Solution Structure Determination of Four *Diploptera punctata* Allatostatins by NMR Spectroscopy and Molecular Modeling

Heejung Yang and Yoongho Lim*

*Bio/Molecular Informatics Center, Konkuk University, Seoul 143-701, Korea. *E-mail: yoongho@konkuk.ac.kr*

Received March 18, 2005

Key Words : NMR, Allatostatins, *Diploptera punctata*, Peptide, Molecular modeling

In 1989 a peptide showing an inhibitory effect against biosynthesis of juvenile hormone III (allatostatic effect) was found in cockroach *Diploptera punctata*. Its sequence was determined and named allatostatin. It has an amidated C-terminal.[1] Four years later its precursor polypeptide was found, which was composed of 370 amino acids (NCBI access number P12764).[2] Since the first discovery of allatostatin, three more analogues were known and named allatostatin I, II, III, and IV.[3] They have a common C-terminal sequence, Tyr-Xaa-Phe-Gly-Leu-NH_2. Based on the fact that there are thirteen common sequences in allatostatin precursor polypeptide, allatostatins of the same number were expected to be found. The precursor peptide isolated from cockroach *D. punctata* and the primary sequences of thirteen allatostatins are shown in Figure 1.[2] They were named according to the order of positions placed in their precursor. Allatostatin I, II, III, and IV were renamed 7, 9, 8, and 5, respectively.

The 50% effective doses against biosynthesis of juvenile hormone III of *Diploptera punctata* allatostatin 7 (Dp-AST7), Dp-AST9, Dp-AST8, and Dp-AST5 found in the early study are 4.1×10^{-11}, 7.2×10^{-9}, 9.4×10^{-9}, and 1.6×10^{-10} M, respectively.[4] The numbers of residues of Dp-AST7, Dp-AST9, Dp-AST8, and Dp-AST5 are 13, 10, 9, and 8, respectively. Therefore, there is no relationship between the number of residues and the effective doses. To provide basic information for the allatostatic effect, authors carried out the structural study of four *D. punctata* allatostatins using Nuclear Magnetic Resonance (NMR) spectroscopy and molecular modeling.

The NMR data of Dp-AST7 obtained in both aqueous solution (90%H_2O/10%D_2O) and TFE/water binary solution (50%trifluoroethanol-d_3/50%H_2O) were assigned based on the standard procedures.[5] Even though all of the backbone amide cross peaks were not observed in COSY and TOCSY, the chemical shifts of the residues could be assigned from the interpretation of the cross peaks between amide protons and alpha protons. In order to confirm the assignments, the sequential walk of the same region was carried out in the NOESY spectrum in TFE/water binary solution (Fig. 2). Assignments of NMR data of Dp-AST7 in both aqueous solution and TFE/water solution are listed in Table 1. While Dp-AST7 in aqueous solution did not show meaningful nOe

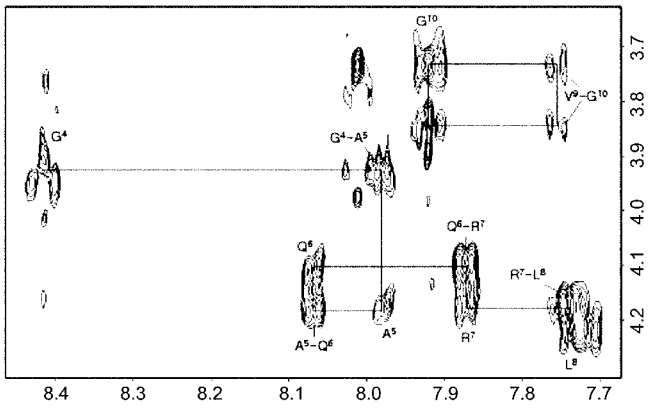

Figure 2. The sequential walk of the cross peaks between amide protons and alpha protons was carried out in the NOESY spectrum of Dp-AST7 in TFE/water binary solution.

```
1       msgprtcfcl psalvlvlls lstsalgtap epsgvheesp agggtdllph pedlsasdnp
61      dlefvkrlyd fglgkraysy vseykrlpvy nfglgkrskm ygfglgkrdg rmysfglgkr
121     dydyygeede ddqqaigded ieesdvgdlm dkrdrlysfg lgkrarpysf glgkrapsga
181     qrlygfglgk rggslysfgl gkrgdgrlya fglgkrpvns grssgsrfnf glgkrsddid
241     freleekfae dkrypqehrf sfglgkreve pseleavrne ekdnssvhdk knntndmhsg
301     erikrslhyp fgirklessy dlnsasslns eendditpee fsrmvrrpfn fglgkripmy
361     dfgigkrser
```

Figure 1. The primary sequence of *D. punctata* allatostatin precursor polypeptide.[2] (The boxes denote allatostatin peptides.)

