. The smooth curve superimposed on the original trajectory consists of $\bar{\varphi}(t)$ for ψ_2. (*b*) Determination of time steps at which a dihedral angle is in transitional state according to the schematic $\bar{\varphi}(t)$ and $\bar{\varphi}'(t)$ curves. At time steps t_3 and t_5, $|\bar{\varphi}'(t)|$ is at local maximum, at t_4 it is at local minimum, and at t_1 and t_7 it is zero. φ at time steps in the periods from t_2 to t_3 and from t_5 to t_6 has near-maximum velocities of dihedral rotation and they are designated to be within local dihedral potential wells; the lengths of time periods $t_3 - t_2$ and $t_6 - t_5$ are 70% of time periods $t_3 - t_1$ and $t_7 - t_5$, respectively. φ in t_1–t_2 and in t_6–t_7 are in dihedral rotational transitional state. In the schematic curves for $\bar{\varphi}(t)$ and $\bar{\varphi}'(t)$ shown here, $t_4 - t_3$ and $t_5 - t_4$ are each < 0.030 ps and φ in such short time segments are conservatively designed to be in transitional state but not located in any local potential wells.

backbone dihedral angle ϕ of Phe_1 is omitted); subscript 1 in 7_1 denotes the reduction of labels at $\mathscr{L}_1$. Lastly, pesudo-$\mathscr{T}_j$ 7_1-tuples are all 7_1-tuples reduced from pseudo-$\mathscr{T}_j$ 8-tuples. Fig. 4 shows the three types of statistics based on 7_1-tuples from a 400 ps MWMD simulation, namely non-$\mathscr{T}$, pseudo-$\mathscr{T}_1$, and pseudo-$\mathscr{T}$ statistics; these statistics are respectively based on (*a*) all proper 7_1-tuples, (*b*) all proper 7_1-tuples plus all pseudo-$\mathscr{T}_1$ 7_1-tuples, and (*c*) all proper 7_1-tuples plus the set of all pseudo-$\mathscr{T}_k$ 7_1-tuples where k is any individual dihedral angle $\varphi_j, j = 1, \ldots, 8$, or any combination of any number (2–8) of dihedral angles. Pseudo-$\mathscr{T}$ statistics thus includes all structures in the simulation and describes the sampling of dihedral conformation space during MWMD. Non-$\mathscr{T}$ statistics describes the distribution of stable conformations in the entire conformational space, whereas pseudo-$\mathscr{T}_1$ statistics also describes distribution of stable conformations but ignores the high rate of rotation of the amino-terminal dihedral angle ϕ_1. The comparison of non-$\mathscr{T}$ and pseudo-$\mathscr{T}_1$ statistics in Fig. 4

