

# The Low-Energy Conformations of Gonadotropin-Releasing Hormone in Aqueous Solution

Matthew R. Pincus · Jannie Woo · Regina Monaco ·  
Jack Lubowsky · Matthew J. Avitable ·  
Robert P. Carty

Published online: 18 November 2014  
© Springer Science+Business Media New York 2014

**Abstract** Using the chain-build-up method based on Empirical Conformational Energies of Peptides Program including solvation, we have computed, the low energy conformations of gonadotropin-releasing hormone, GnRH, whose sequence is Pyro-Glu(PG)-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH<sub>2</sub>. We have found 5,077 solvated conformations with conformational energies that were within 5 kcal/mole of that of the global minimum. These minima were found to occur in 802 distinct conformational classes of which 25 represented 70 % of the Boltzmann energy-weighted structures. Virtually all of these structures adopted bend conformations from Tyr 5-Leu 8, and 3,861 structures adopted bend

conformations at residues 4–7. However, these structures differed significantly from one another, indicating that GnRH does not adopt a well-defined structure in aqueous solution consistent with the absence of a well-defined NMR structure of GnRH in water. A total of 300 of these structures were found to be superimposable on possible NMR structures for GnRH in DMSO with a combined statistical weight of 1.6 %. We found that Gly 6 adopts low energy “starred” states, e.g., C\* and D\*, that are energetically forbidden to L-amino acids but are low energy for D-amino acids, with a statistical weight of 43 %. This can explain why substitutions of L-amino acids for Gly 6 are known to inactivate GnRH while D-amino acid substitutions enhance its activity. Using these findings, in the accompanying manuscript, we compute the low energy conformations for the substituted GnRHs that enable inference of possible receptor-bound conformations.

M. R. Pincus (✉) · R. Monaco  
Department of Pathology, SUNY Downstate Medical Center,  
450 Clarkson Avenue, Brooklyn, NY 11203, USA  
e-mail: matthew.pincus2@va.gov

M. R. Pincus · R. Monaco  
Department of Pathology and Laboratory Medicine, New York  
Harbor VA Medical Center, 800 Poly Place, Brooklyn,  
NY 11209, USA

J. Woo  
Department of Pathology, Upstate Medical University, 750 East  
Adams Street, Syracuse, NY 13210, USA

J. Lubowsky · M. J. Avitable  
Scientific Academic Computing Center, SUNY Downstate  
Medical Center, 450 Clarkson Avenue, Brooklyn, NY 11203,  
USA

R. P. Carty (✉)  
Department of Biochemistry, SUNY Downstate Medical Center,  
450 Clarkson Avenue, Brooklyn, NY 11203, USA  
e-mail: robert.carty@downstate.edu

R. P. Carty  
Department of Anatomy and Cell Biology, SUNY Downstate  
Medical Center, Brooklyn, NY, USA

**Keywords** GnRH · Conformational energies · Low energy minima · Chain build-up procedure · Statistical weight · Superposition of structures · Unique conformations

## Abbreviations

GnRH Gonadotropin-releasing hormone  
ECEPP Empirical Conformational Energies of Peptides Program  
RMS Root mean square  
PGL Pyro-glutamic acid  
NOE Nuclear overhauser effect

## 1 Introduction

GnRH, gonadotropic-releasing hormone (gonadorelin or luteinizing hormone-releasing hormone, LHRH), whose

sequence is P-Gl(pyroglutamyl) 1-His 2-Trp 3-Ser 4-Tyr 5-Gly 6-Leu 7-Arg 8-Pro 9-Gly 10-NH<sub>2</sub>, regulates the production and release of luteinizing hormone, LH, and follicle-stimulating hormone, FSH, in the gonadotropes of the anterior pituitary. Optimal production and release of these sex hormones depend upon the “pulsatile” signaling of the pituitary gonadotropic receptors by GnRH [1]. Continuous delivery of GnRH to its target cells in the pituitary leads to a paradoxical suppression of LH and FSH production, a finding which has been utilized in the treatment of prostatic and breast carcinoma [1].

Over the past 40 years, over 2000 derivatives have been synthesized and analyzed for potential therapeutic efficacy. Some of these derivatives are potent, long-lasting agonists and antagonists of GnRH [2–4]. A variety of clinically important superagonists, that are used to produce a state of “chemical castration” under conditions where orchietomy or estrogen therapy are unacceptable, in treating androgen-sensitive prostate cancer, such as leuprolide, goserelin, nafarelin, histrelin and triptorelin, all possess substitutions for glycine at position-6 with D-amino acids containing bulky sidechains [5–7]. These substitutions are thought to retard peptide bond cleavage of the Tyr [5]-X [6] linkage by endopeptidase and may also increase the inherent biological activity of this peptide. Substitution of D-Val for Gly 6 results in diminished activity of the substituted hormone while substitution of D-Ala, D-Leu, D-Phe and D-Trp results in multifold increases in the activity of these substituted GnRHs [5–7]. In contrast, substitution of L-Ala or L-Val for Gly 6 in GnRH results in loss of most or all activity [5–7]. Thus substitution of many D-amino acids for Gly 6, while inhibiting endopeptidase degradation of GnRH, also promote increased biological activity of this hormone.

A variety of experimental conformational studies have indicated that GnRH exhibits no unique three-dimensional structure in solution [8, 9]. The latter authors found no short- or long-range interactions in GnRH in aqueous solution or in DMSO solution and when the molecule was bound to a lipid membrane. Two-dimensional NMR spectroscopy suggests that these conformers must be rapidly interconverting on a time scale undetectable by the usual spectroscopic methods. In addition, GnRH has resisted crystallization, and no X-ray structure is available.

