Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 18, Issue Number 5, (2001)

Conformational Space Comparison of GnRH and lGnRH-III using Molecular Dynamics, Cluster Analysis and Monte Carlo Thermodynamic Integration.

http://www.adeninepress.com

Charles R. Watts [1], Mihaly Mezei[2], Richard F. Murphy[1] and Sándor Lovas[1]*

[1]Department of Biomedical Sciences, School of Medicine, Creighton University, 2500 California Plaza, Omaha, NE, 68178, USA

[2]Department of Physiology and Biophysics, Mount Sinai School of Medicine, NYU, New York, NY 10029

Abstract

The conformational space available to GnRH and lGnRH-III was compared using 5.2 ns constant temperature and pressure molecular dynamics simulations with explicit TIP3P solvation and the AMBER v. 5.0 force field. Cluster analysis of both trajectories resulted in two groups of conformations. Results of free energy calculations, in agreement with previous experimental data, indicate that a conformation with a turn from residues 5 through 8 is preferred for GnRH in an aqueous environment. By contrast, a conformation with a helix from residues 2 through 7 with a bend from residues 6 through 10 is preferred for lGnRH-III in an aqueous environment. The side chains of His^2 and Trp^3 in lGnRH-III occupy different regions of phase space and participate in weakly polar interactions different from those in GnRH. The unique conformational properties of lGnRH-III may account for its specific anti cancer activity.

Introduction

Mammalian gonadotropin releasing hormone (GnRH) and the sea lamprey gonadotropin releasing hormone type III (lGnRH-III), Figure 1, belong to the class of conserved gonadotropin releasing hormone peptides. lGnRH-III has virtually no endocrine activity in mammals even at high doses and has been shown to suppress directly sex hormone-dependent and -independent growth of breast and prostatic cancer cells *in vitro* (1). These properties make lGnRH-III an excellent starting compound for the development of constrained peptide analogs with increased anticancer activity.

```
              1   2   3   4   5   6   7   8   9   10
GnRH:      pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH2
GnRH-III:  pGlu-His-Trp-Ser-His-Asp-Trp-Lys-Pro-Gly-NH2
```

Figure 1:
The aligned sequences of GnRH and 1GnRH-III. The non-conserved central residues 5 through 8 are underlined and bold.

GnRH and lGnRH-III share similar sequence in the N-terminal region, residues 1 through 4, and C-terminal region, residues 9 through 10, while differing in the central region, residues 5 through 8. Structure-activity studies of GnRH (2) revealed several key features: no single residue is crucial for activity, the N– and C-terminal regions are the most important for receptor binding and activation and the only residues for which there is good evidence for a role in receptor activation are His^2 and Trp^3. The non-conserved residues of the central region are less critical for receptor binding with the exception of Arg^8 which is required for high affinity binding to the human GnRH receptor. Similar information on structure-activity relationships in lGnRH-III is unavailable with the exception of results of studies with constrained analogs that were used to rule out the presence of an electrostatic salt bridge between Asp^6 and Lys^8 (3).

Phone: +1 402 280 5753
Fax: +1 402 280 2690
E-mail: vasz@bif1.creighton.edu

Table I

Atom names, AMBER atom types and partial atomic charges of the L-pyroglutamyl residue.

Atom Name	Atom Type	Partial Charge
N	N	-0.6138
H	H	0.3559
CA	CT	0.1340
HA	H1	0.0266
CB	CT	-0.0223
HB2	HC	0.0206
HB3	HC	0.0206
CG	CT	-0.0506
HG2	HC	0.0401
HG2	HC	0.0401
CD	C	0.6587
OE	O	-0.6067
C	C	0.5569
O	O	-0.5601

The conformation of GnRH and several constrained analogs has been investigated via nuclear magnetic resonance spectroscopy, circular dichroism and molecular modeling. These techniques have indicated that the peptide is either disordered and sampling random conformations or forms a β-turn from residues 5 through 8 (2). The goal of this investigation was to use molecular dynamics, cluster analysis and Monte Carlo thermodynamic integration to elucidate conformational differences between GnRH and lGnRH-III.

Methods

Parameterization. The pyroglutamyl residue was parameterized for use in the 1994 version of the AMBER force field in a manner consistent with that used by Kollman and coworkers (4-7), Table I. The initial geometry of L-pyroglutamyl-N-methylamide was generated using the SYBYL v. 6.4 program (8). Two low energy conformations were generated by rotation around the ψ dihedral so that the carbonyl carbon was either *cis* or *trans* to the amide hydrogen. Partial atomic charges were then calculated using the RESP module (4-7) of AMBER v. 5.0 (9-10) and the HF/6-31G* electrostatic potential surface generated for both conformations by GAMESS (11). All other force field parameters were obtained from previously published AMBER parameters.

Simulated Annealing. All simulated annealing calculations entailed an infinite cutoff distance for generation of the non-bonded list, a constant dielectric of 1.0, GB/SA implicit solvation (12), a tight coupling constant of 0.1 ps for the temperature bath (13), scaling constants of 2.0 and 1.2 for the 1-4 van der Waals and electrostatic interactions respectively, the SHAKE algorithm (14) with a tolerance of 0.0005 Å to constrain all bond lengths, a time constant of 2 fs for the time step and the 1994 version of the AMBER force field (6) as implemented in the program package TINKER v 3.4 (15). Extended conformations of GnRH and lGnRH-III were generated using the program SYBYL v 6.4 (8). These conformations were energy minimized using the MINIMIZE module of TINKER to a gradient of 0.0001 kcal/mol/Å^2. The simulated annealing protocol (16) consisted of 500 repeated cycles of rapidly heating the peptides to a constant temperature of 1050 K for 2 ps and cooling the peptides exponentially from 1050 K to 50 K in 2 ps using the ANNEAL module of TINKER. The starting conformation for each new iteration of the cycle was the annealed conformation from the previous iteration. The 500 annealed conformations were then energy minimized using the NEWTON module of TINKER to a final gradient of 0.0001 kcal/mol/Å^2.

Molecular Dynamics. The lowest energy conformations of GnRH and lGnRH-III obtained from the simulated annealing protocol described above, were used as the starting structures for MD simulations. The peptides were immersed in boxes of TIP3P water (17) so that the boundaries of the box extended 10 Å beyond the furthest extension of the peptide in each of the x, -x, y, -y, z and -z directions. All water molecules with oxygen atoms less than 2.8 Å or hydrogen atoms less than 2.0 Å from the peptide were removed. The GnRH and lGnRH-III systems were solvated with 947 and 1102 water molecules, respectively. MD calculations entailed a 10.0 Å cutoff distance for generation of the non-bonded list, updating of the non-bonded list every 20 steps, a constant dielectric of 1.0, a constant temperature of 300 K, a constant pressure of 1 atm, coupling constants of 0.4 ps for the temperature and pressure (13), scaling constants of 2.0 and 1.2 for the 1-4 van der Waals and electrostatic interactions respectively, the SHAKE algorithm (14) with a tolerance of 0.0005 Å to constrain all bond lengths, a time constant of 2 fs for the time step and the 1994 version of the AMBER force field (6) as implemented in the SANDER module of AMBER v 5.0 (9-10).

Each simulation consisted of 10 steps steepest descent and 100 steps conjugate gradient minimization of the solvent molecules with the solute atoms frozen. The solvent was then allowed to relax under conditions of constant temperature and pres-

sure dynamics for 50 ps. The entire system (solute and solvent) was then energy minimized for 100 steps of steepest descent and 10000 steps of conjugate gradient minimization until the gradient was less than 0.1 kcal/mol/Å^2. The system was then equilibrated for 100 ps constant volume dynamics and 100 ps of constant pressure dynamics. A production run consisted of 5000 ps constant temperature and pressure dynamics with the coordinates recorded every 1.0 ps.

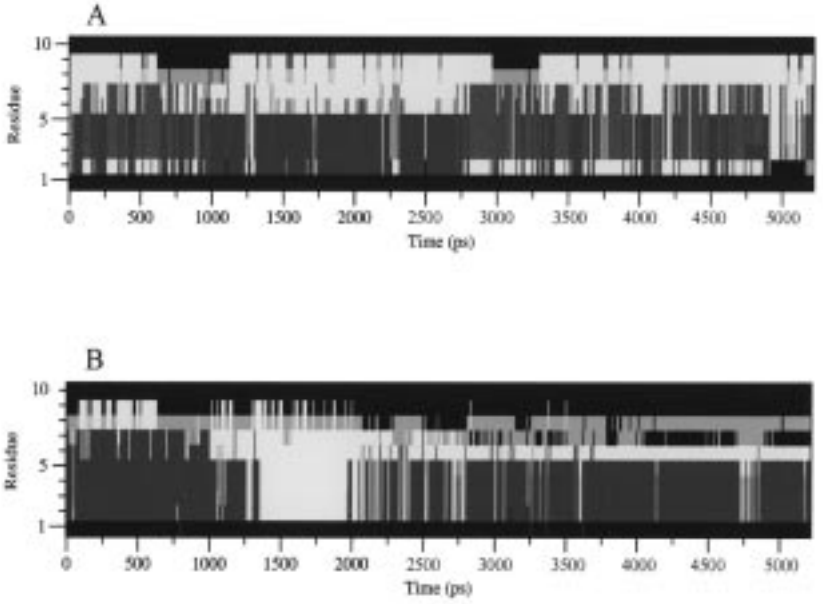


Figure 2: DSSP secondary structure as a function of simulation time for the trajectories of GnRH (A) and lGnRH-III (B).