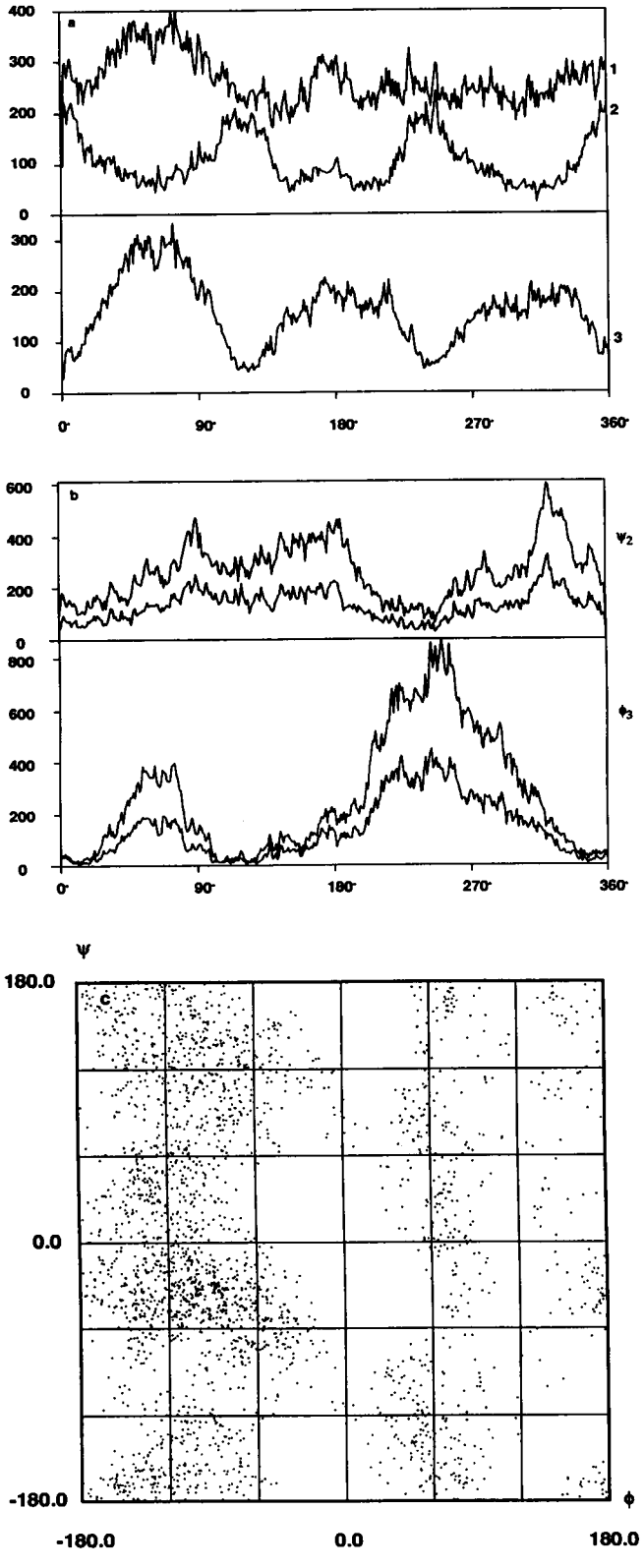

TABLE 1 Potential wells for FMRF-amide backbone dihedral angles

		1	2	3
Phe_1	ϕ	3.4–120.6°	120.6–239.4°	239.4–3.4°
	ψ	35.3–233.4°	233.4–35.3°	—
Met_2	ϕ	350.5–133.1°	133.1–350.5°	—
	ψ	8.9–241.2°	241.2–8.9°	—
Arg_3	ϕ	12.1–115.9°	115.9–12.1°	—
	ψ	14.7–263.9°	263.9–14.7°	—
Phe_4	ϕ	357.4–126.1°	126.1–357.4°	—
	ψ	18.1–244.9°	244.9–18.1°	—

Each potential well is specified by two numbers (in degrees) defining its lower and upper bounds (which are obtained from Fig. 3 *b*). If the leading number is larger, then the range includes 0° (360°). The eight dihedral angles are also referred to as $\varphi_j, j = 1, 2, \ldots, 8$; φ_1 is ϕ in Phe_1 at the amino-terminus.

shows that rotation of the amino-terminal amine group does not provide any significant bias in the distribution of stable conformations.

The convergence of conformational sampling in an MWMD simulation can be monitored as shown in Fig. 5.

From the pseudo-$\mathcal{T}_1$ statistics shown in Fig. 4, the population of some of the stable conformations for FMRF-amide can be obtained (Table 2). The following conclusions can be made with regard to the preferred conformations of the molecule. First, 212•••• is the least populated set of conformations, followed by 111••••, 112••••, and 211••••. Second, as a group, •••2••• accounts for the majority of stable conformations. Third, 12•2••• is more populated than 22•2•••. Last, •22212• is the most populated set of conformations; conformations in either potential well for φ_2 (i.e., ψ_1) and in either well for φ_8 (i.e., ψ_4, which directs the terminal amide group) contribute nearly equally, and Fig. 6 shows a representative structure in this group; the structure specifically belongs in the 7_1-tuple 2222121. As a reference, the starting

FIGURE 3 Conformational sampling of dihedral angles. (*a*) Histograms for φ_1 (i.e., backbone dihedral angle ϕ of Phe_1). Each histogram is based on bins of 1° width. Numbers of time steps at which the value $\varphi(t)$ falls within the 1° bins are accumulated and plotted. Curve 1, histogram from $\varphi(t)$ for all t; curve 2, from $\varphi(t)$ for all time steps at which φ_1 is in transitional state as determined according to the scheme shown in Fig. 2 *b*; curve 3, from $\varphi(t)$ for all t's at which φ_1 is not in transitional state. From curve 3, the dihedral rotation for φ_1 in FMRF-amide thus has three potential wells of a threefold symmetry. (*b*) Histograms for φ_4 (i.e., backbone dihedral angle ψ of Met_2) and φ_5 (i.e., backbone dihedral angle ϕ of Arg_3). For each dihedral angle, the upper curve is the histogram from $\varphi(t)$ for all t, and the lower curve is from $\varphi(t)$ for all t's at which φ is not in transitional state. Potential wells listed in Table 1 are defined by dihedral angle values at which the histograms are at minima in the lower curve of each φ. (*c*) ϕ–ψ plot for residue Arg_3.