More recently, 2D-NMR structures have been obtained for native GnRH [10] and for leuprolide, a GnRH peptide containing a D-Leu residue in place of Gly 6 and that is truncated at Pro 9 that terminates with an N-ethyl group attached to the C=O of this Pro residue [11]. The overall structure for native GnRH consists of two semi-extended domains separated by a reverse turn at Gly 6-Leu 7. One of these structures was fit into a model of the GnRH receptor, that is in the rhodopsin family, that was computed from the

nine transmembrane domains of bovine rhodopsin [12]. Specific contacts made between the NMR structure and critical residues of the receptor could be modeled. However, there is no X-ray or NMR-based structure of GnRH bound to its receptor. It is also not clear that the inferred structures of GnRH in DMSO are the same as the structure of this peptide when bound to the receptor. The model furthermore does not explain how substitution of D-amino acids for Gly 6 enhance binding of GnRH and its activity as noted above.

Theoretical studies have been carried out in attempts to predict the structures of the low-energy conformers of GnRH present in various environments. Initial studies by Momany [13–15] suggested that low-energy conformers of GnRH contained a modified type II bend at Tyr5-Gly6-Leu7-Arg8. One conformer, the CC type, also positioned the N- and C-terminal elements in close proximity. While these computations were restricted to a limited number of starting conformations and the resulting conformers were not solvated, Momany’s model correctly predicted the enhanced biological activity of the [D-Ala6, N-Me-Leu7] GnRH analog [6, 16].

Freidinger et al. [17], constrained the GnRH decapeptide by synthetically introducing a  $\gamma$ -lactam ring connecting the Gly 6 C- $\alpha$  atom and the Leu 7N atom which stabilized the  $\beta$ -turn conformation at residues 5–8. The analog was 2.5 times as active as the parent GnRH. Other studies [18, 19] predicted conformational preferences for the terminal blocked peptides, P-G1-His 2-Trp 3-Ser 4-Tyr 5-NMe and Ac-Gly 6-Leu 7-Arg 8-Pro 9-Gly 10-NH<sub>2</sub>, as well as the central analog fragment, Ac-Ala 4-Tyr 5-Gly 6-Leu 7-Arg 8-pyrrolidyl 9. More recent computations [20] constructed GnRH low energy conformers starting with 5000 Ser 4-Tyr5-Gly 6-Leu 7-Arg 8 structures. Systematic addition of residues to the C-terminal end followed by the addition of residues at the N-terminus gave 300 low-energy conformers. The authors found only eight different low-energy conformers for the central tetrapeptide fragment, Tyr 5-Gly 6-Leu 7-Arg 8. The lowest-energy conformer was found to contain the unusual A\* state [21] for the peptide dihedral angles at Leu 7. This study corroborated previous work indicating that GnRH exhibits a high degree of conformational flexibility and that the central fragment is populated by a small ensemble of reverse turns.

Other workers [22] have attempted to model GnRH conformers based on the identification of conformational patterns in sequences found in protein structure data which are identical to the sequences found in the decapeptide hormone. These constraints, coupled with data on the structure–activity relationships of a variety of synthetic analogs, were used to find a unique conformation representative of GnRH. The putative structure contains a  $\beta$ -turn and a contiguous surface formed by the juxtaposition of the

sidechains of Arg 8, Trp 3 with the C-terminal carboxamide group of Gly 10 and the N-terminal P-G1 group. In 1976, Struthers et al. [23], performed a molecular dynamics simulation to find the minimum energy conformations of GnRH and a cyclic analog. The final structure of the native decapeptide contained a bend at the Gly 6-Leu 7 position, and H-bonds between the Trp 3 HN H atom and the Gly 10 O atom, the Trp 3 O atom and Gly 10 HN H atom, and the His 2 ND1 H atom and Gly10 O atom. Guarnieri and Weinstein [24] using an initial exploratory Monte Carlo-simulated annealing study followed by a biased sampling technique to extract low-energy conformers, found a small family of structures with a  $\beta$ -turn structure at residues 5–8. They also proposed that the analog, Lys 8-GnRH is mostly populated with conformers of an extended backbone structure. In 1998, Mezei and Guarnieri [25] extended their studies by including an explicit solvent model. One aspect of these prior studies that was not addressed is the effects of substitution of D- and L-amino acids for Gly 6 on the affinities of these substituted GnRHs for the receptor and on their relative activities.

In this paper, we extend the earlier ECEPP computational analyses [13–15] and determine the effects of solvation on the conformational preferences of GnRH using the Hodes–Nemethy–Scheraga model of solvation [26]. As noted above, with the exception of D-Val, substitution of D-amino acids for Gly 6 result in enhanced activities of the resulting GnRH peptides while substitution of L-amino acids results in dramatic loss of activities of the resulting peptides. In the accompanying paper, using these findings, we compute the low energy conformations of the Gly6-substituted peptides to infer possible structures of GnRH that may be candidates for the active binding conformations of GnRH and the D-amino acid-substituted peptides.

## 2 Methods

### 2.1 Conformational Energies and Peptide Sequence

All conformational energies for individual conformational states were computed with the program, ECEPP [27]. The GnRH sequence for which we computed the low energy structures was PG-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH<sub>2</sub> where PG is the pyroglutamyl residue. Although considered an end group in ECEPP, the pyroglutamyl residue was considered as residue 1. Conformational sampling was achieved using the chain build-up procedure from the carboxyl to amino terminal end of the GnRH sequence [28]. All sidechains were taken in the uncharged state. During the chain-buildup procedure, the Ser side chain -OH H atom was considered to be a type 1 aliphatic hydrogen. Following completion of the chain-buildup

procedure, the Ser HG sidechain H atom was changed to a type 4 hydrogen, and the low-energy conformations re-minimized. A total of 47 dihedral angles were considered to be variable. The intra-ring dihedral angle,  $\phi$ , for proline, and that for the pyroglutamyl group (defined as C $\delta$ PG-N $\alpha$ PG-C $\alpha$ PG-C'PG) were held constant. Fixed values of  $-75.000^\circ$  and  $140.057^\circ$ , respectively, were used. The Pro 9 residue was generated only in the trans conformation and with the pyrrolidine ring in the “down-B” conformation as defined in in ECEPP [27, 29].

