Journal of Peptide Science
J. Peptide Sci. **8**: 385–397 (2002)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.389

Spatial Organization and Conformational Peculiarities of the Callatostatin Family of Neuropeptides

I. N. ALIEVA, L. I. VELIEVA, D. I. ALIEV* and N. M. GODJAEV

Laboratory of Molecular Biophysics, Baku State University, Baku, Azerbaijan Republic

Received 5 March 2002
Accepted 13 April 2002

Abstract: The structures and conformational peculiarities of five members of the callatostatin family of neuropeptides, i.e. Leu- and Met-callatostatins, ranging in size from 8 to 16 amino acid residues have been investigated by a theoretical conformational analysis method. A comparative analysis of the conformational flexibilities of Met-callatostatin with those of the hydroxylated analogues, [Hyp^2]- and [Hyp^3]-Met-callatostatin has been carried out. Helically packed *C*-terminal pentapeptide in the structure of all investigated Leu-callatostatins are shown to be possible. The reason for the great number low-energy conformers for the callatostatin *N*-terminus is discussed. Copyright © 2002 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: neuropeptide; callatostatin; structure; conformation; conformational analysis

INTRODUCTION

The callatostatins are a family of insect neuropeptides isolated from the nervous tissues of the blowfly *Calliphora vomitoria* [1–3]. Their designation as callatostatins derives from (a) the species of origin; (b) the *C*-terminal homology to the allatostatins of the cockroaches *Diploptera punctata* and *Periplaneta americana* [4–7]. The allatostatins have an important role in the reproductive processes of insects as inhibitors of the synthesis and release of juvenile hormone from the corpus allatum. They also play a variety of roles in neurotransmission and neuromodulation, and act as neurohormones. Callatostatin neurones directly innervate muscles of the hindgut and the heart. The primary structures of these neuropeptides were assigned as shown in Table 1.

A hexadecapeptide designated callatostatin 1 was isolated from the thoracic ganglia, brains and heads. Callatostatins 2 and 3 were isolated from the heads and thoracic ganglia, respectively; they comprise the last 14 and 8 residues of callatostatin 1. Callatostatin 4, isolated from the thoracic ganglia, has a sequence (Table 1) in which Xaa is either Asp or Asn. This peptide, with a serine substitution for glycine at position 5, has a *C*-terminal pentapeptide sequence identical to that of allatostatins 3 and 4 of *D. punctata*. These four blowfly neuropeptides are referred to as Leu-callatostatins 1–4 to distinguish them from the Met-callatostatins, which have an Asp in the Xaa position but, more significantly, a Met-NH_2 for Leu-NH_2 substitution at the *C*-terminus (see Table 1 for sequence data). Met-callatostatin, or callatostatin 5 was identified from extracts of whole flies [3].

Two naturally occurring hydroxylate analogues of Met-callatostatin with a hydroxyproline substitution for proline at position 2 or 3 were designated as [Hyp^2]- or [Hyp^3]-Met-callatostatin. According to Duve *et al.* [8,9], both [Hyp^2]- and [Hyp^3]-Met-callatostatins are much more potent than Met-callatostatin as inhibitors of juvenile hormone synthesis in *D. punctata*. The existence of the hydroxyproline analogues of Met-callatostatin suggests

*Correspondence to: Dr D. I. Aliev, Laboratory of Molecular Biophysics, Physical Department, Baku State University, Z.Khalilov 23, 370148, Baku, Azerbaijan Republic;
e-mail: dshaliev@hotmail.com

that the motif for prolyl hydroxylation in *C.vomitoria* is more variable than those known from mammalian and other invertebrate studies [4–7]. In vertebrates, as well as invertebrates, biologically active peptides containing hydroxyprolyl residues appear to be extremely rare. The only known examples in mammals are the luteinizing hormone releasing hormone, Glp-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2, where in the adult hypothalamus the ratio of hydroxylated Hyp^9 to non-hydroxylated Pro^9 varies from 11% to 70%, and bradykinin (Arg-Pro-Pro^3 or Hyp^3-Gly-Phe-Ser-Pro-Phe-Arg). Both mammalian peptides have Xaa-Pro-Gly in a β-turn conformation, a motif considered ideal for hydroxylation. The same motif is found in collagen where it is favoured by collagen prolyl 4- and prolyl 3-hydroxylases. In almost all instances, the presence of 4-hydroxyproline (and to a more limited extent, 3-hydroxyproline) in collagens and other proteins is thought to occur only in the Y positions of X-Y-Gly triplets [10]. An exception to this preferred prolyl 4-hydroxylase specificity in collagens occurs in the subcuticular epithelium of earthworms where $(Pro\text{-}Ala\text{-}Gly)_n$ is a much better substrate than $(Ala\text{-}Pro\text{-}Gly)_n$. Met-callatostatin, with the sequence Gly-Pro-Pro-Tyr-, obviously presents yet another type of conformation that is acceptable for prolyl hydroxylase activity, as does the neuropeptide from *Aedes aegypti* (Glp-Arg-Pro-Hyp-Ser).

The method of theoretical conformational analysis as used to describe the different types of spatial structures for the Leu-callatostatin and Met-callatostatin neuropeptides. The conformational behaviours of both [Hyp^2]- and [Hyp^3]-Met-callatostatins derivatives have been investigated. The reasons for the large number of members of the callatostatin neuropeptide family are discussed.

METHOD AND MODELS FOR CALCULATIONS

The problem of spatial organization and conformational flexibilities of the callatostatins in an aqueous environment was resolved by use of the theoretical conformational analysis method. The bond distances and the values of valence angles were regarded as invariable. Only dihedral angles were taken to be the intrinsic degrees of freedom. Stable conformations were found by the total conformational energy minimization. The basic features of the computational method are presented below.

Energy Calculations

The total conformational energy (E_{total}) is estimated as a sum of independent contributions of interaction energies for a pair of non-bonded atoms. The following peculiar features were used:

(i) The energy of non bonded atom–atom interaction (E_{NV}) is described by the Lennard-Jones '6–12' potential with the parameters proposed by Momany *et al.* [11];
(ii) electrostatic energy (E_{EL}) was calculated in a monopole approximation corresponding to Coulomb's law with partial charges of atoms as suggested [11];
(iii) the effective dielectric constant (ε) was taken to be equal to 10, as described by Popov *et al.* [12] in all calculations used here;
(iv) the energy of hydrogen bond formation was calculated based on the Morse potential [11] and the dissociation energy of the hydrogen bond was taken to be -6.2 kJmol^{-1} at an NH...OC distance of $r_o = 1.8$ Å;
(v) the intrinsic energy of a molecule also includes the torsion potentials (E_{TOR}), describing the barriers of the inner rotation between atoms that have a 1–4 relationship; the values of the torsional barrier heights are given in [11].