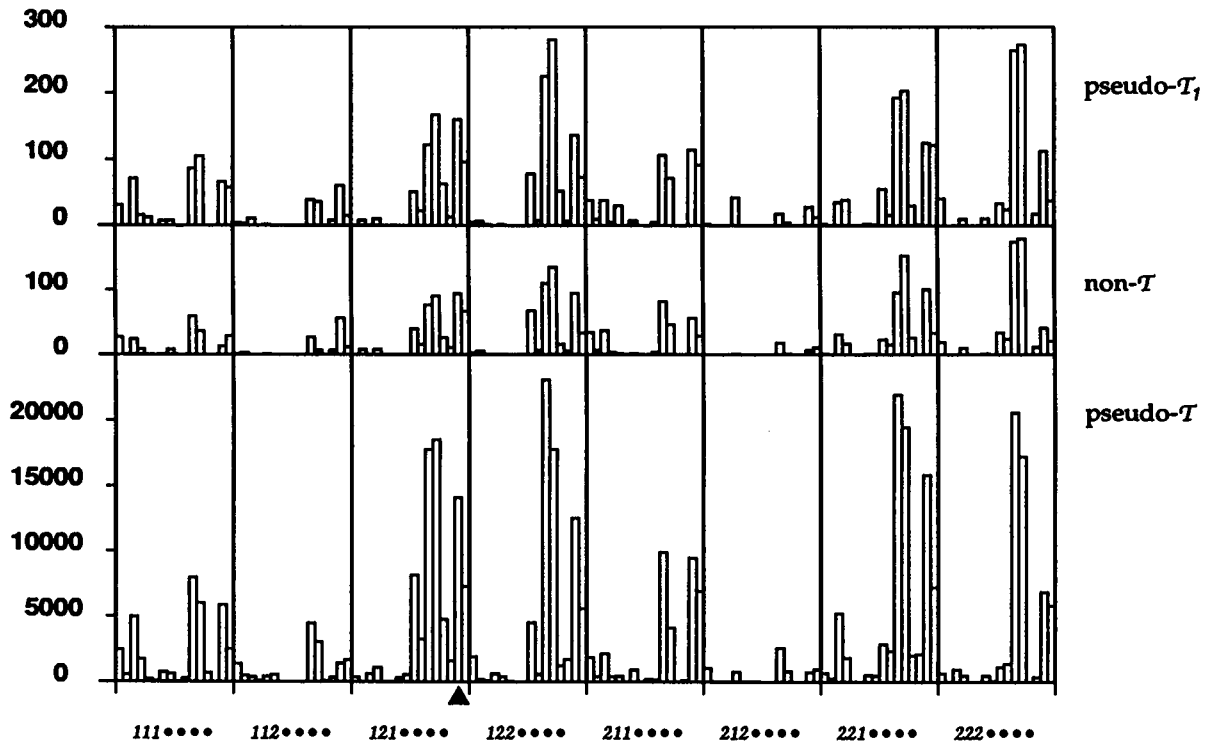

FIGURE 4 Statistics of conformational states of FMRF-amide. Low-resolution structures of the molecule are specified by 7_1-tuples as described in the text. All 128 7_1-tuples are displayed linearly and divided into eight groups; each group is specified by a label in which (•) represents any of the numeric labels for the dihedral angle in its position. For example, the group represented by 111•••• includes 16 7_1-conformations $111\mathcal{L}_5\mathcal{L}_6\mathcal{L}_7\mathcal{L}_8$, where each $\mathcal{L}_j (j = 5, \ldots, 8)$ can be either 1 or 2; 1111111 is the first molecular conformation displayed for the group, followed by 1111112, 1111121, 1111122, ... etc. Pseudo-$\mathcal{T}$, non-$\mathcal{T}$, and pesudo-$\mathcal{T}_1$ statistics are defined in the text. In each type of statistics, the numbers of molecular structures from the MWMD trajectories in 7_1-tuple categories are displayed. The solid triangle points to the 7_1-tuple in which the energy-minimized structure of FMRF-amide (also the starting structure for MWMD) belongs.

conformation of FMRF-amide for MWMD belongs in the 7_1-tuple 1212221, also shown in Fig. 6 and marked by a filled triangle in Figs. 4 and 5. Fig. 7 shows that the molecular conformation of FMRF-amide during MWMD simulation undergoes transitions between various conformations without being trapped in any particular conformation for any length of time, and that the molecular conformation space is sampled continually during the simulation.