### 2.2 Chain Build-Up Procedure

For all calculations, we initiated the calculations by generating all combinations of conformations for the terminally blocked CH<sub>3</sub>CO-Pro 9-Gly 10-NH<sub>2</sub> sequence in which the single residue minima for trans-Pro were combined with those for Gly. Each of these conformations was subjected to energy minimization using the Powell algorithm [30]. All of the resulting minima were sorted such that all structures whose conformational energies that lay within 100 kcal/mole of the energy of the global minimum were retained.

### 2.3 Solvation

We then evaluated the effects of solvation on the conformational energies of these minima using the Hodes–Nemethy–Scheraga hydration program [26]. The total conformational energy ( $E_{\text{total}}$ ) of each minimum was then computed as

$$E_{\text{total}} = E_{\text{ECEPP}} + E_{\text{hydr}}, \quad (1)$$

where  $E_{\text{ECEPP}}$  is the “dry” or unsolvated conformational energy computed using ECEPP, and  $E_{\text{hydr}}$  is the free energy of hydration computed by the Hodes–Nemethy–Scheraga method. All of the resulting conformations whose  $E_{\text{total}}$  values occurred within 5 kcal/mole of that of the global minimum were retained (except at certain steps as described below).

### 2.4 N-Acetyl-Arg-Pro-Gly-NH<sub>2</sub>

All combinations of conformations for these minima with the single residue minima for Arg were then generated and subjected to energy minimization, followed by hydration and selection of minima whose  $E_{\text{total}}$  values occurred within the cutoff energy of 5 kcal/mole.

### 2.5 Non-Degenerate Minima [28]

Following conformer selection based on energy, degenerate minima were removed by conformational sorting. In this routine, minima with identical or near-identical values of the

variable dihedral angles were identified and only those minima with at least one dihedral angle different by  $\geq 20^\circ$  were considered unique and were retained for the next step. If two minima possessed all dihedral angles within  $20^\circ$ , only the one with the lowest energy was retained.

## 2.6 The Low Energy Conformations for the Full Chain

This process was repeated iteratively proceeding toward the amino terminus through Trp 3. The resulting minima were then combined with the minima for PG-His. For this sequence, PG was generated in three different conformations in which the dihedral angle,  $\psi$ , was  $-30^\circ$ ,  $-150^\circ$ , or  $+90^\circ$ . These conformers were combined with the single residue minima for His, and all combinations of the resulting conformations were combined with the low energy conformations computed for Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH<sub>2</sub>, and the resulting conformations were subjected to energy minimization, and the subsequent steps described above were then followed. All of the resulting energy minima were sorted by total conformational energy (Eq. 1) from lowest to highest energy, and, for each minimum, the conformation of each residue was assigned a conformational letter state as defined previously [21].

## 2.7 Statistical Weights for Conformations

We computed the Boltzmann probabilities of occurrence for all of the resulting minima as follows. The partition function,  $Z$ , for the ensemble of low energy conformations for this decapeptide was computed as

$$Z = \sum_{i=1}^N e^{-E_i/RT} \quad (2)$$

where  $E_i$  is the total conformational energy of the  $i$ th conformation computed using Eq. 1,  $R$  is the gas constant, and  $T$  is the temperature in  $^\circ\text{K}$ . The statistical weight of a particular low energy conformation was then computed as,

$$P_j = \left[ \sum_{j=1}^M e^{-E_j/RT} \right] / Z \quad (3)$$

where  $j$  is the  $j$ th particular conformation and  $M$  is the total number of occurrences of the particular conformation. In addition, we computed single residue probabilities for each residue in a given conformational state using Eq. 3. In this case,  $E_j$  in Eq. 3 was the conformational energy of each conformation in which the residue was in the given letter state [20].

## 2.8 Computations

All ECEPP, hydration, and sorting computations were carried out on IBM 9672 Model R24, IBM 9672 Model

R14 and Hewlett-Packard Alpha computers. Coordinates for the structures for the minimum-energy conformers of each of the most probable conformations were generated and models were constructed using either the Molecular Design Inc. (MDI) molecular model system (Memphis, TN) or the Silicon Graphics Indigo system using the program INSIGHT (ACELRYS, San Diego, CA). The Silicon Graphics Iris Indigo Elan 4000 computer running Sybyl, version 6.6 program, was employed to study the superposition of various peptide backbone conformational states.

## 3 Results

### 3.1 Low Energy Minima from the Chain Build-Up Protocol

Table 1 summarizes the computational strategy and the numbers of minima obtained at each stage in the calculation. Based on the number of variable dihedral angles in each residue, the total number of conformations that can be generated is  $1.79 \times 10^{12}$ . As can be seen in Table 1, the chain build-up procedure results in 5,077 low energy structures (0.00000028 % of the total) of which, based on the letter state sequences, there are 802 distinct structures. Of these, 18 possessed an energy-weighted frequency (see Eq. 2) greater than 1 %. The first 25 unique backbone states are shown in Table 2. The percentage of dry minima retained within the 100 kcal/mole cutoff averaged between 90 and 100 % at each step, except for step 7 where 80 % were retained within the cutoff. Conformational sorting eliminated between 2.2 and 31.8 % of the hydrated minima at or within the 5 kcal/mole cutoff.