Simulation of An Aqueous Environment

A molecular conformation is largely determined by its environment. The multipolar array of charges chosen to represent the electrical asymmetry of the structure is largely shielded from the surrounding bulk solvent of high dielectric constant by associated groups of the polypeptide chain and possibly also by solvent molecules complexed (e.g. H-bonded) to the amide group. The effective dielectric constant for the coulombic interactions between a neighbouring pair of peptide units buried within a solvated polypeptide chain may therefore be expected to approximate more nearly to the dielectric constant of the chain than to that of the bulk solvent. However, the dimensions calculated for polypeptide chains are found to be fairly insensitive to the value chosen for ε, provided that it falls within the range of about 2 to 10 [12]. Although $\varepsilon = 4$ is typically used for calculations with proteins [13] it was assumed that small neuropeptides may be less shielded from the bulk solvent in comparison with the globular proteins, thus the highest value of ε ($\varepsilon = 10$) in the recommended range is used

for callatostatins, in agreement with the results of Popov [12] who investigated the effects of various solutions on the conformations of oligopeptides by molecular mechanics calculations. The best agreement between the experimental and theoretical data was found for effective dielectric constant values as follows: for CCl_4 solution, $\varepsilon = 4$; for $CHCl_2$ solution, $\varepsilon = 6–7$; for aqueous solution, $\varepsilon = 10$.

Description and Selection of Initial Conformations

The conformational state of each amino acid residue is conveniently described by the backbone φ, ψ and side chain χ_1, χ_2 dihedral angles. For a stable conformation the φ and ψ dihedral angles of the backbone chain are located in low-energy regions described by the symbols R, B, L and P on the conformational map (Figure 1).

The 'form of a residue' denotes the region of its backbone dihedral angle (R, B, L or P). Each conformational state of a residue is characterized by $X^n_{ij}\ldots$, where X characterizes the backbone φ, ψ angle regions (B, R. L or P), n is the number of a residue in the sequence and subscripts $i,j,\ldots$ specify the position of the side chain χ_1, χ_2 ..., respectively, so that i or $j = 1$ corresponds to the angle χ (χ_1 or χ_2) in the range 0° to 120°; a value of two corresponds to the angle range 120° to −120° and three to −120° to 0°. R, B and L forms may occur with alanine-type residues (16 amino acids with no side chain branching at C^β atoms),as well as with valine and isoleucine (excluding the L-L combination, which is never observed in proteins), but only R and B for proline. The combination of the backbone form of the residue in a given amino acid sequence will specify the backbone form of a fragment.

All backbone forms of a dipeptide can be classified into two types, referred to as shapes: folded (f) and extended (e). The f-shape is represented by R-R, R-B, B-P, L-L, L-P, L-R, P-B and B-L forms and the e-shape by B-B, B-R, R-L, L-B, L-R, R-P, P-L and P-P forms. Forms belonging to a particular shape have an analogous peptide chain contour and a similar mutual arrangement of backbones and side chains and thus should exhibit similar medium-range interaction potentialities. For a tripeptide, all possible backbone forms may be specified by four shapes, i.e. ff, fe, ef and ee. The suggested classification is based on a 'tree' principle, all conformations being specified according to the backbone form with a following specification into shape. The identifier system will be used to describe all structures at the intermediate steps of the calculations with the numerical values of geometry parameters produced only at the final stage. The dihedral angle values corresponding to the lowest energy states of monopeptides were used as the starting conformation. Bond lengths, valence angles and partial charges of the atoms of the neuropeptides were chosen as given in [11]. The nomenclature and conventions adopted are those recommended by IUPAC-IUB [14].

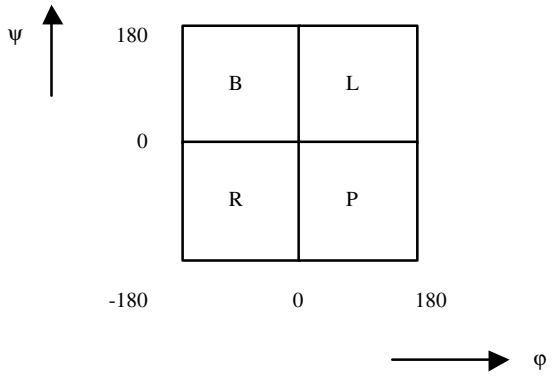

Figure 1 The energy regions of conformational map.

Protocol of Computer Experiment

The conformational analysis of neuropeptides was carried out through a fragmental calculation (Figure 2).

The optimization procedure also has several steps. A final conformation obtained in a preliminary step is taken as an initial one for the next step. A procedure for the minimization of the neuropeptide global energy was conducted by the method of conjugate gradients using the programs written by Godjaev *et al.* [15].

The energy minimization was repeated until the minimal values of the global energy remained at a constant level. The conformations with the lowest minimum of global energy are presented below.

RESULTS AND DISCUSSION

Conformational Analysis of Phe^6-Gly^7-Met^8 Tripeptide

The backbone residues that make up this tripeptide (Figure 2a) can be in R, B, L and P forms. Therefore 320 conformations corresponding to four possible shapes and 36 forms of the peptide fragment skeleton were taken into account for this part