As a control experiment for the MWMD conformational sampling, a reference tetrapeptide amide, Gly-Gly-Gly-Gly-amide or GGGG-amide, is constructed and treated with energy minimization and mass-weighted molecular dynamics simulation in an exactly identical manner as for FMRF-amide. The histograms in Fig. 8*a* for two representative backbone dihedral angles in GGGG-amide show that values of ψ dihedral angle are more evenly distributed in the entire range in comparison with FMRF-amide, and that there is only one potential well for ϕ dihedral angles. When the dihedral space for dihedral angles φ_2 through φ_8 is divided into two potential wells of equal width, the conformational space sampling and the distribution of stable conformations can be analyzed in the same manner as for FMRF-amide. The pseudo-$\mathcal{T}$ statistics (Fig. 8*b*) shows that the conformational space is sampled much more uniformly, as expected for GGGG-amide in which there are no sidechain groups and for which less conformational preference is expected. Stable conformations for GGGG-amide (pseudo-$\mathcal{T}_1$ statistics in Fig. 8*b*) are also more evenly distributed, compared to the preference of some stable structure for FMRF-amide. These comparisons suggest that the preferred conformations for FMRF-amide shown in Fig. 4 result from intramolecular interactions among amino acid sidechain groups in the molecule.

DISCUSSION

We have demonstrated here that, by weighting the atomic mass in a molecular simulation system, the

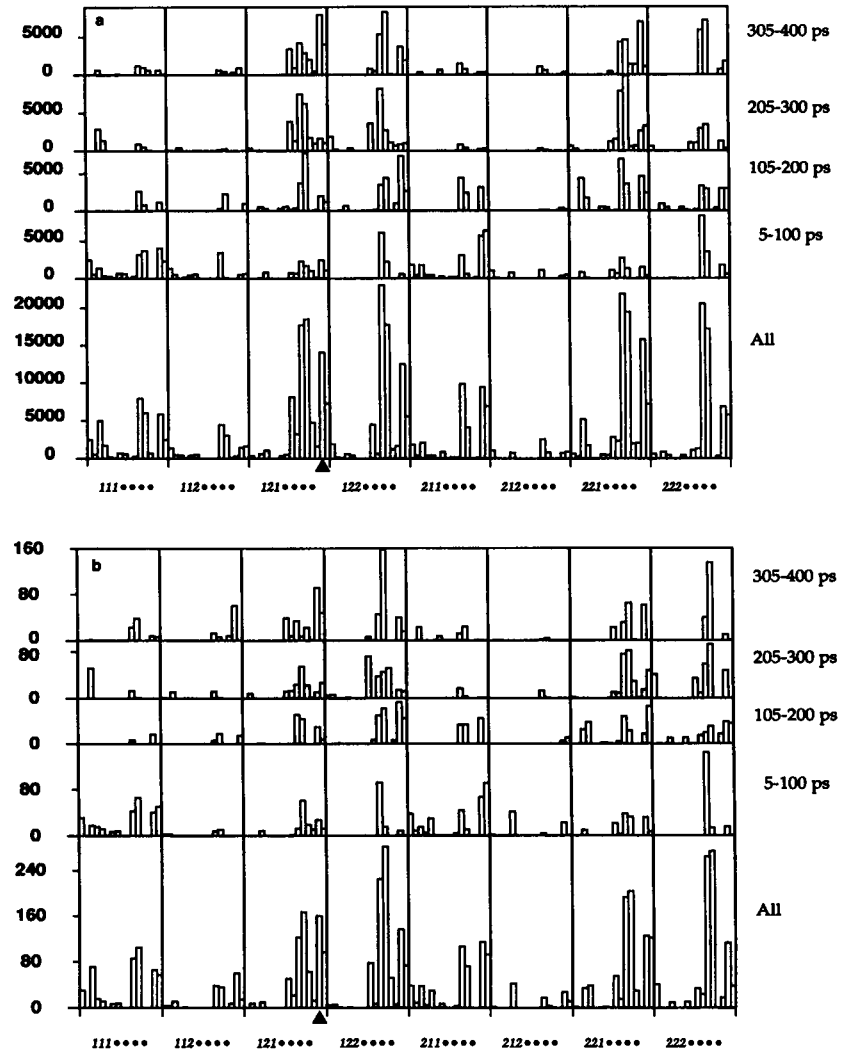

FIGURE 5 Monitoring the convergence of distributions of molecular conformations in different time periods of MWMD simulation. Molecular structures in the first 5 ps are not counted as the system temperature was first being scaled up; for comparing equal numbers of conformations, the first 5 ps periods in the remaining three 100 ps time segments are also not counted. (*a*) Pseudo-$\mathcal{T}$ statistics. (*b*) Pseudo-$\mathcal{T}_1$ statistics.