### 3.2 There are No Unique Low Energy Structures for Native GnRH

As can be seen in Table 2, there are multiple, different low energy minima for GnRH. Eight of the conformers in Table 2 have energies that differ by  $< 1$  kcal/mole compared with that of the global minimum. The most probable structure, conformer 1 in Table 2, and shown in Fig. 1, has only a 14 % probability of occurrence. This conformer has a total conformational energy (Eq. 1) that lies close (0.04 kcal/mole) to that of the global minimum, conformer 5 whose probability of occurrence is about 5 %, or about one-third of that of conformer 1. Conformer 3 has a probability of occurrence of 6.7 % and differs in conformation from structure 1 at Gly 6 (D\* in conformer 3 vs C\* in conformer 1) and C-terminal Gly 10 (C\* in conformer 3 and A in conformer 1). Nonetheless, the backbone and most sidechain coordinates of these two structures superimpose, except for Gly 10 where the two chains diverge as

**Table 1** Chain build-up protocol for the construction of the low-energy conformations of GnRH

Peptide construct	Minima combined			Sorted minima dry 100 kcal/mole cutoff (number)	Sorted minima hydrated 5 kcal/mole cutoff (number)	Sorted minima hydrated nondegenerate (number)
	Single residue (number)	Peptide (number)	Combinations (number)			
NAcPGNH <sub>2</sub>	3	7	21	21	21	
NAcRPGNH <sub>2</sub>	210	21	4,410	4,407	–	3,879
NAcLRPGNH <sub>2</sub>	27	3,879	104,733	104,674	4,325	4,230
NAcGLRPGNH <sub>2</sub>	7	4,230	29,610	29,472	15,971	15,389
NAcYGLRPGNH <sub>2</sub>	30	15,389	461,670	404,226	13,429 <sup>c</sup>	11,661
NAcSYGLRPGNH <sub>2</sub>	44	11,661	513,084	495,908	34,302 <sup>d</sup>	26,330
NAcWSYGLRPGNH <sub>2</sub>	34	26,330	895,220	713,703	22,560	20,184
PG1HWSYGLRPGNH <sub>2</sub> <sup>a</sup>	48	20,184	968,832	2,678,722 <sup>b</sup>	12,557	10,915
PG1HWSer4YGLRPGNH <sub>2</sub>			12,557	12,557	7,447	5,077

<sup>a</sup> Only neutral histidine single-residue minima containing the hydrogen attached to the N<sup>δ</sup> atom of the imidazole ring were considered

<sup>b</sup> This number of minima represents the sum of three separate combinatorial calculations with different values of  $\psi$  for the pyroglutamate residue

<sup>c</sup> Hydration was carried out using a cutoff of 4.5 kcal/mole

<sup>d</sup> Hydration was carried out using a cutoff of 4.0 kcal/mole

can be seen by the divergence of the amide groups of Gly labeled as “11” on the lower right of Fig. 1. If these two structures were considered to be in the same class, the overall probability of occurrence of both structures would be approximately 20 %. Overall, however, the multiplicity of different low energy conformers for this peptide (about 80 % of the low energy structures) corroborates the finding that no unique structure has been found for this peptide in aqueous solution.

As can be seen in Fig. 1, in both conformers 1 and 3 in Table 1, the side chain aromatic rings of Trp 3, Tyr 5 and Pro 9 lie on one face of the structure. This arrangement of the side chains creates a hydrophobic surface. If these structures bind to the GnRH receptor, this side chain arrangement may contribute to the binding if there is a complimentary hydrophobic surface on the receptor.

From Table 2, the first 25 different letter-state conformations constitute about 70 % of the energy-weighted conformers (Eq. 2) but only 31 % of the minima (1,584/5,077). These conformers encompass both structures whose energies lie close to that of the global minimum (conformer 5 in Table 2) and those with energies that lie considerably above that of the global minimum like conformer 23 whose energy is over 3 kcal/mole above that of the global minimum but which has numerous low energy conformations with different side chain conformations.

### 3.3 Common Features of Low Energy Structures

Generation of the lowest energy structures in Table 2 showed that, while the structures differed significantly

from one another, all of these structures possessed reverse turns (bends) at residues 5–8 where a reverse turn is defined as the  $C_i^\alpha - C_{i+3}^\alpha$  distance  $\leq 7$  Å. Analysis of  $C_5^\alpha - C_8^\alpha$  distances in all 5,077 minima revealed that all except 9 of these low energy minima possess a reverse turn at residues 5–8, and 76 % of the conformers possess bends at residues 4–7. Of the 31 backbone conformer types observed at residues 5–8, 29 are bends. Only backbone states, EE\*CA\* and EC\*EA\*, possess  $C_{\alpha i} - C_{\alpha i+3}$  distances greater than 7 Å for residues 5–8. As discussed in the next section below, Leu 7 has close to a 0 probability of adopting the C conformation, making the occurrence of the first of these two conformations very unlikely.

The three most frequently observed bends are AAAD (51.5 % of the minima), EA AE (8.5 % of the minima), and EC\*AE (6.0 % of the minima) corresponding to 286/802 different AAAD peptide backbone states, 61/802 EA AE states, followed by 60/802 EC\*AE states. Some of the more heavily weighted backbone letter states, states 1, 3, 4, 6 and 8, depicted in Table 2, are modestly populated with structures with different sidechain conformations, while others, states 2, 7, 11, 16 and 20 are poorly populated with structures having different sidechain conformations. States 9, 13, 17 and 23 which have an AAAADA conformation at residues 3–9 are highly populated with structures with different sidechain conformations.

Examination of the lowest energy backbone states in Table 2 reveals similarities in structure between various letter-state ensembles. For example, states 2, 5, and 7 all contain the AEAAADAC structure for residues 3–10, and states 4, 6 and 12 are identical in terms of their letter state