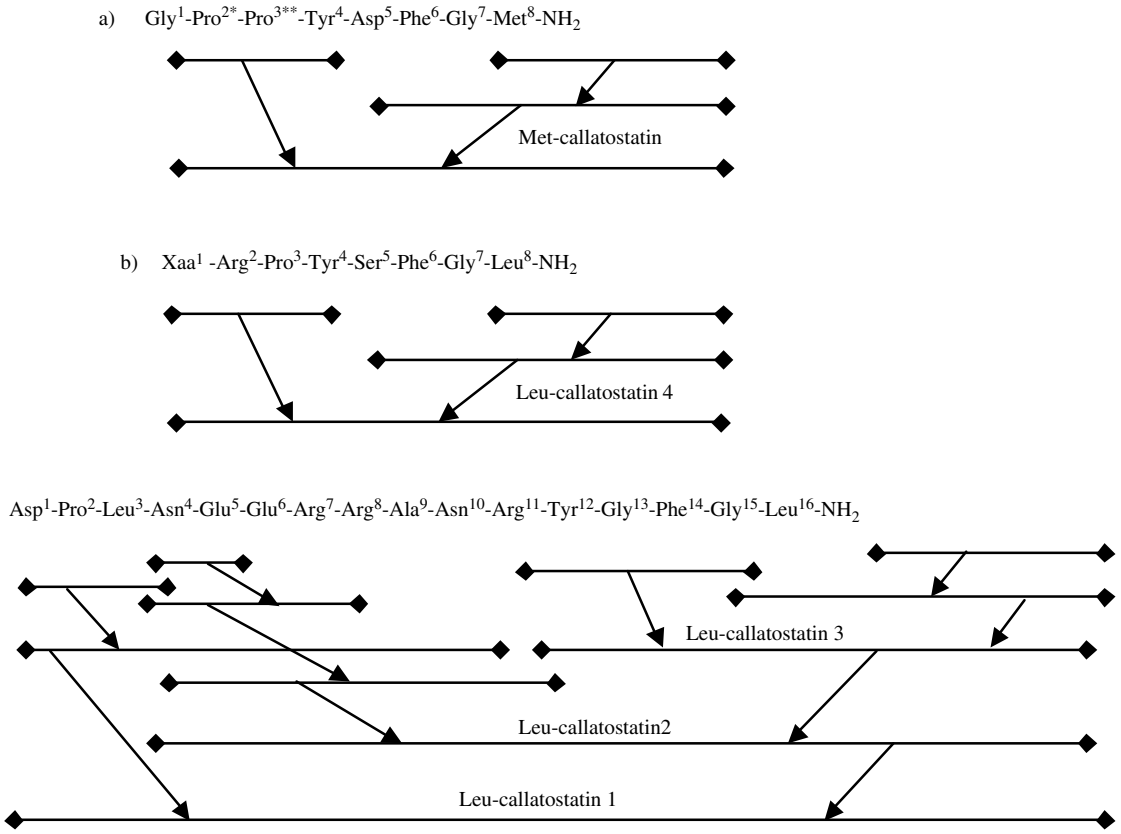

Figure 2 Calculation schemes of the Met-callatostatin (a) and Leu-callatostatins 1–4 (b) (*,**-hydroxyproline substitution for proline at position 2 or 3 of [Hyp^2]- or [Hyp^3]Met-callatostatin, respectively).

Table 1 Amino Acid Sequences of the Callatostatins from *C. vomitoria* (cast, callatostatin; sequence nomenclature for the Callatostatins in Brackets is from [1])

Leu-cast1 (callatostatin1)	Asp-Pro-Leu-Asn-Glu-Glu-Arg-Arg-Ala-Asn-Arg-Tyr-Gly-Phe-Gly-Leu-NH_2
Leu-cast2 (callatostatin2)	Leu-Asn-Glu-Glu-Arg-Arg-Ala-Asn-Arg-Tyr-Gly-Phe-Gly-Leu-NH_2
Leu-cast3 (callatostatin3)	Ala-Asn-Arg-Tyr-Gly-Phe-Gly-Leu-NH_2
Leu-cast4 (callatostatin4)	Xaa-Arg-Pro-Tyr-Ser-Phe-Gly-Leu-NH_2
Met-cast (callatostatin5)	Gly-Pro-Pro-Tyr-Asp-Phe-Gly-Met-NH_2
[Hyp^2]Met-cast	Gly-Hyp-Pro-Tyr-Asp-Phe-Gly-Met-NH_2
[Hyp^3]Met-cast	Gly-Pro-Hyp-Tyr-Asp-Phe-Gly-Met-NH_2

of the calculation. The greatest number of low-energy conformations (49%) corresponds to the relative energy region of 0–3 kJmol^{-1}. They are formed of ef and ff shapes in which the R and B forms of the peptide skeleton are identical. As the energy of the conformations corresponding to the L form was found to be higher than those for the R and B forms, in the following step of the calculations, the L form was not considered. According to the calculation result conformational behaviours of the tripeptide were induced by the side chains of the Phe^6 and Met^8 amino acid residues. Specific interaction between these pairs of residues in the ef and ff shapes made an important contribution to the stabilization of the low-energy conformations. For fe and ee shapes the contact between Phe^6 and Met^8 was negligible. In these shapes, other interactions and torsional contributions were also less than those of other shapes. The distribution of low-energy conformations between shapes of the tripeptide fragment are represented in Table 1.

Table 2 Energetic Distribution of Stable Conformations of Penta- and Tripeptide Fragments of Met-callatostatin Molecules

No	Shape	E_{rel} (kJ mol^{-1})					
		0–1	1–2	2–3	3–4	4–5	5–10
		Tyr^4-Asp^5-Phe^6-Gly^7-Met^8-NH$_2$					
1	efff	—	—	—	1[a]	4	28
2	ffff	—	—	2	6	8	14
3	eeff	—	1	—	2	5	37
4	feff	—	—	—	—	—	34
5	efee	—	—	1	—	1	11
6	ffee	1	—	1	1	2	6
7	eeee	—	—	—	2	2	12
8	feee	—	—	—	1	1	9
9	efef	—	—	—	—	—	4
10	ffef	—	—	—	1	1	1
11	eefe	—	—	—	—	3	20
12	fefe	—	—	—	—	1	17
13	eeef	—	—	—	—	1	7
14	feef	—	—	—	—	—	8
15	effe	—	—	—	—	—	4
16	fffe	—	—	—	—	2	4
		Gly^1-Pro^2-Pro^3					
1	fe	4	—	—	—	—	—
2	ee	2	—	2	—	—	—
		Gly^1-Hyp^2-Pro^3					
1	fe	2	1	9	4	2	2
2	ee	—	—	3	6	1	7
		Gly^1-Pro^2-Hyp^3					
1	fe	10	7	—	1	—	—
2	ee	2	7	3	5	1	—

[a] Indicates the number of calculated conformations.

Tyr^4-Asp^5-Phe^6-Gly^7-Met^8-NH$_2$ Pentapeptide Fragment

The starting conformations of this fragment were obtained by combining the most stable structures of the amino acid residues of Tyr^4, Asp^5 and the calculation results of the Phe^6-Met^8 fragment. Values of ±60°, 180°and value 90° were given to the χ_1 and χ_2 dihedral angles of both tyrosine (Tyr^4) and aspartic acid (Asp^5).