Analysis. A lower triangle dissimilarity matrix of the backbone (-N-Cα-C-) RMSD values was calculated for all pairs of the 5000 sampled conformations of GnRH and lGnRH-III using the CARNAL module of AMBER v. 5.0 (9-10) and a PERL script. The dissimilarity (backbone RMSD) data were then clustered into families of related structures using the method of partitioning around medoids as implemented in the program PAM (18) using a minimum of 2 and a maximum of 10 clusters. The selection of the number of clusters which represented the optimal grouping of the dissimilarity data was based on the clustering with the highest average silhouette width and lowest number of negative silhouette widths (18,19). The average, standard deviation, maximum and minimum silhouette widths and dissimilarities for each cluster were calculated using a PERL script. The medoid conformations were extracted from the trajectories using the CARNAL module of AMBER v. 5.0 for further analysis.

The ϕ, ψ and χ^1 dihedral angles, backbone RMSDs and secondary structure of the sampled conformations of GnRH and lGnRH-III were calculated using the ANAL module of AMBER v. 5.0 (9-10) and the DSSP program (20) integrated with a PERL script and generating graphical output using a PERL script and the xpm2ps utility from the GROMACS v 2.0 package(21). The DSSP program assigns secondary structure on the basis of hydrogen bonding patterns and simple geometric rules. A turn of either three, four or five residues is assigned on the basis of the presence of a hydrogen bond between residues *i* and *i*+2 for a three residue turn, *i* and *i*+3 for a four residue turn and *i* and *i*+4 for a five residue turn. The turn is marked at the *i*+1 position for a three residue turn, *i*+1 and *i*+2 for a four residue turn and *i*+1, *i*+2 and *i*+3 for a five residue turn. A helix is assigned on the basis of the presence of two consecutive turns at positions *i*-1 and *i*. The helix is marked at positions *i*, *i*+1 and *i*+2 for a 3_{10}-helix, *i*, *i*+1, *i*+2 and *i*+3 for an a-helix and i, *i*+1, *i*+2, *i*+3 and *i*+4 for a π-helix. A bend is defined as a non hydrogen bonded five residue turn with a curvature of at least 70° between the first three residues, *i*-2, *i*-1 and *i*, and the last three residues, *i*, *i*+1 and *i*+2- and is marked at position *i*. Those residues of low curvature without recognized hydrogen bonding patterns are

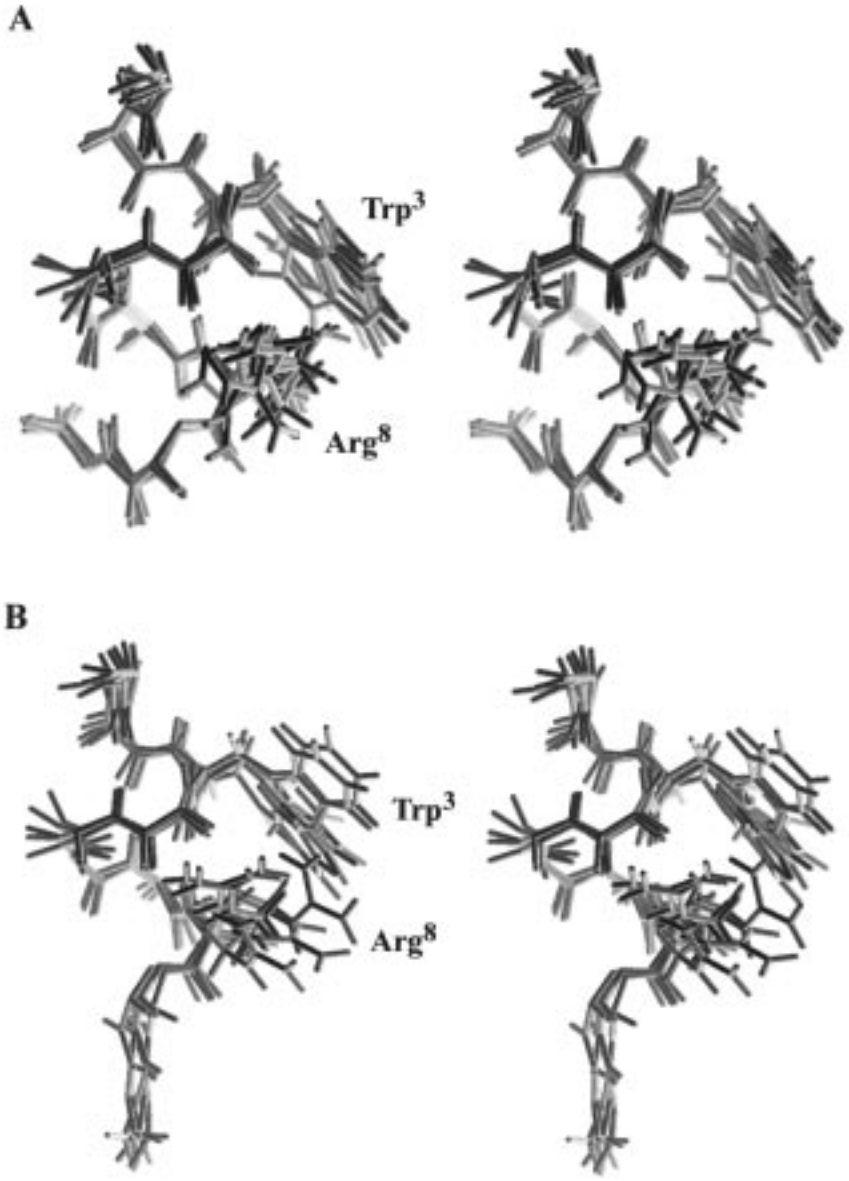


Figure 3: Stereo overlays of the five GnRH conformations with the highest silhouette width for cluster 1 (A) and cluster 2 (B).

assigned as random coil.

Weakly polar aromatic-aromatic (Ar-Ar) and aromatic-cationic (Ar-N^+) interactions between side chains were analyzed using the CARNAL module of AMBER v. 5.0 and were considered to occur when the average distance between the centers of mass of two groups was less than 7.0 Å and the average angle formed between the plane of the two rings, for Ar-Ar interactions, was greater than or equal to 33° and less than or equal to 147° (22). 1.0 Å was added for each Trp residue involved in an interaction.

The solvent contribution to the free energy difference between the medoid conformations of GnRH and lGnRH-III was calculated using thermodynamic integration (23) over a polynomial path (24). This path parameterizes the solute-solvent contribution to the system's energy as

$$E(\lambda)=\lambda^{k12}E_1^{12}+\lambda^{k6}E_1^{6}+(1-\lambda)^{k12}E_0^{12}+(1-\lambda)^{k6}E_0^{6}+(1-\lambda)^{k1}E_0^{1} \quad [1]$$

where λ is the coupling parameter between the free and hydrated solute molecule and E_i^{12}, E_i^{6} and E_i^{1} are the energy contributions containing the $1/r^{12}$, $1/r^{6}$ and $1/r$ terms, respectively, for the system i = 0 and 1. This parameterization leads to the

Table II

Backbone (-N-Cα-C-) RMSDs (Å) between all pairs of sampled conformations of the GnRH trajectory from 200 ps to 5200 ps.

Region	Avg.	Dev.	Max.	Min.
A. All conformations.				
whole (residues 1 to 10)	1.09	0.41	3.13	0.16
N-terminal (residues 1 to 4)	0.42	0.19	1.58	0.05
central (residues 5 to 8)	0.33	0.11	1.23	0.05
C-terminal (residues 9 to 10)	0.59	0.24	1.39	0.03
B. Cluster 1, 4138 conformations.				
whole (residues 1 to 10)	0.90	0.27	2.85	0.16
N-terminal (residues 1 to 4)	0.43	0.20	1.58	0.05
central (residues 5 to 8)	0.32	0.10	0.94	0.05
C-terminal (residues 9 to 10)	0.52	0.24	1.35	0.03
C. Cluster 2, 862 conformations.				
whole (residues 1 to 10)	0.97	0.33	3.13	0.22
N-terminal (residues 1 to 4)	0.35	0.14	1.35	0.05
central (residues 5 to 8)	0.34	0.12	1.23	0.07
C-terminal (residues 9 to 10)	0.58	0.22	1.31	0.04
D. Inter cluster dissimilarities.				
whole (residues 1 to 10)	1.54	0.32	3.05	0.26
N-terminal (residues 1 to 4)	0.39	0.18	1.49	0.06
central (residues 5 to 8)	0.35	0.12	1.11	0.06
C-terminal (residues 9 to 10)	0.75	0.14	1.39	0.04