**Table 2** Conformational states and statistical weights for the low-energy backbone conformations of GnRH<sup>a</sup>

	P-G	His	Trp	Ser	Tyr	Gly	Leu	Arg	Pro	Gly	Statistical weight <sup>b</sup>	Frequency <sup>c</sup>	Energy difference (kcal/mole) <sup>d</sup>
1	E*	A	A	A	E	C*	A	E	C	A	0.14254	52	+0.04
2	E*	C	A	E	A	A	A	D	A	C	0.07008	6	+0.16
3	E*	A	A	A	E	D*	A	E	C	C*	0.06732	36	+0.82
4	E*	F	C	A	E	A	A	E	F	C	0.05523	67	+0.28
5	D*	A*	A	E	A	A	A	D	A	C	0.04899	11	+0.00
6	E*	F	C	A	E	A	A	E	F	C*	0.04028	84	+0.75
7	G*	C	A	E	A	A	A	D	A	C	0.03921	6	+0.62
8	E*	A	A	A	D	D*	A	F	C	A*	0.03533	80	+1.23
9	E*	A	A	A	A	A	A	D	A	E	0.02159	196	+1.83
10	E*	D	A	E	A	A	A	D	A	E	0.02062	34	+1.09
11	E*	C	A	E	A	A	A	D	A	G	0.02019	5	+0.54
12	E*	F	C	A	E	A	A	E	C	C	0.01996	25	+1.15
13	D*	E	A	A	A	A	A	D	A	E	0.01347	236	+1.68
14	E*	A	A	A	A	A	A	D	A	C	0.01293	156	+2.47
15	E*	A	A	A	A	A	A	D	A	D*	0.01133	97	+2.40
16	G*	E	C	G	E	C*	A	E	C	C	0.01125	6	+0.95
17	E*	A	A	A	A	A	A	D	A	G	0.01089	168	+2.81
18	G*	E	C	G	E	C*	A	E	F	E*	0.01037	14	+1.65
19	E*	C	A	E	A	A	A	D	A	E	0.00903	14	+1.20
20	E*	A	C	E	F	A*	G	D	C	G*	0.00776	4	+1.13
21	E*	D	A	E	A	A	A	D	A	D*	0.00773	21	+2.01
22	E*	D	A	E	A	A	A	D	C	C*	0.00720	20	+1.96
23	D*	E	A	A	A	A	A	D	A	C	0.00711	175	+3.32
24	D*	A	C	G	E	C*	A	E	F	E*	0.00694	14	+1.87
25	E*	A	E	E	F	C*	A	A*	C	A*	0.00675	57	+2.53

<sup>a</sup> Conformational states are defined as follows: The familiar states are A ( $\alpha$ -helix) and E (extended). States E, E\*, F, F\*, D, and D\* are collectively called  $\beta$  conformations. The range of dihedral angles for all letter conformational states follows: A,  $-110^\circ \leq \phi < -40^\circ$ ,  $-90^\circ \leq \psi < -10^\circ$ ; C,  $-110^\circ \leq \phi < 40^\circ$ ,  $50^\circ \leq \psi < 130^\circ$ ; D,  $-180^\circ \leq \phi < -110^\circ$ ,  $20^\circ \leq \psi < 110^\circ$ ; E,  $-180^\circ \leq \phi < -110^\circ$ ,  $-180^\circ \leq \psi < -140^\circ$  or  $110^\circ \leq \psi < 180^\circ$ ; F,  $-110^\circ \leq \phi < -40^\circ$ ,  $-180^\circ \leq \psi < -140^\circ$  or  $130^\circ \leq \psi < 180^\circ$ ; G,  $-180^\circ \leq \phi < -110^\circ$ ,  $-90^\circ \leq \psi < -40^\circ$ ; H,  $-180^\circ \leq \phi < 0^\circ$ ,  $-90^\circ \leq \psi < -140^\circ$  or  $180^\circ \leq \psi < -180^\circ$ . States indicated by an asterisk are obtained by multiplying the corresponding single-letter value by -1 and reversing the inequalities

<sup>b</sup> The statistical weight is calculated from the Boltzmann distribution function,  $v_i = \exp(-\Delta E_i/RT)/\sum \exp(-\Delta E_i/RT)$

<sup>c</sup> The frequency is the number of times the backbone conformational type appears in the list of 5077 hydrated minima with energies at or within 5 kcal/mole of the global minimum and represents the number of different sidechain conformations within the group

<sup>d</sup> Energies are expressed as the difference between the conformational energy of the lowest-energy conformer in the group and the global minimum

designation for residues 1–8, E\*FCAEAAE. However, Pro 9 and Gly-NH<sub>2</sub> 10 do not superimpose. These two residues may be critical in interacting with the GnRH receptor [5–7, 12].

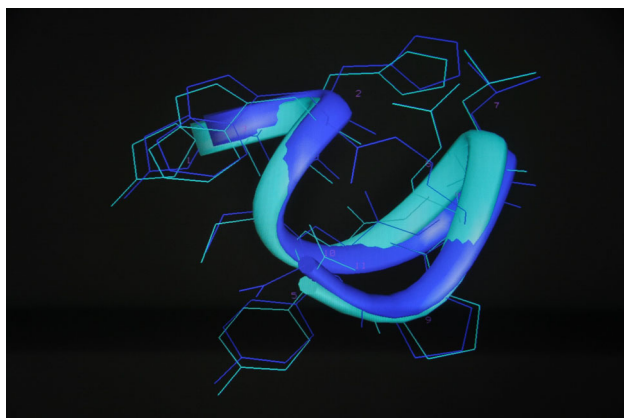
### 3.4 Single Residue Probabilities

Table 3 shows the single residue probabilities for the GnRH sequence. As can be seen from this table, the pyroglutamyl residue adopts the E\* conformation (70 %) although the D\* state has a significant probability of 20 %. His 2 can adopt either the alpha (A) state or beta states

(C,D,E,F) with almost equal probability (41 and 42 %, respectively). Trp 3 is computed to have a high probability for adopting the A state (70 %) while Ser 4 and Tyr 5 have approximately the same probabilities for alpha and beta states. Gly 6 is seen to adopt the A state with the highest probability (55 %) but also adopts “starred” states (A\*, C\* and D\*) with a combined probability of 43 %. The C\* and D\* states are high energy states for L-amino acids but low energy states for D-amino acids,

As discussed in the accompanying paper [31], substitution of a number of D-amino acids for Gly 6 in GnRH results in significantly increased activity of GnRH while





**Fig. 1** Ribbon structures for the two structures of highest statistical weight for GnRH (conformers 1 and 3 in Table 2) as computed by ECEPP and the chain build-up procedure including the effects of solvation. The blue structure is for the higher probability structure (E\*A A A E D\*A E C C\*) while the cyan structure is for the lower probability structure (E\*A A A E C\*A E C A) which are seen to be superimposable (Color figure online)

substitution of L-amino acids for Gly 6 results in low or no GnRH activity [6, 7]. While D-amino acid residues can adopt the A state, L-amino acids are more likely to adopt the A state [21], making it unlikely that Gly 6 would adopt the A state in receptor-bound GnRH. Since the starred states are low energy ones for D-amino acids, these findings would suggest that conformations of GnRH with Gly 6 in a starred state, such as conformers 1, 3, 8, 16 in Table 2, would be the ones most likely to bind to its receptor even

though Gly 6 has a higher probability of adopting the A state (Table 3).