Therefore, as starting conformations of the pentapeptide fragment Tyr^4-Met^8, 441 conformations belonging to 16 different shapes and 72 forms of backbone skeleton were considered and their geometry was optimized through energy minimization. The calculation results indicated that about 20 conformations corresponded to the relative energy region of 0–5 kJmol^{-1} (Table 2). No significant energy difference was found between the shapes and as a result conformational rigidity of the *C*-terminal pentapeptide was assumed. Low-energy conformations are made up ffee, eeff and ffff shapes. In the ffff shape the terminal residues Tyr^4 and Met^8 are close to each other and an effective interaction is found to be the contact between them. Hydrogen bonding between carbonyl oxygen of Tyr^4 and amide hydrogen of Met^8 is characteristic for all the low-energy structures of this shape. The lowest energy conformer, namely the global conformer of the pentapeptide has the ffee shape. The contacts between Tyr^4 and Asp^5 or Phe^6 make an important contribution towards the dispersion (non-bonded)

Table 3 Energetic Distribution of Met-callatostatin Molecule Conformations

Molecule	Met-callatostatin																[Hyp²] Met-callatostatin												[Hyp³] Met-callatostation											
Shape	fefeff	feeffff	fefeeff	eefeff	feeeff	fefeeee	eeeffff	eefeeff	feeeeff	eeeeff	fefeefe	fefeff	feeffff	fefeeff	eefeff	feeeff	fefeeee	eeeffff	eefeeff	feeeeff	eeeeff	fefeefe	fefeff	feeffff	fefeeff	eefeff	feeeff	fefeeee	eeeffff	eefeeff	feeeeff	eeeeff	fefeefe	fefeff	feeffff	fefeeff	eefeff	feeeff	fefeeee	eeeffff	eefeeff	feeeeff	eeeeff	fefeefe
Number of conformations	8	11	16	4	8	–	3	7	8	3	6	12	15	7	10	4	8	4	4	6	3	4	28	20	23	6	9	8	7	2	10	7	8	8	7	2	10	7	8	8
E total (kJ mol^{-1})	–	2	–	–	–	–	3	2	1	1	2	3	7	1	3	1	2	3	3	5	3	4	4	17	6	3	6	8	7	2	10	7	8	8	7	2	10	7	8	8
0	1	1	1	1								4	1	2									3																	
1	3	1	5	–	1							1	1	–									1	2																
2	4	2	–	1	3	1	2					–	2	1	2	1							1	–															
3	–	3	6	2	4	2	1	1	3			4	2	3	3	2	1						–	1															
4	–	2	4	–	–	1	3	1	1	3		–	2	–	2	–	–	1	1	1			4	3	3	2													
5		2	–	–	–	3	2	1	2	1		3	7	1	3	1	7	3	3	5	3	4	20	17	6	6	7	2	10	7	8								
10	8	11	16	4	8	7	8	3	6	12	15	7	10	4	8	4	4	6	3	4	28	23	9	8	7	2	10	7	8	8								

interaction (−14.6 and −11.7 kJmol^{-1}, respectively) in this conformation. The fully extended form of the pentapeptide skeleton as well as the conformations of the other shapes are favourable only by dispersion contacts between backbone and amino acid side chains.

Gly^1-Pro^2-Pro^3, Gly^1-Hyp^2-Pro^3 and Gly^1-Pro^2-Hyp^3 Tripeptides

These tripeptides form the *N*-terminal parts of Met-callatostatin molecule and its hydroxylate analogues (Figure 2,a). The glycine residue at the first position has two forms (R and B) of backbone skeleton which are equivalent to L and P forms, respectively. Proline and oxyproline can be in R and B forms only. So the starting conformations belonging to the four shapes and 16 forms of the main chain include 12 conformations of Gly^1-Pro^2-Pro^3 and 70 conformations for both Gly^1-Hyp^2-Pro^3 and Gly^1-Pro^2-Pro^3 (the value of the oxyproline χ angle is a variable parameter). The calculation results (Table 2) indicate that hydroxyproline substitution for proline at position 3 does not lead to the re-distribution of low-energy conformations of the tripeptide. In both cases folded-extended backbone chains belonging to the fe shape are energetically favourable. The identical substitution for proline at the second position leads to the decreasing of number low-energy conformational states of the tripeptide Gly^1-Hyp^2-Pro^3 (Table 2). Quantitative evaluation of interresidue contacts showed that non-bonded interactions and torsional contributions for the extended form of the backbone skeleton (ee shape) are less than those of other shapes.

(Hyp^2)-, and (Hyp^3)-Met-callatostatins

On the basis of the knowledge of the conformational possibilities of the above penta- and tripeptide fragments in free states, conformational investigation of the Met-callatostatin molecule and its modified analogues was carried out. The starting conformations for these molecules were obtained by combining the most stable structures of the respective fragment and the dihedral angles were optimized through energy minimization. The structures of the investigated molecule were classified into 40 different shapes of the peptide skeleton, obtained as combinations of 16 shapes of the Tyr^4-Met^8 fragment

Table 4 The Geometrical Parameters (degrees) of the Met-callatostatin Molecules and Hydroxylate Analogues for Low-energetic Conformations