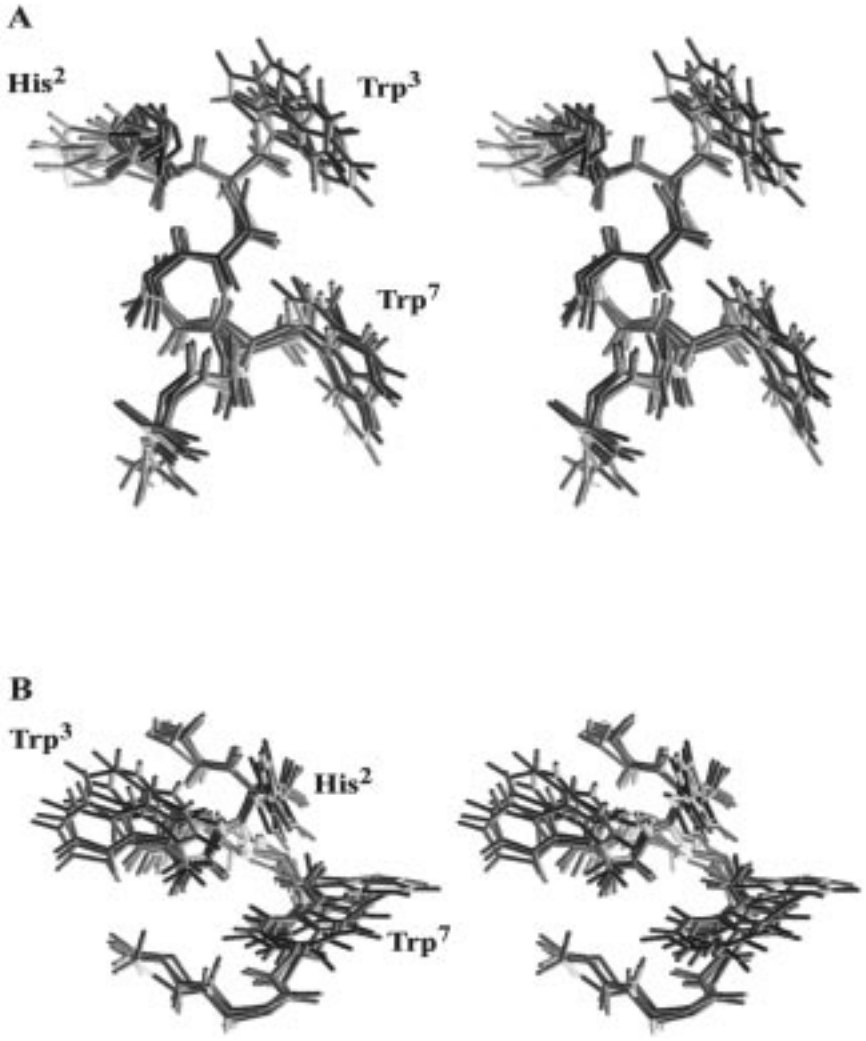


Figure 4: Stereo overlays of the five lGnRH-III conformations with the highest silhouette width for cluster 1 (A) and cluster 2 (B).

following expression for the solvation free- energy difference:

$$\Delta A=\int_0^1 k_{12}[\lambda^{k_{12}-1}\langle E_1^{12}\rangle_\lambda-(1-\lambda)^{k_{12}-1}\langle E_0^{12}\rangle_\lambda] + k_6[\lambda^{k_6-1}\langle E_1^{6}\rangle_\lambda-(1-\lambda)^{k_6-1}\langle E_0^{6}\rangle_\lambda] + k_1[\lambda^{k_1-1}\langle E_1^{1}\rangle_\lambda-(1-\lambda)^{k_1-1}\langle E_0^{1}\rangle_\lambda]d\lambda \quad [2]$$

Table III

Backbone (-N-Cα-C-) RMSDs (Å) between all pairs of sampled conformations of the lGnRH-III trajectory from 200 ps to 5200 ps.

Region	Avg.	Dev.	Max.	Min.
A. All conformations.				
whole (residues 1 to 10)	2.43	0.98	4.61	0.16
N-terminal (residues 1 to 4)	0.32	0.13	1.50	0.04
central (residues 5 to 8)	1.04	0.53	2.45	0.05
C-terminal (residues 9 to 10)	0.55	0.21	1.38	0.03
B. Cluster 1, 2436 conformations.				
whole (residues 1 to 10)	1.87	0.65	4.31	0.16
N-terminal (residues 1 to 4)	0.36	0.14	1.50	0.06
central (residues 5 to 8)	0.61	0.28	1.60	0.05
C-terminal (residues 9 to 10)	0.59	0.22	1.37	0.03
C. Cluster 2, 2564 conformations.				
whole (residues 1 to 10)	1.43	0.58	3.71	0.18
N-terminal (residues 1 to 4)	0.28	0.10	1.00	0.04
central (residues 5 to 8)	0.60	0.31	2.21	0.06
C-terminal (residues 9 to 10)	0.49	0.20	1.23	0.03
D. Inter cluster dissimilarities.				
whole (residues 1 to 10)	3.21	0.50	4.61	0.25
N-terminal (residues 1 to 4)	0.33	0.12	1.43	0.05
central (residues 5 to 8)	1.47	0.32	2.45	0.07
C-terminal (residues 9 to 10)	0.57	0.21	1.38	0.03

Here, $\langle ... \rangle_\lambda$ indicates that the ensemble averages were computed using $E(\lambda)$ as the solute-solvent energy contribution in the Boltzmann factor. Following earlier work on a similar system (25) the exponents k_{12} , k_6 and k_1 were chosen as 4, 3 and 2, respectively. This method neglects the change in intramolecular entropy upon the conformational change, and the difference in the degrees of diversity in the two clusters.

The integral of [2] was evaluated using a 5 point Gaussian quadrature. At each quadrature point, a canonical ensemble Monte Carlo calculation was performed on a system containing the combined solute and 3799 water molecules using the program MMC (26). During the simulations, the solute molecules were kept fixed in an orientation that had been optimized (27) using the program Simulaid (28). Face-centered cubic periodic boundary conditions with an inscribed sphere radius of 27.334 Å were used. The 1994 AMBER force field (6) was used for the peptides and the TIP3P model was used for the water (17). All calculations entailed 1×10^8 Monte Carlo steps at each quadrature point.

Results and Discussion

A. GnRH. Simulated annealing suggests that the lowest energy conformation of GnRH is a 3_{10}- helix from residues 2 through 5, a turn from residues 5 through 8 and a turn from residues 7 through 10. The regions of ϕ/ψ phase space were assigned using the method of Zimmerman and coworkers (29). Residues 2 through 5 are in the right handed α-helical region, residue 6 is in a left handed contiguous region, residues 7 through 9 are in the right handed α-helical region and residue 10 is in the right handed β-bridge region. The secondary structure is stabilized by the

Table IV.

Percentages of sampled DSSP secondary structure for the GnRH trajectory from 200 ps to 5200 ps.

Residue	Coil	π-helix	α-helix	3_{10}-helix	Turn	Bend
A. All conformations.						
His^2	5.5	3.1	52.8	0.4	38.1	0.0
Trp^3	0.0	37.0	53.4	2.0	6.9	0.6
Ser^4	0.0	37.0	53.5	2.0	7.5	0.0
Tyr^5	0.0	37.0	53.5	1.8	7.7	0.0
Gly^6	0.0	37.0	8.9	0.0	54.1	0.0
Leu^7	0.0	37.0	0.1	0.0	62.9	0.0
Arg^8	0.6	0.0	0.0	0.0	78.6	20.8
Pro^9	21.4	0.0	0.0	0.0	78.6	0.0
B. Cluster 1, 4138 conformations.						
His^2	6.3	2.4	53.2	0.4	37.7	0.0
Trp^3	0.0	35.7	53.8	2.4	7.5	0.6
Ser^4	0.0	35.7	53.8	2.4	8.1	0.0
Tyr^5	0.0	35.7	53.8	2.1	8.4	0.0
Gly^6	0.0	35.7	7.4	0.0	56.9	0.0
Leu^7	0.0	35.7	0.1	0.0	64.2	0.0
Arg^8	0.0	0.0	0.0	0.0	93.6	6.5
Pro^9	6.5	0.0	0.0	0.0	93.6	0.0
C. Cluster 2, 862 conformations.						
His^2	1.7	6.6	51.0	0.2	40.4	0.0
Trp^3	0.0	43.6	51.3	0.4	3.9	0.8
Ser^4	0.0	43.6	51.4	0.4	4.6	0.0
Tyr^5	0.0	43.6	51.4	0.4	4.6	0.0
Gly^6	0.0	43.6	15.7	0.0	40.7	0.0
Leu^7	0.1	43.6	0.1	0.0	56.2	0.0
Arg^8	3.7	0.0	0.0	0.0	6.6	89.7
Pro^9	93.4	0.0	0.0	0.0	6.6	0.0

Table V

Percentages of sampled DSSP secondary structure for the lGnRH-III trajectory from 200 ps to 5200 ps.