TABLE 2 Population of stable conformations for FMRF-amide from the statistics of pseudo-$\mathcal{T}_1$ 7_1-tuples

···1···	7.75%	···2···	92.25%
111····	5.21%	112····	4.74%
121····	17.43%	122····	23.64%
211····	6.30%	212····	1.07%
221····	21.23%	222····	20.38%
12·2···	40.39%	22·2···	37.54%
·21212·	16.93%	·21222·	13.23%
·22212·	24.33%	·22222·	10.47%

The total number of MWMD structures in this category in the last three time periods in Fig. 5 *b* (i.e., 105–200, 205–300, and 305–400 ps) is 3,189. Percentages of this total for 7_1-tuples in the table are listed.

sampling of dihedral conformational space of a short peptide molecule can be significantly improved. Moreover, it was shown that, from the mass-weighted molecular dynamics trajectories, molecular conformational space sampling can be analyzed comprehensively by numerical analyses of dihedral angle trajectories. Whereas it is necessary to increase the strength of covalent interactions in this mass-weighted molecular system, i.e. by placing constraints on terms in the potential energy function for bond stretching, bond angle bending, and tetrahedral carbon stereochemistry and amide planarity, intramolecular interactions relevant to dihedral conformational space sampling (i.e., dihedral angle rotations, hydrogen bond interactions, and nonbonded electro-

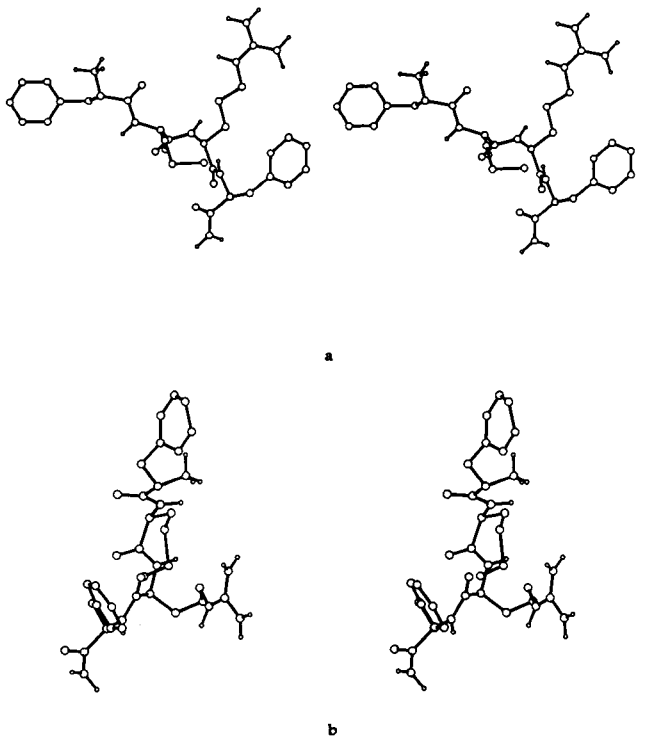

FIGURE 6 Molecular structures for FMRF-amide. (*a*) Energy-minimized conformation; also the initial structure for dynamics simulations. Backbone dihedral angles (ϕ, ψ) are (60.8, 138.4) for Phe_1, (278.5, 144.0) for Met_2, (280.8, 327.0) for Arg_3, and (266.5, 128.9) for Phe_4. This conformation belongs in the 8-tuple 32222121. (*b*) The conformation of FMRF-amide randomly extracted from the MWMD simulation (at t = 87.162 ps); the (ϕ, ψ) pairs are (343.2, 25.9), (285.2, 314.5), (266.9, 161.3), and (159.6, 133.2) in this structure. The conformation belongs in the 8-tuple 11212221, and is displayed without reorientation from that in the simulation.

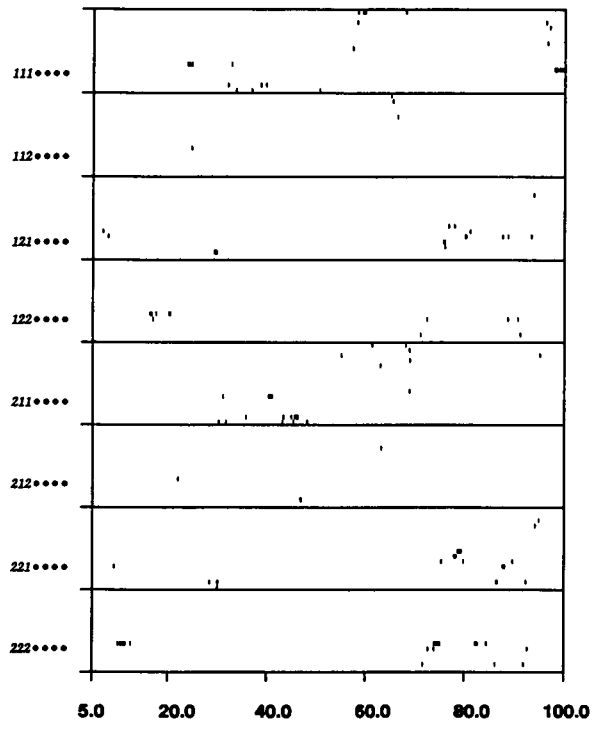

FIGURE 7 Time evolution of molecular conformation during the first 100 ps of MWMD simulation. The 0–5 ps period again is not displayed as the system temperature was being increased. Each short vertical bar indicates that the molecular conformation at the given t can be designated the given proper 7_1-tuple. The dihedral conformational space of the molecule is shown to be sampled continually throughout the time period of simulation.