Amino acid	Conformation								
	Met-callatostatin			[Hyp^2]Met-callatostatin			[Hyp^3]Met-callatostatin		
Gly^1	$\varphi = 138$	$\psi = -72$	$\omega = 181$	$\varphi = -70$	$\psi = -73$	$\omega = 175$	$\varphi = 137$	$\psi = -72$	$\omega = 180$
Pro^2 (Hyp^2)	$\psi = 164$		$\omega = 175$	$\psi = 161$	$\omega = 179$	$\chi = -86$	$\psi = 162$		$\omega = 175$
Pro^3(Hyp^3)		$\psi = -42$	$\omega = 179$		$\psi = -45$	$\omega = 177$	$\psi = -54$	$\omega = 180$	$\chi = 154$
Tyr^4	$\varphi = -121$	$\psi = 160$	$\omega = 183$	$\varphi = -118$	$\psi = 161$	$\omega = 182$	$\varphi = 101$	$\psi = -38$	$\omega = 177$
	$\chi_1 = 51$	$\chi_2 = 94$		$\chi_1 = 51$	$\chi_2 = 94$		$\chi_1 = 59$	$\chi_2 = 87$	
Asp^5	$\varphi = -91$	$\psi = -32$	$\omega = 191$	$\varphi = -89$	$\psi = -31$	$\omega = 193$	$\varphi = -93$	$\psi = -30$	$\omega = 174$
	$\chi_1 = 61$	$\chi_2 = 97$		$\chi_1 = 61$	$\chi_2 = 97$		$\chi_1 = 61$	$\chi_2 = 95$	
Phe^6	$\varphi = -70$	$\psi = 145$	$\omega = 189$	$\varphi = -69$	$\psi = 145$	$\omega = 189$	$\varphi = -75$	$\psi = -48$	$\omega = 178$
	$\chi_1 = -57$	$\chi_2 = 101$		$\chi_1 = -57$	$\chi_2 = 101$		$\chi_1 = -56$	$\chi_2 = 99$	
Gly^7	$\varphi = 70$	$\psi = -72$	$\omega = 181$	$\varphi = 71$	$\psi = -74$	$\omega = 180$	$\varphi = -67$	$\psi = -57$	$\omega = 180$
Met^8	$\varphi = -116$	$\psi = -51$	$\omega = 174$	$\varphi = -117$	$\psi = -52$	$\omega = 174$	$\varphi = -101$	$\psi = -52$	$\omega = 178$
	$\chi_1 = 62$	$\chi_2 = 181$	$\chi_3 = 180$	$\chi_1 = 62$	$\chi_2 = 181$	$\chi_3 = 180$	$\chi_1 = 184$	$\chi_2 = 177$	$\chi_3 = 180$
Form	PBRB$_1$R$_1$B$_3$PR$_{3222}$			RBRR$_1$R$_1$R$_3$PR$_{3222}$			PBRR1R1R3RR2222		
Shape	fefefff			fefefff			feffffff		
	Energy contribution (kJ mol^{-1})								
E_{NV}[a]	−130.5			−153.1			−137.6		
E_{EL}	−4.6			15.1			15.1		
E_{TOR}	15.9			18.0			10.9		
E_{TOTAL}	−119.2			−120.0			−111.6		

[a] Energy of hydrogen bonding was included in E_{NV}.

Table 5 Inter-residue Interaction Energies (kJ mol^{-1}) of Met-callatostatin Molecules for the Low-energy Conformations[a]

	Gly^1	Pro^2[Hyp^2]	Pro^3[Hyp^3]	Tyr^4	Asp^5	Phe^6	Gly^7	Met^8	
1	5.73	−13.26	−1.88	−0.46	−0.17	−0.04	0.00	−0.04	GLY^1
2	5.64	−1.80	−2.09	−0.33	−0.25	−0.08	0.04	0.00	
3	5.69	−13.59	−1.21	−0.37	−0.33	0.00	0.00	0.00	
		3.18	−1.25	−11.00	−0.88	−0.17	0.00	−5.81	PRO^2
		3.09	−18.73	−14.60	3.26	0.46	−0.01	−11.12	[HYP^2]
		3.05	−14.76	−8.86	−1.80	−5.35	0.04	0.04	
			0.88	−10.16	−0.71	−0.08	0.00	0.04	PRO^3
			1.00	−10.50	−0.75	−0.08	0.00	0.04	[HYP^3]
			3.93	−10.66	0.54	−7.32	−5.02	−1.38	
				−0.17	−8.32	−10.25	−1.04	−1.04	TYR^4
				0.46	−8.14	−12.88	−1.00	−4.06	
				0.17	−13.76	−8.72	−3.97	−7.15	
					−3.47	−10.75	−1.96	−4.85	ASP^5
					−3.43	−10.29	−1.96	−4.56	
					−3.39	−9.66	−3.68	−6.27	
						2.50	−1.38	−5.81	PHE^6
						2.59	−1.42	−6.31	
						0.58	−0.96	−4.00	
							5.44	−2.93	GLY^7
							5.39	−2.84	
							5.60	−2.30	
								0.58	MET^8
								0.58	
								0.96	

together with four shapes of the *N*-terminal tripeptide fragments. Also for each backbone form (shape) every possible combination of side chain angles of the residues was taken into account.

The distribution of conformations of the low-energy states Met-callatostatins and the contribution of non-bonded (including energy of hydrogen bonding), electrostatic and torsional interaction to the most stable conformations of the molecules are given in Tables 3 and 4.

The global conformation of Met- and [Hyp^2]-Met-callatostatins is in the $PBRB_1R_1B_3PR_{3222}$ and $RBRB_1R_1B_3PR_{3222}$ form, respectively, and belongs to the fefefff backbone shape. This conformation is optimal with regard to the dispersion energy contribution which plays an important role in the stabilization of the density packing structures of the molecules. The electrostatic energy contribution is found to be small because contacts between charged groups of atoms are weak in water solutions. Consideration of the energy of interresidue interaction of the Phe^6 residue (Table 5) showed that it may be considered to be the centre of dispersion stabilization. The total energy of interaction of this residue with Tyr^4, Asp^5 and Met^8 is found to be −26.8,−29.5, −20.7 kJmol^{-1} for Met-, [Hyp^2]-Met- and [Hyp^3]-Met-callatostatins.

The $PBRR_1R_1R_3RR_{2222}$ conformation is the most stable state for the [Hyp^3]-Met-callatostatin molecule and belongs to the feffff shape (Figure 3). This is characterized by the folding of the majority of the conformations of the residues and is considered to be a helical-type structure of [Hyp^3]-Met-callatostatin. All three molecules are characterized by an identical structure of the *C*-terminal part, having a folded form of skeleton. The conclusion from this study is that the geometry of the pentapeptide fragment including the amino acid sequence Gly^1-Pro^2-Pro^3 could be essential for the biological activity of callatostatins. The replacement of amino acid residues in the naturally occurring sequence, i.e. modification of the chemical structure of the peptide molecule is accompanied by a change of conformational possibilities, intra- and interresidue interactions and also stability of analogues to an peptidase action. The strong interresidual interactions between Hyp^2 and Pro^3 (−18.8 kJmol^{-1}) and Pro^2-Hyp^3 (−14.6 kJmol^{-1}) allows the assumption

Table 6 The Energies (kJ mol^{-1}) and the Shapes of Preferred Conformations of the Leu-callatostatins 1–3 C-terminus Pentapeptide