Residue	Coil	π-helix	α-helix	3_{10}-helix	Turn	Bend
A. All conformations.						
His^2	1.1	3.2	71.9	3.2	20.6	0.0
Trp^3	0.0	3.3	72.0	3.5	21.1	0.1
Ser^4	0.0	3.3	72.3	4.1	20.2	0.1
His^5	0.0	3.3	72.3	1.0	23.3	0.1
Asp^6	0.0	3.3	16.3	1.0	77.8	1.6
Trp^7	24.1	0.0	14.5	0.1	30.1	31.2
Lys^8	14.5	0.0	0.0	0.0	15.4	70.4
Pro^9	84.9	0.0	0.0	0.0	15.4	0.0
B. Cluster 1, 2436 conformation						
His^2	0.9	4.4	57.8	0.9	36.0	0.0
Trp^3	0.0	4.4	57.9	1.4	36.0	0.2
Ser^4	0.0	4.4	58.5	2.7	34.2	0.3
His^5	0.0	4.4	58.5	2.1	34.9	0.0
Asp^6	0.0	4.4	31.9	2.1	60.3	1.3
Trp^7	0.0	0.0	29.7	0.2	58.4	11.7
Lys^8	13.8	0.0	0.0	0.0	30.1	56.1
Pro^9	69.9	0.0	0.0	0.0	30.1	0.0
C. Cluster 2, 2564 conformation						
His^2	1.2	2.1	85.2	5.5	6.1	0.0
Trp^3	0.0	2.2	85.4	5.5	7.0	0.0
Ser^4	0.0	2.2	85.4	5.5	7.0	0.0
His^5	0.0	2.2	85.4	0.0	12.3	0.2
Asp^6	0.0	2.2	1.5	0.0	94.4	2.0
Trp^7	46.9	0.0	0.0	0.0	3.2	49.9
Lys^8	15.2	0.0	0.0	0.0	0.9	83.9
Pro^9	99.1	0.0	0.0	0.0	0.9	0.0

Table VI.

Distance (Å) between groups and atoms in weakly polar and hydrogen bond interactions, respectively, for the GnRH trajectory from 200 ps to 5200 ps*.

Interaction	Avg.	Dev.	Max.	Min.
A. All conformations.				
$Trp^{3}_{ind}\cdots Arg^{8}_{gua}$	6.18	1.10	10.30	3.25
$Tyr^{5}_{NH}\cdots pGlu^{1}_{C=O}$	3.01	0.21	3.99	2.59
$Gly^{6}_{NH}\cdots Trp^{3}_{C=O}$	3.46	0.27	4.00	2.70
$Leu^{7}_{NH}\cdots His^{2}_{C=O}$	3.18	0.34	4.00	2.62
$Arg^{8}_{NH}\cdots Trp^{3}_{C=O}$	2.96	0.17	3.98	2.59
$Gly^{10}_{NH}\cdots Gly^{6}_{C=O}$	3.13	0.29	4.00	2.58
B. Cluster 1, 4138 conformations.				
$Trp^{3}_{ind}\cdots Arg^{8}_{gua}$	6.16	1.13	10.30	3.25
$Tyr^{5}_{NH}\cdots pGlu^{1}_{C=O}$	3.01	0.21	3.99	2.60
$Gly^{6}_{NH}\cdots Trp^{3}_{C=O}$	3.46	0.28	4.00	2.70
$Leu^{7}_{NH}\cdots His^{2}_{C=O}$	3.18	0.34	4.00	2.64
$Arg^{8}_{NH}\cdots Trp^{3}_{C=O}$	2.97	0.18	3.98	2.60
$Gly^{10}_{NH}\cdots Gly^{6}_{C=O}$	3.13	0.29	4.00	2.58
C. Cluster 2, 862 conformations.				
$Trp^{3}_{ind}\cdots Arg^{8}_{gua}$	6.24	0.97	9.41	4.08
$Tyr^{5}_{NH}\cdots pGlu^{1}_{C=O}$	3.01	0.20	3.91	2.63
$Gly^{6}_{NH}\cdots Trp^{3}_{C=O}$	3.48	0.25	4.00	2.76
$Leu^{7}_{NH}\cdots His^{2}_{C=O}$	3.19	0.32	4.00	2.62
$Arg^{8}_{NH}\cdots Trp^{3}_{C=O}$	3.18	0.24	3.94	2.69
$Gly^{10}_{NH}\cdots Gly^{6}_{C=O}$	2.92	0.14	3.72	2.65

backbone hydrogen bonds, $Ser^{4}_{NH}\cdots pGlu^{1}_{C=O}$, $Tyr^{5}_{NH}\cdots His^{2}_{C=O}$, $Gly^{6}_{NH}\cdots Trp^{3}_{C=O}$, $Arg^{8}_{NH}\cdots Trp^{3}_{C=O}$, $Gly^{10}_{NH}\cdots Gly^{6}_{C=O}$ and $Gly^{10}_{NH2}\cdots Arg^{8}_{C=O}$; the side chain to backbone hydrogen bonds, $Ser^{4}_{OH}\cdots pGlu^{1}_{C=O}$ and $Arg^{8}_{NH2}\cdots Arg^{8}_{C=O}$; an Ar-Ar interaction, $His^{2}\cdots Tyr^{5}$ and an Ar-N^{+} interaction, $Trp^{3}\ldots Arg^{8}$. These results are consistent with those of previous computational and spectroscopic studies which indicate the presence of a β-turn structure from residues 5 through 8 (2,16).

During the MD simulations, the total potential energy, temperature and density of the GnRH peptide/solvent system came to equilibrium within the first 100 ps. The average backbone RMSDs of the peptide, Table IIA, are low and indicate a stable ensemble of closely related conformations with the greatest degree of conformational flexibility in the C-terminal region. The DSSP secondary structure of GnRH, as a function of simulation time, Figure 2A, and the percentages of sampled secondary structure, VA, show the presence of a helix from residues 2 through 5, a turn from residues 5 through 8 and a turn from residues 7 through 10. The secondary structures are stabilized by backbone hydrogen bonds, Table VIA, and an Ar-N^{+} interaction, $Trp^{3}\cdots Arg^{8}$. No side chain to side chain hydrogen bonds, side chain to backbone hydrogen bonds or Ar-Ar interactions between His^{2} and Tyr^{5} were observed, by contrast with the low energy conformation obtained from simulated annealing. The MD sampled conformations, however, are still consistent with the proposed central β-turn structure from residues 5 through 8 (2,16). These differences in conformation may be attributed to the less efficient sampling in the presence of explicit TIP3P waters in the MD simulations whereas an implicit GB/SA water model was employed in the simulated annealing calculations (30).

Clustering of the GnRH trajectory resulted in two groups of conformations with a silhouette coefficient of 0.40±0.13. Cluster 1 contains 4138 conformations with an average silhouette width of 0.40±0.19, a maximum of 0.58 and a minimum of -0.08. The average backbone RMSDs, Table IIB, indicate that the conformations within cluster 1 are closely related, Figure 3A, with the greatest degree of conformational stability in the central region and the greatest degree of conformational

flexibility in the C-terminal region. The medoid conformation of cluster 1 is an a-helix from residues 2 through 5, a turn from residues 5 through 8 and a turn from residues 7 through 10. This conformation is stabilized by the backbone hydrogen bonds, $Tyr^5_{NH}\cdots pGlu^1_{C=O}$, $Gly^6_{NH}\cdots His^2_{C=O}$, $Leu^7_{NH}\cdots His^2_{C=O}$, $Arg^8_{NH}\cdots Tyr^5_{C=O}$ and $Gly^{10}_{NH}\cdots Gly^6_{C=O}$ and an Ar-N^+ interaction, $Trp^3\cdots Arg^8$. The percentages of sampled secondary structure, Table IVB, are similar to those described previously for all