static and van der Waals forces) remain unchanged from the normal unweighted molecular systems. Thus, the potential energy surface that represents noncovalent intramolecular interactions remain the same whereas the kinematics of dihedral rotations are modified by weighted atomic masses in the system. The mass weighting increases the effective momentum of dihedral rotation in barrier crossing and thus increases the sampling of conformations in different potential wells of backbone dihedral rotations. The mass weighting and internal coordinate constraints also increase the fluctuation of the total energy of the microcanonical ensemble generated in MWMD simulations which contributes to the crossing of dihedral rotational barriers.

The mass-weighted molecular dynamics trajectory can be analyzed comprehensively by the determination of potential wells in each dihedral angle space and the assignment of low-resolution stable conformers for FMRF-amide molecule. The analyses showed that stable conformers (i.e., proper 8-tuples and 7_1-tuples) exist in the simulation. Such stable conformations a priori may not have existed in the simulation in which case the molecule would have been flexible in the sense that at any given time step some backbone dihedral angles will be in rotational transitional state. As described in the previous section however, stable molecular conformations exist, and some of the stable backbone conformations for FMRF-amide are occupied more frequently than others, most likely due to intramolecular interactions involving sidechain groups of FMRF-amide. In contrast, both conformational sampling and the distribution of stable conformations for tetraglycyl amide are uniform among various molecular conformations. In the procedure described here, the stable conformation with a highest statistics (cf. Fig. 4) represents the most probable structure for the molecule.

Short peptides are expected to have quite flexible molecular structures and yet they can be induced to adopt specific conformations (Kaiser and Kézdy, 1984)