Number	Shape	Conformation	$\sum E_{int}ra$ (mono)[a]	Inter-residue interaction energies										Energy contribution			
				Tyr^4 Gly^5	Gly^5 Phe^6	Phe^6 Gly^7	Gly^7 Leu^8	Tyr^4 Phe^6	Gly^5 Gly^7	Phe^6 Leu^8	Tyr^4 Gly^7	Gly^5 Leu^8	Tyr^4 Leu^8	E_{NV}	E_{EL}	E_{TOR}	E_{total}
1	ffff	$B_2PR_2RR_{21}$	4.56	−8.15	−4.01	−3.22	−4.85	−5.60	−2.72	−3.89	−2.72	−5.81	−10.71	−99.1	18.4	9.2	−71.5
2	ffef	$B_1PR_1PR_{21}$	−1.13	−6.15	−4.26	−4.56	−3.42	−10.54	−2.84	−13.67	−9.95	−0.08	0.29	−89.1	13.0	8.4	−67.7
3	effe	$B_3RB_2PR_{32}$	6.94	−3.09	−4.18	−5.56	−6.77	−1.80	−1.38	−7.53	−0.29	−2.59	−19.40	−92.8	8.8	13.4	−70.7
4	efee	$R_2PB_3BR_{32}$	4.77	−8.49	−7.32	−1.63	−7.15	−16.35	−0.79	−17.65	−7.78	−0.21	−0.63	−89.9	15.0	8.8	−66.1

[a] The total intra-residue interaction energies of the monopeptides in a given sequence.

Table 7 The Energies (kJ mol^{-1}) and Shapes of Preferred Conformations of the Leu-callatostatins 4 *C*-terminus Pentapeptide

Number	Shape	Conformation	$\sum E_{intra}$ (mono)[a]	Inter-residue interaction energies										Energy contribution			
				Tyr^4 Ser^5	Ser^5 Phe^6	Phe^6 Gly^7	Gly^7 Leu^8	Tyr^4 Phe^6	Ser^5 Gly^7	Phe^6 Leu^8	Tyr^4 Gly^7	Ser^5 Leu^8	Tyr^4 Leu^8	E_{NV}	E_{EL}	E_{TOR}	E_{total}
1	efff	$B_3R_1B_2PR_{31}$	5.39	−6.36	−7.11	−5.69	−7.02	−2.17	−1.50	−8.03	−0.38	−2.63	−20.91	−103.3	13.4	10.4	−79.5
2	ffff	$R_2R_1B_3PR_{31}$	4.93	−11.75	−8.28	−5.65	−6.77	−2.05	−1.50	−7.82	−0.25	−2.55	−3.51	−100.8	12.5	11.3	−76.9
3	feff	$R_1B_1B_3PR_{21}$	4.06	−12.55	−7.11	−1.63	−4.39	−6.98	−0.96	−7.48	0	−2.84	0.08	−89.1	15.0	8.8	−65.2
4	efef	$B_2R_1B_1RB_{21}$	−0.17	−15.60	−9.82	−6.48	−2.93	−11.58	−0.92	−7.28	−6.77	0	−0.17	−95.3	15.9	5.8	−73.6
5	eeee	$B_2B_2R_2BR_{21}$	1.38	−13.72	−4.60	−4.01	−3.64	−5.48	−5.90	−8.15	−0.12	−15.68	−15.68	−96.6	15.0	10.0	−71.5

[a] The total intra-residue interaction energies of the monopeptides in a given sequence.

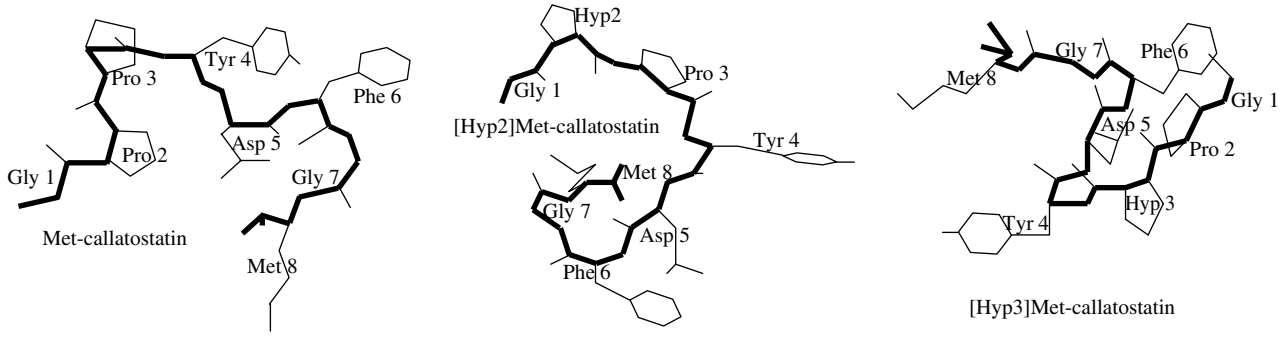

Figure 3 The stereoviews of the Met-callatostatin neuropeptides structures in the optimized conformations based on the coordinates of atoms.

Table 8 The Forms and Shapes of the Leu-callatostatins Backbone (cast, callatostatin)

Shape	Conformation	Energy contribution (kJ/mol)			
		E_{NV}	E_{EL}	E_{TOR}	E_{total}
Leu-cast 4					
fefffff	$R_{21}B_3RR_1R_2R_2RR_{21}$	−180.0	13.2	21.2	−145.6
fefeeff	$R_{22}B_1RB_2B_1B_3PR_{21}$	−165.9	12.7	11.2	−142.0
fefefff	$R_{21}B_1RB_3R_1B_2PR_{32}$	−159.0	10.4	10.8	−137.8
Leu-cast 3					
fefffff	$RB_{11}R_2B_3PB_2PB_{21}$	−165.5	22.8	11.1	−131.6
ffeffef	$RR_{22}B_3B_1PR_1PR_{32}$	−171.4	24.7	13.9	−132.8
fffefff	$RR_{11}R_2R_3PB_2PB_{32}$	−177.4	25.9	21.4	−130.1
ffefffff	$RR_{11}B_1B_3PR_2RR_{21}$	−166.8	24.6	12.5	−129.7
Leu-cast 2					
ffffeffffff	$R_2R_{11}R_3R_2B_3R_2RB_{11}R_2B_3PB_2PB_{21}$	−274.5	−17.4	102.5	−154.6
ffffeefffff	$R_2R_{11}R_3R_2R_3L_2BL_{11}R_2B_3PB_2PB_{21}$	−273.8	17.6	121.7	−134.5
ffffefffefff	$R_2R_{11}R_3R_2B_1R_2RR_{11}R_2R_3PB_2PB_{32}$	−273.7	15.8	113.8	−144.1
Leu-cast 1					
eefffeffefffff	$B_1BR_2R_{11}R_3R_2B_3R_2RB_{11}R_2B_3PB_2PB_{21}$	−317.7	46.9	70.7	−200.1
eefffeefefffff	$B_3BR_2R_{11}R_3R_2R_3L_2BL_{11}R_2B_3PB_2PB_{21}$	−315.6	65.9	50.8	−198.9
eefffefffefff	$B_1BR_2R_{11}R_3R_2B_1R_2RR_{11}R_2R_3PB_2PB_{32}$	−311.0	63.2	49.0	−198.8
effffeefefffff	$B_2RR_2R_{11}R_3R_2R_3L_2BL_{11}R_2B_3PB_2PB_{21}$	−331.4	82.1	51.2	−198.1