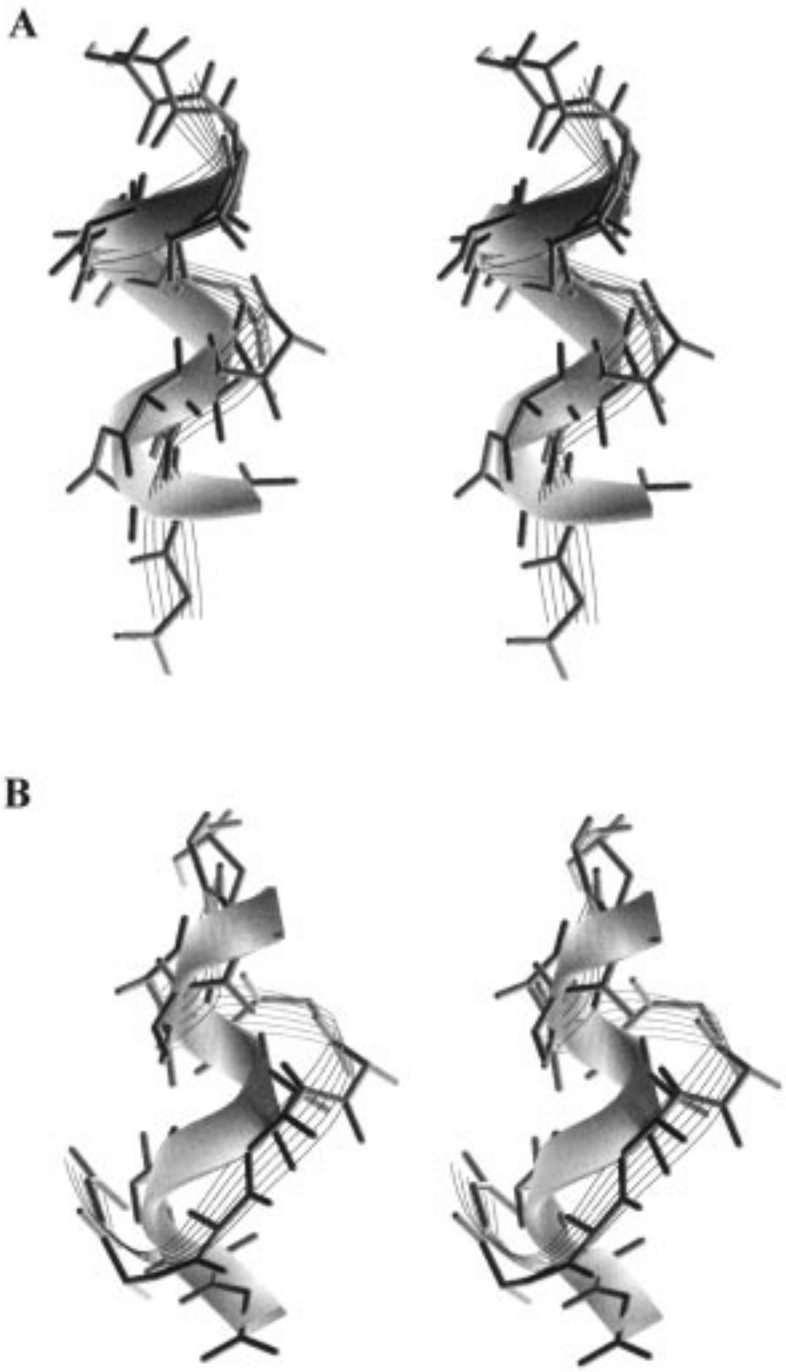


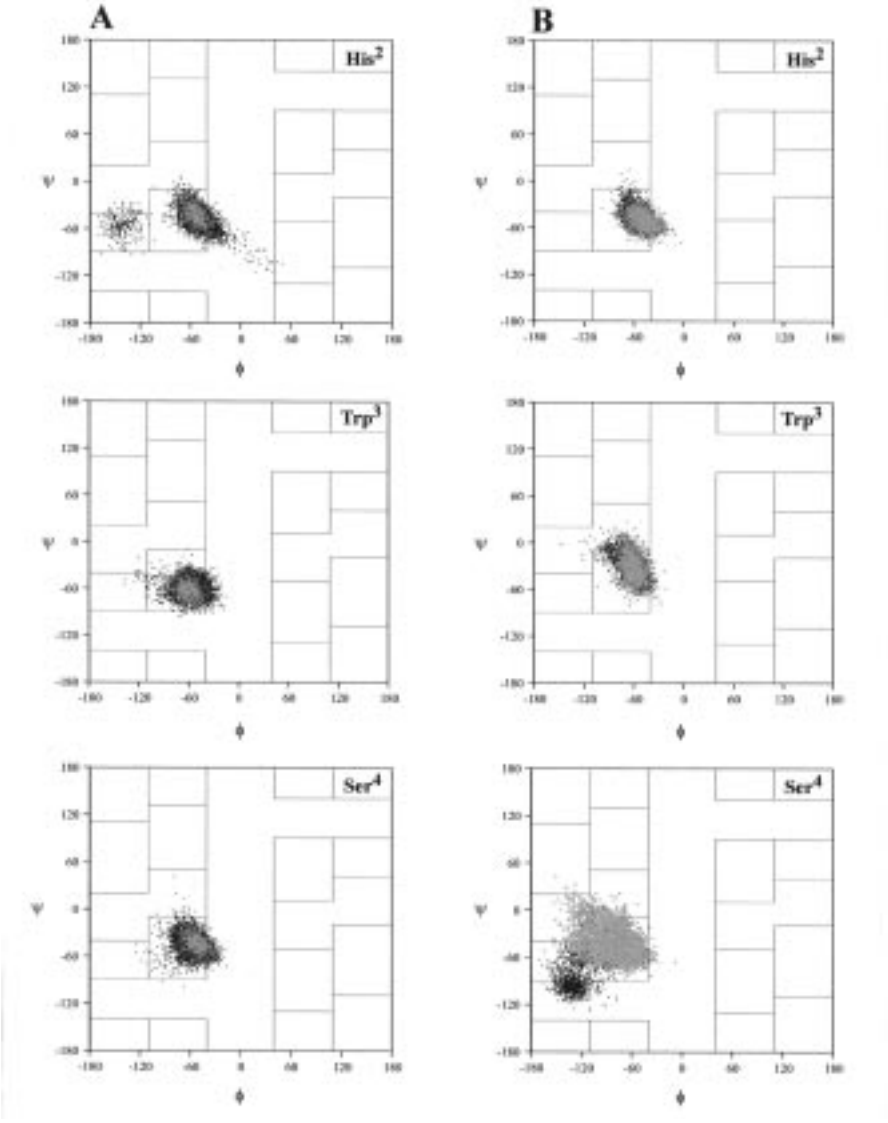
Figure 5: Stereo overlays of the medoid conformations of cluster 1, solid ribbon, and cluster 2, lined ribbon, of GnRH (A) and lGnRH-III (B).

conformations of GnRH and show the same pattern of backbone hydrogen bonds and weakly polar interactions, Table VIB.

Cluster 2 contains 862 conformations with an average silhouette width of 0.36±0.15, a maximum of 0.58 and a minimum of -0.09. The average backbone RMSDs, Table IIC, suggest that the conformations within cluster 2 are closely related, Figure 3B. The greatest degree of conformational stability is in the N-terminal and central regions with increased conformational flexibility in the C- ter-

minal region The medoid conformation of cluster 2 is a turn from residues 1 through 3, a π-helix from residues 3 through 7 and a bend from residues 6 through 10. This conformation is stabilized by the backbone hydrogen bonds, $Ser^4_{NH}\cdots pGlu^1_{C=O}$, $Tyr^5_{NH}\cdots pGlu^1_{C=O}$, $Leu^7_{NH}\cdots His^2_{C=O}$ and $Arg^8_{NH}\cdots Trp^3_{C=O}$; a side chain to backbone hydrogen bond, $Arg^8_{NH2}\cdots Ser^4_{C=O}$ and an Ar-N^+ interaction, Trp3···Arg8. The percentages of sampled secondary structure, Table IVC, are similar to those of cluster 1 with the exception of a change from a turn structure from residues 7 through 10 in the conformations of cluster 1 to a bend from residues 6 through 10 in the conformations in cluster 2. The backbone hydrogen bonds of the conformations in cluster 2 differ from those in cluster 1 by the absence of the $Gly^{10}_{NH}\cdots Gly^6_{C=O}$ hydrogen bond and formation of a $Leu^7_{NH}\cdots Trp^3_{C=O}$ hydrogen bond, Table VIC.

The average backbone RMSDs between the two clusters, Table IID, is larger than the average backbone RMSDs within each cluster. This observation confirms the distinct conformational differences between the two groups (18,19). The average regional backbone RMSDs show that the greatest difference between the two groups is due to the conformation adopted by the C-terminal region, Figure 5A. The ϕ/ψ dihedrals of clusters 1 and 2, Figure 6A, show that residues 2 through 8 sample similar regions of phase space. The clusters differ because the ϕ/ψ dihedrals of Pro^9 occupy the right handed α-helical region of phase space in cluster 1 and a righted handed contiguous region of phase space in cluster 2 (29). The χ^1 dihedral of Trp^3, Figure 7A, occupies only the trans region of phase space. This is consis-