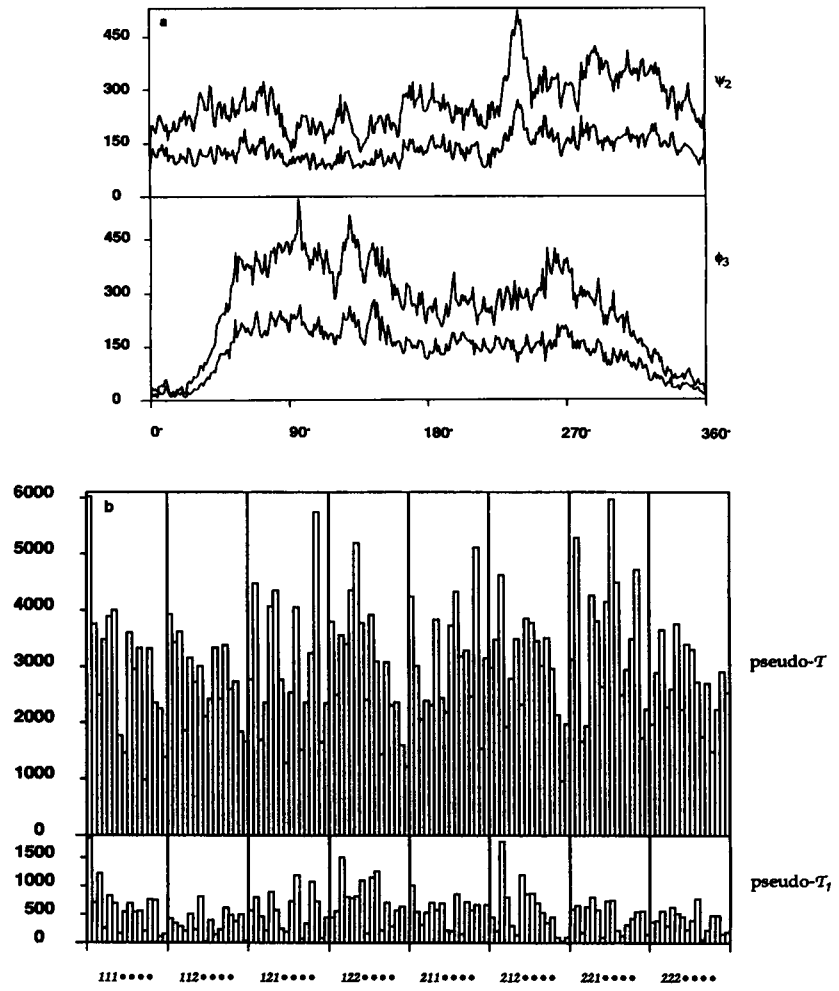

FIGURE 8 Conformational sampling of GGGG-amide in MWMD simulation. (*a*) Representative histograms for values of dihedral angles in GGGG-amide, obtained in the manner similar to those in Fig. 3 *b*. (*b*) Statistics for molecular conformations in MWMD trajectory of GGGG-amide, obtained in the manner similar to those shown in Fig. 4. The range of potential wells for GGGG-amide backbone dihedral angles are 0.0–116.6°, 116.6–244.1°, and 244.1–0.0° for ϕ of Phe_1, 94.7–275.2° and 275.2–94.7° for ψ of Phe_1; 0.0–180.5° and 180.5–0.0° for ϕ of Met_2; 131.8–312.3° and 312.3–131.8° for ψ of Met_2; 0.0–180.5° and 180.5–0.0° for ϕ of Arg_3; 163.7–344.2° and 344.2–163.7° for ψ of Arg_3; 16.2–196.7° and 196.7–16.2° for ϕ of Phe_4; 31.2–211.7° and 211.7–31.2° for ψ of Phe_4.

or to raise specific antibodies (Pain et al., 1990), and functional conformations of biologically active peptides can be determined by a variety of techniques (Taylor and Ösapay, 1990). On the other hand, the extensive conformational sampling by the mass-weighted simulation method made it possible to identify stable conformations independent of the starting structure and to monitor the convergence of simulation results. Moreover, the molecular flexibility can be quantitatively described on the basis of distribution and statistics of different stable conformations available for the molecule. Molecular structures in the MWMD simulation identified to be in low-resolution stable conformations can be further characterized by extending the molecular simulation via energy minimization and/or conventional molecular dynamics methods.

With the mass weighting, the MWMD method is expected to further enhance the efficiency of molecular dynamics-based methods in macromolecular conformational sampling in comparison with the Monte Carlo method (Northrup and McCammon, 1980). In the velocity-based approach for increasing conformational space sampling (i.e., high temperature simulations [Di Nola et al., 1984; Bruccoleri and Karplus, 1990]), the molecular dynamics integration becomes numerically unstable at high temperatures, with or without harmonic constraints on covalent interactions (Mao, B., in preparation). The MWMD method in contrast modifies the mass in the

momentum expression instead, and average atomic velocities in the simulation can thus be maintained in the regime where numerical integration remains stable; due to this numerical stability, the simulation does not require a smaller integration time step, which would have improved the numerical integration of the velocity-based high temperature simulations but would have required more processor time for the same total simulation time length (say 200 ps).

Whereas peptide conformation sampling has been studied by other methods (Bruccoleri and Karplus, 1990; Vasquez and Scheraga, 1988; Vajda and Delisi, 1990), the MWMD simulation and numerical analytical methods described here requires no predetermined parameters and appear to be more efficient. Although a continuum dielectric constant was chosen here to represent solvent environment of the peptide molecule, the mass-weighted simulation approach can be similarly applied if solvent effects were treated in a different manner. Whereas the presence of sidechain groups in FMRF-amide results in some preferred backbone conformations for the molecule (in contrast to the comparative results of GGGG-amide simulation), a detailed analyses of sidechain χ angles and hydrogen bond interactions will be essential in determining the details of intramolecular contributions from sidechain groups to the preferred backbone conformations shown in this study; conformation analyses of sidechain dihedral angles will be reported separately. Other future studies include the investigation of the nature of unpopulated conformations of FMRF-amide shown in Fig. 4.

The author would like to acknowledge helpful discussion with J. A. McCammon, A. R. Friedman, K. C. Chou, and G. M. Maggiora, and thank M. Karplus for making the Harvard macromolecular computation program CHARMM available.

Received for publication 3 January 1991 and in final form 21 May 1991.

REFERENCES

Al-Obeidi, F., M. E. Hadley, B. M. Pettitt, and V. J. Hruby. 1989. Design of a new class of superpotent cyclic α-melanotropins based on quenched dynamic simulations. *J. Am. Chem. Soc.* 111:3413–3416.

Bennett, C. H. 1975. Mass tensor molecular dynamics. *J. Comput. Phys.* 19:267–279.

Brooks, B., and M. Karplus. 1983. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA.* 80:6571–6575.

Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* 4:187–217.

Brooks, C. L. III, M. Karplus, and B. M. Pettitt. 1988. Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. John Wiley & Sons, New York. 259 pp.

Bruccoleri, R. E., and M. Karplus. 1990. Conformational sampling using high-temperature molecular dynamics. *Biopolymers.* 29:1847–1862.

Di Nola, A., H. J. C. Berendsen, and E. Edholm. 1984. Free energy determination of polypeptide conformations generated by molecular dynamics. *Macromolecules.* 17:2044–2050.