Leu 7 is seen in Table 3 to have a high propensity (95 %) to adopt the A (alpha) state while Arg 8 does not adopt the A state but rather adopts beta conformations with high probability. The last two residues, Pro and Gly, appear to have several conformations that can be alpha or beta. Since the single residue probability for Leu 7 is 95 % for the A state and since there is a high probability for reverse turn formation to occur at Gly 6-Leu 7, our calculations suggest that almost all bends that form at these two residues must have Leu 7 in the A state. In prior calculations of the low energy conformations of GnRH, without solvation, these two residues were both computed to adopt the C conformation [13–15]. As can be seen in Table 3, based on our calculations including solvation, we find neither residue adopts the C conformation (Table 3).

### 3.5 Superposition of the 5,077 Low Energy GnRH Structures on Structures for Native GnRH Determined by NMR in DMSO

Structures of GnRH have been determined using 2D NMR for GnRH in dimethylsulfoxide (DMSO). Twenty one structures were found to fit the NOE constraints, and the coordinates for these structures have been deposited in the protein data base as Structure Set 1YY1 [10, 11]. There is no structure, however, for GnRH bound to its receptor.

A model for the structure of GnRH bound to its receptor has been proposed using the 1YY1 NMR coordinates [12]. The receptor structure was modeled using homology and

**Table 3** Computed single residue probabilities for GnRH

Conf. state <sup>a</sup>	Residues									
	1 PG <sup>b</sup>	2 HIS	3 TRP	4 SER	5 TYR	6 GLY	7 LEU	8 ARG	9 PRO	10 GLY
A	0.00	0.41	0.70	0.55	0.41	0.55	0.95	0.00	0.35	0.16
B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C	0.00	0.17	0.23	0.02	0.00	0.00	0.00	0.00	0.45	0.30
D	0.00	0.09	0.01	0.01	0.07	0.00	0.01	0.46	0.00	0.03
E	0.00	0.10	0.01	0.34	0.46	0.00	0.00	0.44	0.00	0.09
F	0.00	0.16	0.01	0.01	0.06	0.00	0.00	0.05	0.20	0.00
G	0.00	0.01	0.01	0.06	0.00	0.02	0.04	0.00	0.00	0.04
H	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A*	0.00	0.07	0.02	0.00	0.00	0.04	0.00	0.05	0.00	0.09
B*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C*	0.00	0.00	0.00	0.00	0.00	0.27	0.00	0.00	0.00	0.15
D*	0.20	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.05
E*	0.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05
F*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
G*	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
H*	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

<sup>a</sup> Conformational states as defined in ref. 20

<sup>b</sup> PG = L-Pyroglyutamyl residue

molecular dynamics from the known structural domains of the photoactivated, deprotonated intermediate state of bovine rhodopsin to whose class of structures the GnRH receptor belongs. The modeling of the binding conformation of GnRH was based on the results of prior studies on site-specific mutagenesis of the GnRH receptor to determine which residues were involved in ligand binding. It was found that the NMR structure could fit the predicted receptor structure and explain the results of the prior site-specific mutagenesis studies. It was found, for example, that the ring C=O group of PG 1 can hydrogen bond to the side chain NH of Asn 212 of the receptor, His 2 of GnRH with Lys 121 and Asp 98, Arg 8 of GnRh with Asp 302 and Pro 9-Gly 10 of GnRH with Trp 101 and Asp 102. In the latter case, it is not clear why the GnRH derivative lacking the C-terminal Gly-NH<sub>2</sub> but having an *N*-ethyl substituent attached to the carboxyl group of Pro 9 binds to the receptor. As far as we are aware, there has been no experimental confirmation of the proposed receptor structure.

We superimposed the coordinates for the backbone atoms for residues 2–10 and all of the heavy atoms for residue 1, the pyro-Glu residue, for each of our 5,077 computed lowest energy structures on the corresponding coordinates for each of the twenty one structures that satisfied the NOEs for the NMR structure of GnRH in DMSO. Structures were considered superimposable if the RMS deviations were  $\leq 3$  Å.

Of the total number of 5,077 computed low energy structures, we found that 300 of these were superimposable on at least one of the 21 NMR-determined structures. The lowest energy among these was structure 440, E\*A C A E A A E F C\*, that was superimposable on multiple PDB structures, whose conformational energy was 2.5 kcal/mole above that of the global minimum. The low energy structure, number 3,603 (G\*G E C A D\*A E C E) in the list of low energy structures, that gave the lowest RMS deviation of 2.38 Å when superimposed on the eighth NMR structure listed in the PDB, was computed to have a conformational energy of almost 5 kcal/mole higher than that for the global minimum. The C<sup>α</sup> trace for this computed structure (purple) is shown in Fig. 2 superimposed on that for NMR structure 8 (red).

All of these structures contain reverse turns at Gly 6-Leu 7 as can be seen in Fig. 2 for structure 3,603 and the 8th NMR structure. The fact that this bend occurs in the NMR structure supports the findings of prior studies and of our present study that residues 5–8 form a bend in GnRH.