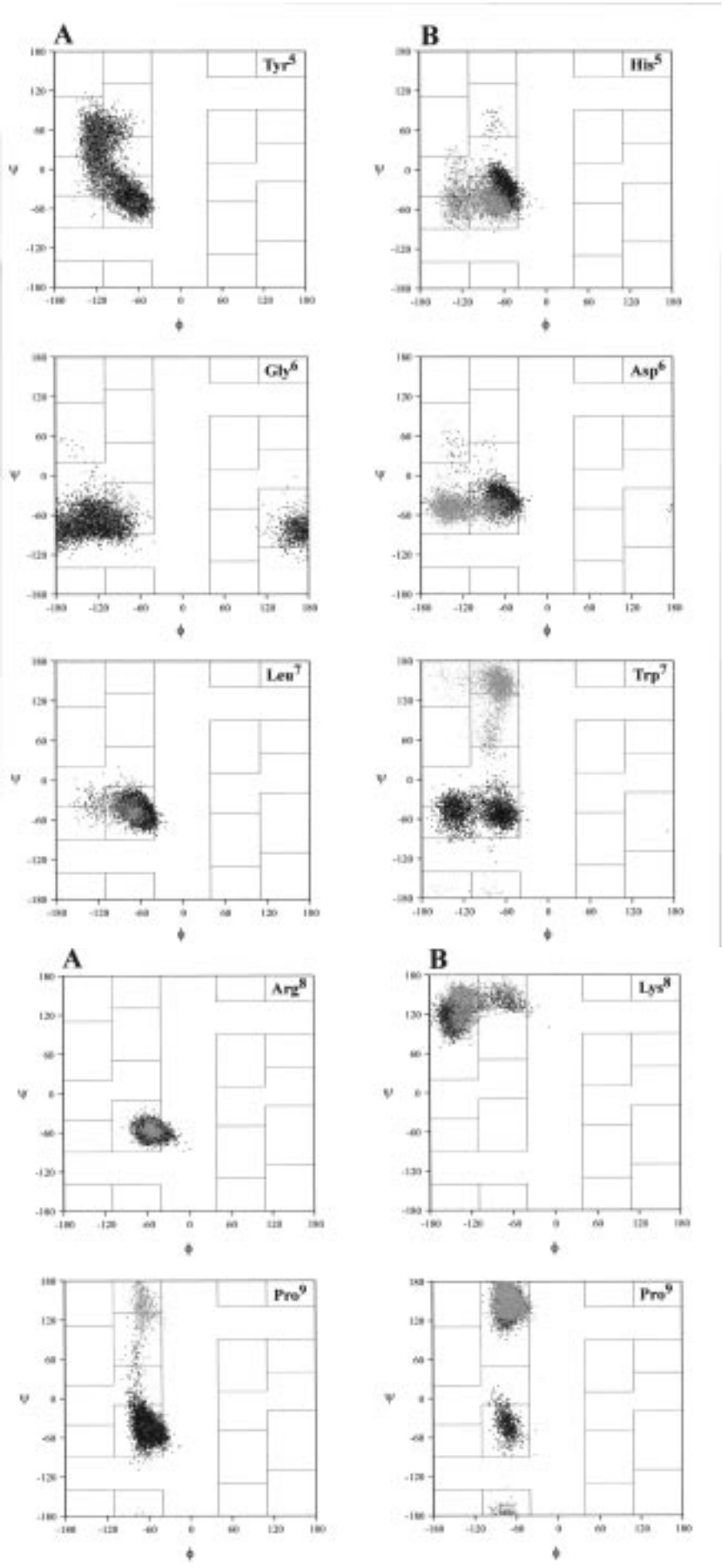
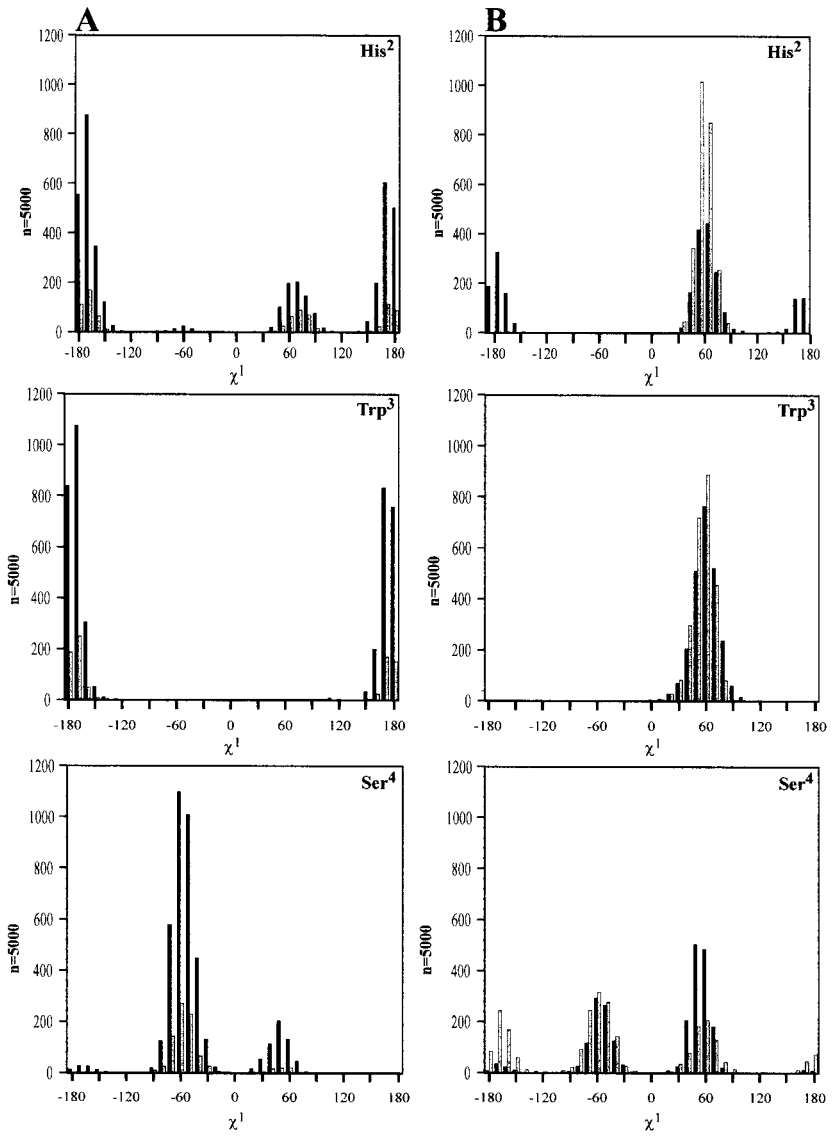


Figure 6: The ϕ/ψ dihedrals of residues 2 through 8 for cluster 1, black, and cluster 2, grey, of GnRH (A) and lGnRH-III (B) overlaid on to the regions of phase space as described by Zimmerman and coworkers (29).

tent with this residue being involved in a stable interaction that restricts its side chain flexibility. Otherwise, the phase space of the χ^1 dihedrals of clusters 1 and 2 differ only minimally suggesting that the change in backbone conformation is not due to differences in side chain to side chain interactions.

The solvent contribution to the free energy difference between the medoid conformations of cluster 1 and cluster 2 is $\Delta A_{\text{cluster 1}\rightarrow\text{cluster 2}} = 0.3\pm1.6$ kcal/mol, which indicates that solute-solvent interactions play only a minor role in stabilizing the conformations sampled in cluster 1 compared to those sampled in cluster 2. The net dipole moment of the medoid conformation in cluster 1 was 8.06 D and 7.40 D in cluster 2. Similar net dipole moments were expected because of the similar backbone geometry and low backbone RMSDs of the two medoid conformations.

B. lGnRH-III. Simulated annealing showed that the lowest energy conformation of lGnRH-III is an a-helix from residues 2 through 6, a bend from residues 5 through 9 and a bend from residues 6 through 10. The regions of phase space (29) occupied by the individual residues are, residues 2 through 6 are in the right handed a-helical region, residue 7 is in the right handed b-bridge region, residue 8 is in a left handed contiguous region, residue 9 is in a right handed contiguous region and residue 10 is in the left handed extended region. The secondary structure is stabilized by the backbone hydrogen bonds, $His^5{}_{NH}\cdots pGlu^1{}_{C=O}$, $Asp^6{}_{NH}\cdots His^2{}_{C=O}$,