Harvey, S. C., and J. A. McCammon. 1981. Intramolecular flexibility in phenylalanine transfer RNA. *Nature* (Lond.). 294:286–287.

Ichiye, T., and M. Karplus. 1987. Anisotropy and anharmonicity of atomic fluctuations in proteins: analysis of a molecular dynamics simulation. *Proteins: Struct. Funct. Genet.* 2:236–259.

Kaiser, E. T., and F. J. Kézdy. 1984. *Science (Wash. DC).* 223:249–255.

Kitson, D. H., and A. T. Hagler. 1988. Theoretical studies of the structure and molecular dynamics of a peptide crystal. *Biochemistry.* 27:5246–5257.

Levy, R. M., R. P. Sheridan, J. W. Keepers, G. S. Dubey, S. Swaminathan, and M. Karplus. 1985. Molecular dynamics of myoglobin at 298°K. Results from a 300-ps computer simulation. *Biophys. J.* 48:509–518.

Man-Son-Hing, H., M. J. Zoran, K. Lukowiak, and P. G. Haydon. 1989. A neuromodulator of synaptic transmission acts on the secretory apparatus as well as on ion channels. *Nature (Lond.).* 341:237–239.

Mao, B., and A. R. Friedman. 1990. *Biophys. J.* 58:803–805.

Mao, B., and J. A. McCammon. 1983. Theoretical study of hinge bending in L-arabinose-binding protein. *J. Biol. Chem.* 258:12543–12547.

Mao, B., M. R. Pear, J. A. McCammon, and S. H. Northrup. 1982. Molecular dynamics of ferrocytochrome *c*: anharmonicity of atomic displacements. *Biopolymers.* 21:1979–1989.

McCammon, J. A., and S. C. Harvey. 1987. Dynamics of Proteins and Nucleic Acids. Cambridge University Press, Cambridge. 234 pp.

Northrup, S. H., and J. A. McCammon. 1980. Simulation methods for protein structure fluctations. *Biopolymers.* 19:1001–1016.

Northrup, S. H., M. R. Pear, J. D. Morgan, and J. A. McCammon. 1981. Molecular dynamics of ferrocytochrome *c*. Magnitude and anisotropy of atomic displacements. *J. Mol. Biol.* 153:1087–1109.

Pain, D., H. Murakami, and G. Blobel. 1990. Identification of a receptor for protein import into mitochondria. *Nature (Lond.).* 347:444–449.

Pomès, R., and J. A. McCammon. 1990. Mass and step length optimization for the calculation of equilibrium properties by molecular dynamics simulation. *Chem. Phys. Lett.* 166:425–428.

Post, C. B., C. M. Dobson, and M. Karplus. 1989. A molecular dynamics analysis of protein structural elements. Proteins: Struct. Funct. Genet. 5:337–354.

Ramachandran, G.N., C. M. Venkatachalam, and S. Krimm. 1966. Stereochemical criteria for polypeptide and protein chain conformations. III. Helical and hydrogen-bonded polypeptide chains. *Biophys. J.* 6:849–872.

Sessions, R. B., P. Dauber-Osguthorpe, and D. J. Osguthorpe. 1988. Filtering molecular dynamics trajectories to reveal low-frequency collective motions: phospholipase A_2. *J. Mol. Biol.* 209:617–633.

Sweatt, J. D., A. Volterra, B. Edmonds, K. A. Karl, S. A. Siegelbaum, and E. R. Kandel. 1989. FMRFamide reverses protein phosphorylation produced by 5-HT and eAMP in *Aplysia* sensory neurons. *Nature (Lond.).* 342:275–278.

Taylor, J. W., and G. Ösapay. 1990. Determining the functional conformations of biologically active peptides. *Acct. Chem. Res.* 23:338–344.

Thaisrivongs, S., B. Mao, D. T. Pals, S. R. Turner, and L. T. Kroll. 1990. Renin inhibitory peptides. A β-aspartyl residue as a replacement for the Histidyl residue at the P-2 site. *J. Med. Chem.* 33:1337–1343.

Vajda, S., and C. Delisi. 1990. Determining minimum energy conformations of polypeptides by dynamic programming. *Biopolymers.* 29:1755–1772.

Vasquez, M., and H. A. Scheraga. 1988. Calculation of protein conformation by the build-up procedure. Application to bovine pancreatic trypsin inhibitor using limited simulated nuclear magnetic resonance data. *J. Biomol. Struct. Dyn.* 5:705–755.

Verlet, L. 1967. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* 159:98–103.

Wood, D. W. 1979. Computer simulation of water and aqueous solutions. *In* WATER A Comprehensive Treatise. Vol. 6. Franks, F., editor. Plenum Publishing Corp., New York. 279–410.