The NMR structures, as represented by structure 8 in Fig. 2, have residues on both sides of the reverse turn at Gly 6-Leu 7 that are in the extended or in the  $\beta$ -region of conformational space and have overall features of a  $\beta$ -pleated sheet. This type of structure is of low probability in



**Fig. 2** C<sup>α</sup> trace of the energy minimum structure (3,603), shown in purple, of the 5,077 lowest energy structures computed for GnRH, of lowest rms deviation (2.38Å) when superimposed on one of the 21 NMR-determined structures (number 8 in the 1YY1 PDB file, ref.10 for GnRH in DMSO), shown in red (Color figure online)

our calculated low energy structure list. The overall probability that our computed 300 structures would superimpose on at least one NMR structure was 1.6 %. This low probability may reflect differences in solvent: the NMR structures were determined in DMSO; our computed structures are for GnRH in water, in which solvent no NMR structure has been determined, likely due to the existence of multiple, different conformations as we found in our calculations. Thus if the receptor “environment” is accurately simulated by DMSO solvent, we surmise that this low percentage of GnRH molecules that adopt the superimposable pleated sheet conformations in water would make receptor binding a low probability event.

#### 4 Discussion

We have performed conformational analysis on GnRH using the chain build-up method with high energy cut-offs for the unsolvated states and 5 kcal/mole for most stages of the chain build-up except for two stages in which we used 4.5 and 4 kcal/mole after hydration, respectively. While it is possible that these latter cutoffs may have resulted in the exclusion of “viable” structures, this process resulted in our generating over 2 million structures in the last step of the build-up procedure, suggesting that this process resulted in an adequate sampling of the conformation space for this peptide. In the build-up procedure, we used the method of combination of non-degenerate minima, i.e., if two conformations have corresponding dihedral angles that differ by  $<20^\circ$ , they are considered to be identical, and the one with the higher energy is excluded. We applied this method to all of the dihedral angles, i.e., the three backbone and all of the side chain angles. Therefore we



excluded only essentially identical structures of higher energy. Thus we generated a true representative sampling of low energy conformations that could be used in our computations of the probabilities of the low energy conformations.

Analysis of the resulting low energy minimum revealed that there was no distinct lowest energy structure or a structure with high probability of occurrence with solvation included. This result is compatible with the results of prior NMR studies of GnRH in aqueous solvents that showed no distinct structures. However, our calculations suggest that the most probable structure, i.e., conformer 1 in Table 2, E\*A A A E C\*A E C A, which has a probability of about 14 %, is similar to another structure of slightly higher conformational energy, giving a combined probability of occurrence of 20 %. This is still a low probability structure that would probably not be detected in multidimensional NMR experiments. In addition, all, except a few, high energy structures, form bends at Gly 6-Leu 7; these bends have different conformations so that again it may be impossible to detect a bend conformation in solution.

The structure of GnRH has been determined by NMR in DMSO that has a high dielectric constant (47) that is not so high as that for water (80) and, unlike water, is aprotic. The 21 structures that satisfied the NOE constraints were semi-extended structures on either side of bends at Gly 6-Leu 7 supporting the occurrence of a bend at these two residues. Although 300 of our 5,077 computed low energy structures superimposed on at least one of these 21 structures, the overall statistical weight for these structures was 1.6 %, that would be undetectable in an NMR experiment. Interestingly, even for these superimposable structures, there were several different conformations for the two bend residues, e.g., A A and D\*A.

It is important to bear in mind that the only other prior published NMR study of GnRH in DMSO and water [9] concluded that there were no NOEs beyond intrasidue interactions and that GnRH has an open structure, i.e., an ensemble of multiple structures in these solvents. The more recent 2D-NMR structures for GnRH were determined in DMSO but not in water, the solvent that was modeled in our calculations. The finding that there were no NOEs in water is consistent with the results of our calculations, i.e., that there is no one preferred structure, i.e., one of high probability of occurrence. Thus, in water, the one or more structures that bind(s) to the GnRH receptor must be one(s) of low probability.

Based on the results we have obtained, it is impossible to infer the structure(s) of GnRH that bind(s) to the GnRH receptor. However, GnRH peptides containing D- and L-amino acid substitutions for Gly 6 have been synthesized and tested. It has been found that almost all of the D-amino acid-substituted peptides have higher GnRH-induced

hormone release activity and/or enhanced receptor affinities while the L-amino acid substituted peptides have low or no activities and receptor affinities [6, 7]. We therefore have extended our computations to include these substituted GnRH peptides to enable computing low energy structures that would be strong candidates for receptor binding as we discuss in the succeeding paper [31].

**Acknowledgments** We wish to thank the Scientific Academic Computing Center, particularly Dr. Matthew Avitable and the University Hospital Computing Center of SUNY Downstate Medical Center for their excellent support for this project. We also acknowledge the support of the United States Department of Veteran Affairs and of the staff of the research service of the VA New York Harbor Health Care System for this project.