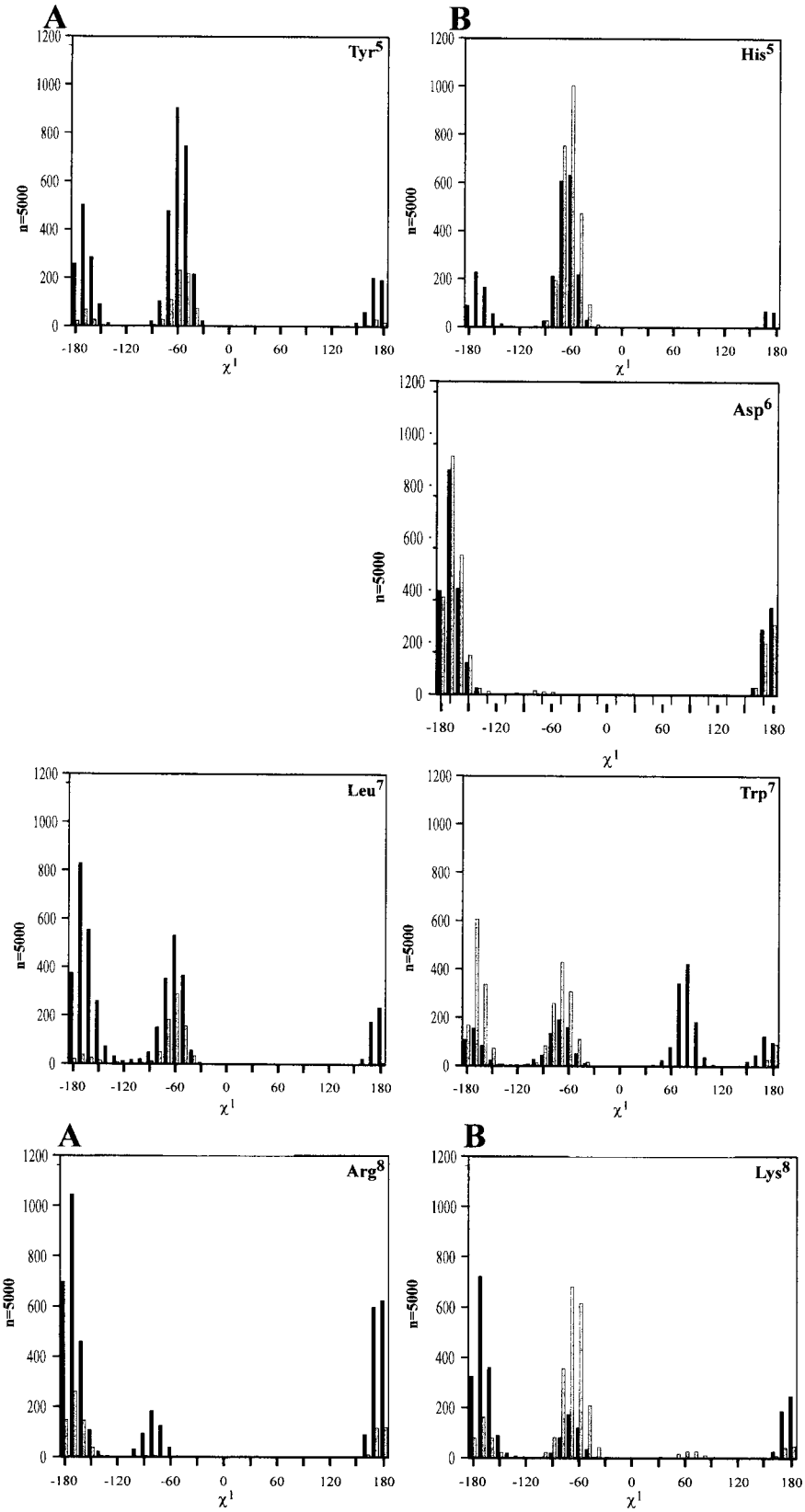


Figure 7: The χ^1 dihedrals of residues 2 through 8 for cluster 1, black, and cluster 2, grey, of GnRH (A) and lGnRH-III (B).

$Trp^7{}_{NH}\cdots Trp^3{}_{C=O}$ and $Lys^8{}_{NH}\cdots His^5{}_{C=O}$; the side chain to backbone hydrogen bonds, $pGlu^1{}_{NH}\cdots His^5{}_{NE2}$, $His^5{}_{ND1H}\cdots Pro^9{}_{C=O}$ and $Lys^8{}_{NH3}\cdots Gly^{10}{}_{C=O}$ and an Ar-Ar interaction, $Trp^3\cdots Trp^7$. A salt bridge between Asp^6 and Lys^8 was not found in any of the 500 annealed conformations. These results are consistent with those of a previous simulated annealing investigation (16) and show substantially different conformational features of lGnRH-III from those of GnRH.

During the MD simulation, the total potential energy, temperature and density of the lGnRH-III peptide/solvent system came to equilibrium within the first 100 ps. The average backbone RMSDs of the peptide, Table IIIA, are higher than those of GnRH with the greatest degree of conformational flexibility in the central region. The DSSP secondary structure of lGnRH-III as a function of simulation time, Figure 2B, and the percentages of sampled secondary structure, Table VA, show the presence of a helix from residues 2 through 5, a turn from residues 5 through 7 and a bend from residues 6 through 10. The secondary structures are stabilized by backbone hydrogen bonds, Table VIIA, and Ar-Ar interactions, $His^{2}\cdots Trp^{3}$, $His^{2}\cdots Trp^{7}$ and $Trp^{3}\cdots Trp^{7}$. By contrast, no interaction between His^{2} and Trp^{3} is observed in GnRH. A cationic side chain is present at position 8 in both peptides, Arg^{8} in GnRH and Lys^{8} in lGnRH-III, but, an Ar-N^{+} interaction is observed between Trp^{3} and Arg^{8} in GnRH whereas an Ar-Ar interaction is observed between Trp^{3} and Trp^{7} in lGnRH-III. This results in different backbone geometry including stabilization of an extended helical conformation in lGnRH-III.

Clustering of the lGnRH-III trajectory resulted in two groups of conformations with a silhouette coefficient of 0.48±0.15. Cluster 1 contains 2436 conformations with an average silhouette width of 0.41±0.26, a maximum of 0.56 and a minimum of -0.11. The average backbone RMSDs, Table IIIB, show that the conformations within cluster 1 are related, Figure 4A, with the greatest degree of conformational stability in the N-terminal region and the greatest degree of conformational flexibility located in the central and C-terminal regions. The medoid conformation of cluster 1 is an a-helix from residues 2 through 7 and a bend from residues 6 through 10. This conformation is stabilized by the backbone hydrogen bonds, $Ser^{4}{}_{NH}\cdots pGlu1_{C=O}$, $His^{5}{}_{NH}\cdots pGlu^{1}{}_{C=O}$, $Asp^{6}{}_{NH}\cdots His^{2}{}_{C=O}$, $Trp^{7}{}_{NH}\cdots His^{5}{}_{C=O}$ and $Lys^{8}{}_{NH}\cdots His^{5}{}_{C=O}$; a side chain to backbone hydrogen bond, $pGlu^{1}{}_{NH}\cdots His^{5}{}_{NE2}$ and Ar-Ar interactions, $His^{2}\cdots Trp^{3}$ and $Trp^{3}\cdots Trp^{7}$. The percentages of sampled secondary structure, Table VB, show the presence of an a-helix from residues 2 through 5, a turn from residues 5 through 8 and a bend from residues 6 through 10. The stabilizing backbone hydrogen bonds and weakly polar interactions, Table VIIB, differ from those of all sampled conformations of lGnRH-III due to the absence of the $Asp^{6}{}_{NH}\cdots His^{2}{}_{C=O}$ hydrogen bond, an Ar-Ar interaction, $His^{2}\cdots Trp^{7}$ and the formation of the $Trp^{7}{}_{NH}\cdots Ser^{4}{}_{C=O}$ and $Lys^{8}{}_{NH}\cdots Ser^{4}{}_{C=O}$ hydrogen bonds.

Cluster 2 contains 2564 conformations with an average silhouette width of 0.55_0.13, a maximum of 0.69 and a minimum of 0.03. The average backbone RMSDs, Table IIIC, show that the conformations within cluster 2 are related, Figure 4B, with the greatest degree of conformational stability in the N-terminal region and increased conformational flexibility in the central and C-terminal regions. The medoid conformation of cluster 2 is an α-helix from residues 2 through 5, a turn from residues 5 through 7, a bend from residues 5 through 9 and a bend from residues 6 through 10. This conformation is stabilized by the backbone hydrogen bonds, $His^{5}{}_{NH}\cdots pGlu^{1}{}_{C=O}$, $Asp^{6}{}_{NH}\cdots His^{2}{}_{C=O}$, $Trp^{7}{}_{NH}\cdots His^{2}{}_{C=O}$; the side chain to backbone hydrogen bonds, $pGlu^{1}{}_{NH}\cdots His^{5}{}_{NE2}$, $Lys^{8}{}_{NH}{}^{3}\cdots Trp^{3}{}_{C=O}$, $Lys^{8}{}_{NH3}\cdots Ser^{4}{}_{C=O}$ and $Lys^{8}{}_{NH3}\cdots Gly^{10}{}_{C=O}$ and Ar-Ar interactions, $His^{2}\cdots Trp^{3}$, $His^{2}\cdots Trp^{7}$ and $Trp^{3}\cdots Trp^{7}$. The percentages of sampled secondary structure, Table VC, show the presence of a helix from residues 2 through 5, a turn from residues 5 though 7 and a bend from residues 6 through 10. The secondary structures are stabilized by backbone hydrogen bonds, Table VIIC, and weakly polar interactions, differing from cluster 1 by the absence of the $Lys^{8}{}_{NH}\cdots Ser^{4}{}_{C=O}$ hydrogen bond, an Ar-Ar interaction, $His^{2}\cdots Trp^{7}$ and formation of the $Asp^{6}{}_{NH}\cdots His^{2}{}_{C=O}$ and $Trp^{7}{}_{NH}\cdots His^{2}{}_{C=O}$ hydrogen bonds.

The average backbone RMSDs between the two clusters, Table IIID, are larger than the average backbone RMSDs within each cluster and confirm the distinct conformational differences between the two groups (18,19). The regional backbone

Table VII

Distance (A) between groups and atoms in weakly polar and hydrogen bond interactions, respectivley, for the 1GnRH-III trajectory from 200 ps to 5200 ps*.