that [Hyp^2]-Met- and [Hyp^3]-Met-callatostatins are more stable against peptidases capable of degrading the Pro-Pro- bond in all known regulatory peptides. Therefore these molecules should have a prolonged inhibitory effect compared with Met-callatostatin. The hydroxyprolyl residue most probably imparts a special conformational status that is important not only in receptor recognition but also (at least when the peptides are tested in cockroach haemolymph) as a protection against degradation, a factor that increases relative potency.

Conformational Analysis of the Leu-callatostatin 1–4 Neuropeptides

The spatial structure and conformational properties of four Leu-callatostatins are shown in Table 1 and Figure 2b. All the neuropeptides are *C*-terminally amidated. The *C*-terminal pentapeptide sequence -Tyr-Gly-Phe-Gly-Leu-NH_2 of callatostatins 1–3 is identical to that of allatostatin 1 from *D. punctata*. The octapeptide callatostatin 4 shares with allatostatins 3 and 4 the substitution of serine for glycine at position 4 from the *C*-terminus.

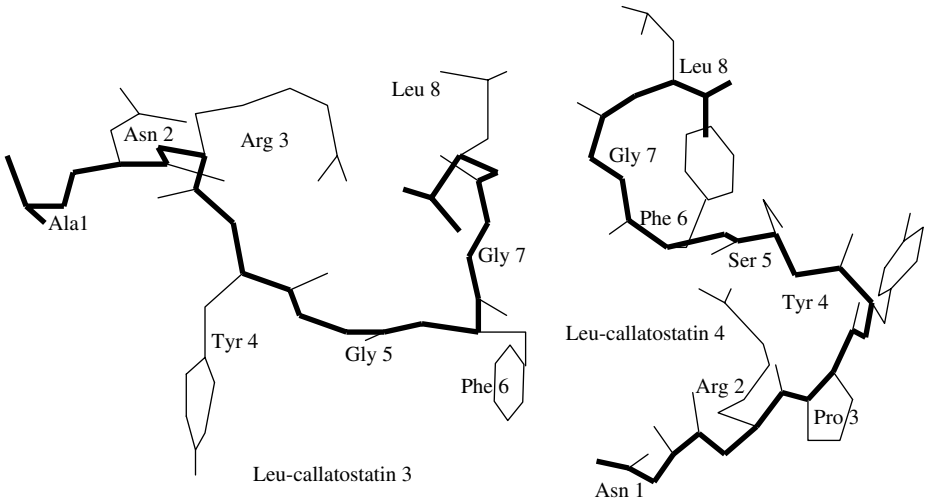

Figure 4 The stereoviews of the Leu-callatostatin 3 and Leu-callatostatin 4 structures in the optimized conformations based on the coordinates of atoms.

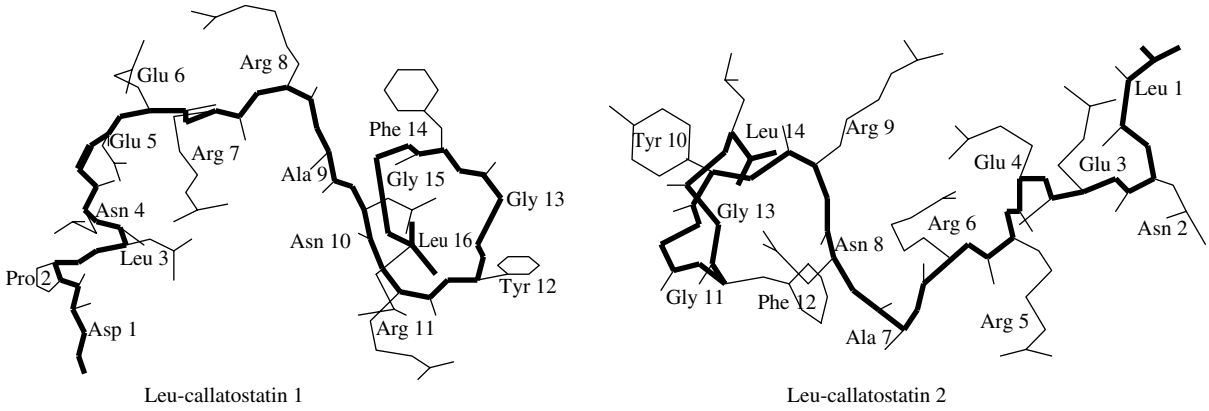

Figure 5 The stereoviews of the Leu-callatostatin 1 and Leu-callatostatin 2 structures in the optimized conformations based on the coordinates of atoms.

An interesting feature of hexadecapeptide callatostatin 1 is the presence of a pair of arginine residues at positions 7 and 8. A similar potential dibasic cleavage site (-Lys-Arg-) occurs at positions 9 and 10 of the cockroach octadecapeptide allatostatin 5. The tetradecapeptide callatostatin 2 lacks the first two *N*-terminal residues — Asp-Pro-. Of the other peptides, callatostatins 1 and 2 were more potent than the shorter octapeptide callatostatin 3. This finding [1] is in agreement with structure–activity studies of the cockroach allatostatins [27]. Conformational analysis of the Leu-callatostatins ranging in size from 8 to 16 amino acid residues was carried out through a fragmental calculation based on the conformational possibilities of the complicated molecular fragments as shown in Figure 2b. Only conformations are retained whose energies are smaller than the cut-off values. This cut-off value is usually taken as 3 kJmol^{-1} above the lowest energy, but it can be varied to make sure than higher energy states can justifiably be neglected.