## References

- Klingmüller D, Schweikert HU (1992) Gonadotropin-releasing hormone: physiological aspects in mechanisms of action of LH-RH agonists. In: Höffken K, Kath R (eds) Recent results in cancer research 124. Peptides in oncology 1. LH-RH agonists and antagonists. Springer, New York, pp 1–6
- Karten MJ, Rivier JE (1986) Gonadotropin-releasing hormone analog design. structure-function studies toward the development of agonists and antagonists: rationale and perspective. *Endocr Rev* 7:44–66
- Henzl MR (1992) Gonadotropin-releasing analogs: update on new findings. *Am J Obstet Gynecol* 166:757–761
- Conn PM, Crowley WF Jr (1994) Gonadotropin-releasing hormone and its analogs. *Ann Rev Med* 45:391–405
- Bentley PJ (1980) Endocrine pharmacology structural basis and therapeutic applications. Press Syndicate of the University of Cambridge, London, p 63
- Monahan MW, Amoss MS, Anderson HA, Vale W (1973) Synthetic analogues of the hypothalamic luteinizing hormone releasing factor with increased agonist or antagonist properties. *Biochemistry* 12:4616–4620
- Heber D, Odell WD (1978) Pituitary receptor binding activity of active, inactive, superactive and inhibitory analogues of gonadotropin-releasing hormone. *Biochem Biophys Res Comm* 82:67–73
- Wessels PL, Feeney J, Gregory H, Gormley JJ (1973) High resolution nuclear magnetic resonance studies of the conformation of luteinizing hormone releasing hormone (LH-RH) and its component peptides. *J Chem Soc Perkin* 2:1691–1698
- Chary KVR, Srivastava S, Hosur RV, Roy KB, Govil G (1986) Molecular conformation of gonadoliberin using two-dimensional NMR spectroscopy. *Eur J Biochem* 158:323–352
- Pappa EV, Zompra AA, Spyrianti Z, Diamantopoulou Z, Pairas G, Lamari FN, Katsoris P, Spyroulias GA, Cordopatis P (2010) Enzymatic stability, solution structure and antiproliferative effect on prostate cancer cells of leuprolide and new gonadotropin-releasing hormone peptide analogues. *Biopolymers (Peptide Sci)* 96:260–272
- Pappa EV, Zompra AA, Diamantopoulou Z, Spyrianti Z, Pairas G, Lamari FN, Katsoris P, Spyroulias GA, Cordopatis P (2012) Structure-activity studies of IGnRHIII through rational amino acid substitution and NMR conformational studies. *Biopolymers (Peptide Sci)* 98:525–534
- Stewart A, Sellar R, Wilson DJ, Millar RP, Zhi-Liang L (2008) Identification of a novel ligand binding residue Arg38(1.35) in

- the human gonadotropin-releasing hormone receptor. *Mol Pharm* 73:75–81
13. Momany FA (1976) Conformational energy analysis of the molecule, luteinizing hormone—releasing hormone 1. Native decapeptide. *J Am Chem Soc* 98:2990–2996
  14. Momany FA (1976) Conformational energy analysis of the molecule, luteinizing hormone releasing hormone 2. Tetrapeptide and decapeptide analogs. *J Am Chem Soc* 98:2996–3000
  15. Momany FA (1978) Conformational analysis of the molecule, luteinizing hormone-releasing hormone 3. Analog inhibitors and antagonists. *J Med Chem* 21:63–68
  16. Ling N, Vale W (1975) Analogs of luteinizing hormone releasing factor (LRF). Synthesis and biological activity of [(N<sup>2</sup>-Me)Leu<sup>7</sup>]LRF and [D-Ala<sup>6</sup>,(N<sup>2</sup>-Me)Leu<sup>7</sup>]LRF. *Biochem Biophys Res Commun* 63:801–806
  17. Freidinger RM, Verber DF, Perlow DS, Brooks JR, Saperstein R (1980) Bioactive conformation of luteinizing hormone-releasing hormone: evidence from a conformationally constrained analog. *Science* 210:656–658
  18. Akhrem AA, Golubovich VP, Kimarskii LI, Galaktionov SG (1978) Conformational energy analysis of the luliberin molecule. In: Symposium paper-IUPAC international symposium chemistry of natural product, 11th., vol 1, pp 29–30
  19. Golubovich VP, Kirnarskii LI, Akhrem AA, Galaktionov SG (1981) Spatial structure of luliberin molecules. *Bioorg Khim* 7, 819–831 (in Russian). *Chem. Abstr.* 95, 75744r
  20. Nikiforovich GV, Marshall GR (1993) Conformation-function relationships in LHRH analogs I. Conformations of LHRH peptide backbone. *Int J Pept Protein Res* 42:171–180
  21. Zimmerman SS, Scheraga HA (1977) Influence of local interactions on protein structure. I. Conformational energy studies of N-Acetyl-N'-Methylamides of Pro-X and X-Pro dipeptides. *Biopolymers* 16:811–843
  22. Gupta HM, Talwar GP, Salunke DM (1993) A novel computer modeling approach to the structures of small bioactive peptides: the structure of gonadotropin releasing hormone. *Proteins Struct Funct Genet* 16:48–56
  23. Struthers RS, Rivier J, Hagler A (1985) Molecular dynamics and minimum energy conformations of GnRH and analogs A. Methodology for computer-aided drug design. *Ann N Y Acad Sci* 439:81–96
  24. Guarnieri F, Weinstein H (1996) Conformational memories and the exploration of biologically relevant peptide conformations: an illustration for the gonadotropin-releasing hormone. *J Am Chem Soc* 118:5580–5589
  25. Mezei M, Guarnieri F (1998) Computer simulation studies of the fully solvated wild-type and mutated GnRH in extended and  $\beta$ -turn conformations. *J Mol Struct Dyn* 16:723–732
  26. Hodes ZI, Nemethy G, Scheraga HA (1979) Model for the conformational analysis of hydrated peptides. Effect of hydration on the conformational stability of the terminally blocked residues of the 20 naturally occurring amino acids. *Biopolymers* 18: 1565–1610
  27. Scheraga HA, Liwo A, Oldziej S, Czaplowski A, Pillardy J, Ripoll DR, Vila JA, Kazmierkiewicz R, Saunders JA, Arnautova YA, Jagielska A, Chinchio M, Nancias M (2004) The protein folding problem: global optimization of the force fields. *Front Biosci* 9:3296–3323
  28. Pincus M (1988) The chain build-up procedure in computing the structures of biologically active polypeptides and proteins. *Int J Quantum Chem Quantum Biol Symp* 15:209–220
  29. Miller MH, Scheraga HA (1976) Calculation of the structures of collagen models. Role of interchain interactions in determining the triple-helical coiled coil conformation. I. Poly-(glycyl-prolyl-prolyl). *J Polym Sci Symp* 54:171–200
  30. Powell MJD (1964) An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput J* 7:155–162
  31. Pincus MR, Woo J, Monaco R, Lubowsky J, Avitable M, Carty RP (2014) Conformational analysis of the effects of D- and L-amino acid substitutions for Gly 6 in GnRH; inference of the receptor-bound conformation. Accompanying paper. *Prot J* (submitted)