Interaction	Avg.	Dev.	Max.	Min.
A. All conformations.				
$His^{2}_{imi}\cdots Trp^{3}_{ind}$	6.93	1.21	10.44	4.98
$His^{2}_{imi}\cdots Trp^{7}_{ind}$	7.88	2.94	14.57	4.14
$Trp^{3}_{ind}\cdots Trp^{7}_{ind}$	8.90	1.16	13.66	4.80
$Ser^{4}_{NH}\cdots pGlu^{1}_{C=O}$	3.13	0.24	4.00	2.57
$His^{5}_{NH}\cdots pGlu^{1}_{C=O}$	3.06	0.24	4.00	2.55
$Asp^{6}_{NH}\cdots His^{2}_{C=O}$	3.03	0.22	4.00	2.58
B. Cluster 1, 2436 conformations.				
$His^{2}_{imi}\cdots Trp^{3}_{ind}$	7.60	1.41	10.44	5.15
$Trp^{3}_{imi}\cdots Trp^{7}_{ind}$	8.95	1.58	12.04	5.10
$Ser^{4}_{NH}\cdots pGlu^{1}_{C=O}$	3.20	0.25	4.00	2.62
$His^{5}_{NH}\cdots pGlu^{1}_{C=O}$	3.09	0.25	4.00	2.55
$Trp^{7}_{NH}\cdots Ser^{4}_{C=O}$	3.13	0.27	3.99	2.64
$Lys^{8}_{NH}\cdots Ser^{4}_{C=O}$	3.02	0.25	3.99	2.61
C. Cluster 2, 2564 conformations.				
$His^{2}_{imi}\cdots Trp^{3}_{ind}$	6.29	0.37	7.98	4.98
$His^{2}_{imi}\cdots Trp^{7}_{ind}$	5.58	0.75	10.09	4.14
$Trp^{3}_{ind}\cdots Trp^{7}_{ind}$	8.90	0.92	12.41	6.38
$Ser^{4}_{NH}\cdots pGlu^{1}_{C=O}$	3.06	0.22	3.98	2.57
$His^{5}_{NH}\cdots pGlu^{1}_{C=O}$	3.10	0.23	4.00	2.63
$Asp^{6}_{NH}\cdots His^{2}_{C=O}$	3.00	0.18	3.92	2.58
$Trp^{7}_{NH}\cdots His^{2}_{C=O}$	3.04	0.22	3.99	2.59

*imi,imidazolyl-group;ind, indolyl-group; gua,guanidino-group

RMSDs, show that the greatest difference between the two groups is due to the conformation adopted by the central region, Figure 5B. The ϕ/ψ dihedrals of clusters 1 and 2, Figure 6B, show that residues 2 through 5 sample similar regions of phase space, with the differences between the two groups being due to the ϕ/ψ dihedrals of residues Asp^{6} and Trp^{7}. The c1 dihedral of Trp^{3}, Figure 7B, occupies only the *gauche(+)* region of phase space in all conformations. This is consistent with involvement of this residue in a stable interaction that restricts its side chain flexibility. Cluster 2 shows restricted sampling of the χ^{1} dihedrals of His^{2} in the *gauche(+)* and of His^{5} in *gauche(-)* regions of phase space.

The solvent contribution to the free energy difference between the medoid conformations of cluster 1 and cluster 2 is $\Delta A_{\text{cluster 1}\rightarrow\text{cluster 2}} = 72.6\pm1.7$ kcal/mol, which indicates that solute-solvent interactions play a greater role in stabilizing the conformations sampled in cluster 1 than those in cluster 2. Though the number of conformations sampled in each cluster is approximately equal, the greater solvent stabilization of cluster 1 could only be explained by a more favorable enthalpic conformational energy for cluster 2. The free energy calculation was used to study which medoid conformation is more stable in aqueous environment. During simulations, than no direct transition from medoid conformation 1 to 2 took place. This is consistent with the sign of the calculated free energy differences. Furthermore, the fact that no transition from medoid conformation 1 to 2 was observed indicates that the free energy difference must be substantial and this was revealed by the calculations. The net dipole moment of the medoid conformation in cluster 1 is 11.81 D compared with 7.78 D in cluster 2. The difference in the net dipole moments between the two conformations was expected because of the more compact medoid conformation of cluster 2 than the extended helical medoid conformation of cluster 1. The large difference in net dipole moments may also contribute to the large solvent influence on free energy difference between the two conformations.

Conclusion

Important conformational features are the differences in the sampled ϕ/ψ dihedral phase space of residues 5, 7 and 8 and the χ^{1} dihedral phase space of residues 2 and 3. GnRH forms a turn from residues 5 through 8 whereas lGnRH-III has an extended helical con-

formation from residues 2 through 7. The side chains of His^2 and Trp^3 in lGnRH-III participate in weakly polar interactions different from those in GnRH. The unique conformational properties of lGnRH-III may account for its different biological activity.

Acknowledgments

This work was supported by the US-Hungarian Science & Technology Joint Fund, No. 455, the Carpenter chair in Biochemistry, Creighton University and The Nebraska Medical Foundation Inc.

References and Footnotes

1. S. Lovas, I. Pályi, B. Vincze, J. Horváth, M. Kovács, I. Mezö,, G. Tóth, I. Teplán, R. F. Murphy, *J. Pept. Res. 52*, 384-389 (1998).
2. S.C. Sealfon, H. Weinstein, R.P. Millar, *R.P. Endocrine Rev. 18*, 180-205 (1997).
3. I. Mezö, S. Lovas, I. Pályi, B.Vincze, A.Kálnay, G.Turi, Z.Vadász, J.Seprödi, M.Idei, G.Tóth, E.Gulyás, F.Ötvös, M.Mák, J.Horváth, I.Teplán, R.F.Murphy, *J. Med. Chem. 40*, 3353-3358 (1997).
4. C.I. Bayly, P.Cieplak, W.D.Cornell, P.A.Kollman, *J. Phys. Chem. 97*, 10269-10280 (1993).
5. W.D.Cornell, P.Cieplak, C.I.Bayly, P.A.Kollman, *J. Am. Chem. Soc*. 115, 9620-9631 (1993).
6. W.D. Cornell, P. Cieplak, C.I. Bayly, I.R.Gould, K. Merz, Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox,,J.W. Caldwell, P.A. Kollman, *J. Am. Chem. Soc. 117*, 5179-5197 (1995).
7. P. Cieplak, W.D. Cornell, C.I., Bayly, P.A. Kollman, J. Comp. Chem. 16, 1357-1377 (1995).
8. Tripos Inc. Sybyl version 6.4 Users Manual, Tripos Inc., St. Louis (1996).
9. D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, W.S. Ross, C. Simmerling, T. Darden, K.M. Merz, R.V. Stanton, A. Cheng, J.J. Vincent, M. Crowley, D.M. Ferguson,R. Radmer, G.L. Seibel, U.C. Singh, P. Weiner, P.A. Kollman, AMBER version 5.0, University of California, San Francisco (1997).
10. D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham III, S. DeBolt,D.M. Ferguson, G. Seibel, P.A. Kollman, *Comp. Phys. Commun. 91*, 1-41 (1995).
11. M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S.J. Su, T.L. Windus, M. Dupuis, J.A. Montgomery. *J. Comp. Chem. 14*, 1347-1363 (1993).
12. W.C. Still, A. Tempczyk, R.C. Hawley, T. Hendrickson, *J. Am. Chem. Soc. 112*, 6127-6129 (1990).
13. H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, *J. Chem. Phys. 81*, 3684-3690 (1984).
14. M.P. Allen, D.J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, New York (1987).
15. Ponder, J.W. Tinker User's Manual, Washington University School of Medicine, St. Louis, http://dasher.wustl.edu (1997).
16. S. Lovas,R.F. Murphy, in *Molecular Modeling of Neuropeptides. 73*, 209-217 (1997).
17. W.L. Jorgensen, J. Chaudreskhar, J.D. Madura, R.W. Impey, M.L. Klein, *J. Chem. Phys. 79*, 926- 935 (1982) .
18. L. Kaufman, P.J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, John Wiley & Sons, Inc., New York (1990).
19. C.R.Watts, G. Tóth, R.F. Murphy, S. Lovas, *J. Mol. Struct. (THEOCHEM) 535*, 171-182 (2001).
20. W. Kabsch, C. Sander, *Biopolymers 22*, 2577-2637 (1983).
21. D. van der Spoel, A.R. van Bunren, E. Apol, P.J. Meulenhoff, D.P. Tielman, A.L.T.M. Sijbers, B. Hess, K.A. Feenstra, E. Lindahl, R. van Drunen, H.J.C. Berendsen, GROMACS version 2.0 UsersManual, Groningen, The Netherlands, http://md.chem.rug.nl/~gmx (1999).
22. C.A.Burley, G.A. Petsko, *Science 229*, 23-29 (1985).
23. M. Mezei, D.L. Beveridge, Ann. Acad. Sci. N.Y. 482, 1-23 (1986).
24. M. Mezei, *M.J. Comp. Chem. 13*, 651-656 (1992).
25. M. Mezei, F. Guarnieri, *J. Biomol. Struct. Dyn. 16*, 723-732 (1998).
26. M. Mezei, MMC, Molecular Monte Carlo User's Guide, Mount Sinai School of Medicine, New York, http://inka.mssm.edu/~mezei/mmc (2000)
27. M. Mezei, *J. Comp. Chem. 18*, 812-815 (1997).
28. M. Mezei, Simulaid, Simulation Setup Utilities, Mount Sinai School of Medicine, New York, http://inka.mssm.edu/~mezei/simulaid (2000).
29. S.S. Zimmerman, M.S. Pottle, G. Nemethy, H.A. Scheraga,. *Macromolecules 10*, 1-9 (1977).
30. P.I. Nagy, J.E. Bitar, D.E. & Smith, *J. Comp. Chem. 15*, 1228-1240 (1994).

Date Received: August 24, 2000

Communicated by the Editor Ramaswamy H Sarma