Tyr-Gly-Phe-Gly-Leu-NH_2 Pentapeptide Fragment

Some 200 initial conformations corresponding to 16 shapes of the Leu-callatostatins 1–3 pentapeptide segment were taken into account for the first stage calculation experiment. The results showed only a very restricted set of low-energy conformations. The lowest conformer forms a helical-type structure with the ffff backbone shape (Table 6). Hydrogen bonding between the pair of residues at

the opposite end of the pentapeptide *C*-terminus NH (Leu)... CO (Tyr) was shown to be possible. The efff and ffef shape conformations were also found to be important for energy aspects of the Leu-callatostatins 1–3 neuropeptides. Table 6 represents energies of interresidue interactions in the most preferable conformations of the *C*-terminus pentapeptide. The Phe residue displays effective dispersion interactions in all the available conformations of the pentapeptide. So the conclusion was made that Phe is important for the conformational stability of the Leu-callatostatins 1–3.

Leu-callatostatin 4 is different from any of the Leu-callatostatins 1–3 in having a serine substitution at position 5, so some changes in the conformational properties were observed (Table 7). In addition to folded structures the particularly extended conformations (eeff shape) were shown to be possible. The conformational study of the all Leu-callatostatins 1–4 *N*-terminus was carried out according to the scheme in Figure 2b. In the final stage the most favourable conformations were obtained as results of both *N*-terminal and *C*-terminal pentapeptide fragments overlapping. The available conformations of all Leu-callatostatins are presented in Table 8.

All the low-energy conformations of Leu-callatostatin 1 and Leu-callatostatin 2 may be divided into two groups that differ in location of the helically packed segments in the neuropeptide spatial organization. The first group consists of two helically packed segments at opposite ends of the molecule. They are separated by a rigid structure formed by the Arg^{7}-Arg^{11} or Arg^{5}-Arg^{9} amino acid sequences in Leu-callatostatin 1 and Leu-callatostatin 2, respectively.

Positively charged arginine side chains are located on the surface of the folded neuropeptide structures (Figures 4 and 5). They obviously play an important role as a centre of the electrostatic stabilization at the ligand–receptor binding.

The low-energy conformations of the second group consist of three helically packed segments, separated by an arginine residue in positions 7, 11 for Leu-callatostatin 1 and in positions 5, 9 for Leu-callatostatin 2. The following conclusions were made from the calculation experiment:

(i) helically packed *C*-terminal pentapeptides are identical for the all callatostatins spatial organization, so they are members of a family, possible originating from a single gene;

(ii) positively charged amino acid side chains may be considered as the centre of the electrostatic stabilization at the ligand–receptor binding.

REFERENCES

1. Duve H, Johnsen AH, Scott AG, Yu CG, Yagi KJ, Tobe SS, Thorpe A. Callatostatins: neuropeptides from the blowfly *Calliphora vomitoria* with sequence homology to cockroach allatostatins. *Proc. Natl Acad. Sci. USA* 1993; **90**: 2456–2460.
2. Duve H, Johnsen AH, Maestro JL, Scott AG, East PD, Thorpe A. Identification of the dipteran Leu-callatostatin peptide family: the pattern of precursor processing revealed by isolation studies in *Calliphora vomitoria*. *Regul. Pept.* 1996; **67**: 11–19.
3. Duve H, Thorpe A. Distribution and functional significance of Leu-callatostatins in the blowfly *Calliphora vomitoria*. *Cell Tissue Res.* 1994; **276**: 367–379.
4. Woodhead AP, Stay B, Seidel SL, Khan MA, Tobe SS. Primary structure of four allatostatins: neuropeptide inhibitors of juvenile hormone synthesis. *Proc. Natl Acad. Sci. USA* 1989; **86**: 5997–6001.
5. Pratt GE, Farnsworth DE, Siegel NR, Fok KF, Feyereisen R. Identification of an allatostatin from adult *Diploptera punctata*. *Biochem. Biophys. Res. Commun.* 1989; **163**: 1243–1247.
6. Pratt GE, Farnsworth DE, Fok KF, Siegel NR, Mc Cormack AL, Shabanowitz J, Hunt DF, Feyereisen R. Identity of a second type of allatostatin from cockroach brains: an octadecapeptide amide with a tyrosine-rich address sequence. *Proc. Natl Acad. Sci. USA* 1991; **88**: 2412–2416.
7. Donly BC, Ding Q, Tobe SS, Bendena WG. Molecular cloning of the gene for the allatostatin family of neuropeptides from the cockroach *Diploptera punctata*. *Proc. Natl Acad. Sci. USA* 1993; **90**: 8807–8811.
8. Duve H, Johnsen AH, Scott AG, East PD, Thorpe A. [Hyp^{3}]Met-callatostatin. Identification and biological properties of a novel neuropeptide from the blowfly *Calliphora vomitoria*. *J. Biol. Chem.* 1994; **269**: 21 059–21 066.
9. Duve H, Johnsen AH, Scott AG, Thorpe A. Isolation, identification and functional significance of [Hyp^{2}]Met-callatostatin and des Gly-Pro Met-callatostatin, two further post-translational modifications of the blowfly neuropeptide Met-callatostatin. *Regul. Pept.* 1995; **57**: 237–245.
10. Kivirikko KI, Myllyla R, Pihlajaniemi T. In *Post-translational Modifications of Proteins*, Harding JJ, Grabbe MJC (eds). CRC Press: Boca Raton, FL, 1992; 1–51.
11. Momany FA, McGuire RF, Burgess AW, Scheraga HA. Energy parameters in polypeptides: VII. Geometric parameters partial atomic charges, nonbonded

interaction for naturally occurring amino acid. *J. Phys. Chem.* 1975; **79**: 2361–2381.
12. Popov EM. *The Structural Organization of Proteins* (in Russian). Nauka: Moscow, 1989.
13. Cantor CR, Schimmel PP. *Biophysical Chemistry, Part 1, The Conformation of Biological Macromolecules*. W.H.Freeman: San Francisco, CA, 1980.
14. IUPAC-IUB Commission on Biochemical Nomenclature Abbreviations and symbols for description of conformation of polypeptide chains. *Pure Appl. Chem.* 1974; **40**: 291–308.
15. Godjaev NM, Maksumov IS, Ismailova LI. The program for semiempirical calculation of conformations of the molecular complexes. *J. Chem. Struct.* (in Russian) 1983; **24**: 147–148.
16. Pratt GE, Farnsworth DE, Fok KF, Siegel NR, McCormack AL, Shabanowitz J, Hunt DF, Feyereisen R. Identity of a second type of allatostatin from cockroach brain: an octadecapeptide amide with a tyrosine-rich address sequence. *Proc. Natl Acad. Sci. USA* 1991; **88**: 2412–2416.