Table 1. ^{1}H chemical shifts (ppm) of Dp-AST7

Medium	Residue	Chemical shifts of ^{1}H/ppm			
		NH	CαH	CβH	CγH and Others
TFE[a]	Ala[1]	–[c]	4.25	1.50	
Water[b]		–[c]	4.30	1.46	
TFE	Pro[2]		4.52	1.92, 2.29	2.00; Cδ*H* 3.54, 3.66
Water			4.48	1.88, 2.28	1.98; Cδ*H* 3.55, 3.66
TFE	Ser[3]	8.33	4.43	3.87, 3.95	
Water		8.49	4.39	3.84	
TFE	Gly[4]	8.41	3.90, 3.94		
Water		8.42	3.91		
TFE	Ala[5]	7.97	4.18	1.38	
Water		8.12	4.23	1.32	
TFE	Gln[6]	8.07	4.10	2.05	2.35; δNH_2 6.61, 7.28
Water		8.24	4.19	1.92, 2.00	2.29; δNH_2 6.79, 7.44
TFE	Arg[7]	7.87	4.18	1.78	1.58; Cδ*H* 3.14; δ*NH* 7.07
Water		8.20	4.19	1.66	1.47; Cδ*H* 3.08; δ*NH* 7.05
TFE	Leu[8]	7.74	4.23	1.43	1.53; Cδ*H* 0.76, 0.83
Water		8.09	4.25	1.38	1.46; Cδ*H* 0.75, 0.81
TFE	Tyr[9]	7.71	4.44	2.90, 3.03	Cδ*H* 7.06; Cε*H* 6.77
Water		8.07	4.48	2.85, 2.97	Cδ*H* 7.04; Cε*H* 6.76
TFE	Gly[10]	7.92	3.73, 3.85		
Water		8.16	3.70, 3.82		
TFE	Phe[11]	7.75	4.51	3.03, 3.14	CδH 7.20; Cε*H*[c]; Cζ*H*[c]
Water		8.00	4.51	2.98, 3.10	CδH 7.19; Cε*H* 7.31; Cζ*H* 7.24
TFE	Gly[12]	8.01	3.74, 3.92		
Water		8.32	3.76, 3.85		
TFE	Leu[13]	7.60	4.30	1.57	1.57; Cδ*H* 0.83, 0.87
Water		7.95	4.26	1.62	1.54; Cδ*H* 0.80, 0.84
TFE	NH				6.71, 7.28
Water					6.98, 7.54

[a]50% trifluoroethanol-d_3/50% H_2O; [b]90% H_2O/10% D_2O; [c]no data

peaks, several meaningful nOe cross peaks were observed in TFE/water solution. These results suggest that the flexible structure of Dp-AST7 in aqueous solution may be stabilized in TFE/water binary solution providing the environments of the biological membrane. Based on the integration of nOe cross peaks observed in TFE/water solution, the structure of Dp-AST7 was obtained. Simulated annealing was applied to obtain the refined structure of Dp-AST7 using InsightII software. The best 18 conformers were collected and superimposed as shown in Figure 3. The root mean-square deviation (RMSD) of 18 conformers was 0.98 Å. The conformer with the lowest total energy was validated using PROCHECK. The statistical analysis of Ramachandran plot showed that 42.9% of the residues are in the most favored regions, 28.6% in additional allowed regions, 28.6% in generously allowed regions, and 0% in disallowed regions. The structure obtained in the binary solution has two hinges at residues, Ala5, Gln6, and Tyr9.

The structure of Dp-AST7 without nOe constraints was calculated based on the molecular modeling method. The peptide was generated using InsightII software and was embedded in a 5 Å shell of 287 water molecules to imitate

Figure 3. Superimposition of the best 18 refined structures of Dp-AST7 obtained from NMR experiments carried out in TFE/water solution.

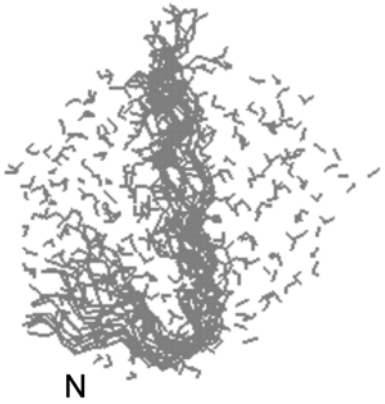

Figure 4. Superimposition of the 20 conformers of Dp-AST7 obtained from molecular modeling calculation. Small molecules show water molecules.

aqueous solvent conditions. Among 200 conformers obtained from molecular dynamics, 20 conformers showing the low total energy were selected and they were superimposed (Fig. 4). Their RMSD value was 1.72 Å. The increment of RMSD values from 0.98 Å obtained from the nOe data to 1.72 Å obtained from the MD calculation gives information that the structure in aqueous solution is flexible. This result agrees with the suggestion that Dp-AST7 structure obtained by NMR in aqueous solution is flexible.

The NMR data of Dp-AST9, Dp-AST8, and Dp-AST5 in both aqueous solution and TFE/water binary solution were assigned based on the standard procedures.[5] Their assignments are listed in Table 2, Table 3, and Table 4, respectively. Their NOESY experiments were carried in both aqueous solution and TFE/water binary solution. Since Dp-AST7 composed of 13 amino acids had a random coil structure in aqueous solution, these three Dp-ASTs composed of shorter residues than that of Dp-AST7 were expected to have flexible structures in aqueous solution.

Table 2. ^{1}H chemical shifts (ppm) of Dp-AST 9

Medium	Residue	Chemical shift			
		NH	CαH	CβH	CγH and Others
TFE[a]	Gly[1]	–[c]	–[c]		
Water[b]		–[c]	–[c]		
TFE	Asp[2]	8.44	4.73	2.81	
Water		8.64	–[c]	2.78	
TFE	Gly[3]	8.45	3.90		
Water		8.51	3.91		
TFE	Arg[4]	8.09	4.20	1.75	1.61; CδH 3.14; δNH 7.38
Water		8.11	4.23	1.71	1.52; CδH 3.12; δNH 7.16
TFE	Leu[5]	7.77	4.22	1.51	1.51; CδH 0.80, 0.88
Water		8.12	4.25	1.39	1.47; CδH 0.78, 0.85
TFE	Tyr[6]	7.56	4.42	2.87, 2.98	CδH 7.05; CεH 6.78
Water		8.02	4.48	2.82, 2.94	CδH 7.05; CεH 6.78
TFE	Ala[7]	7.69	4.12	1.22	
Water		8.00	4.20	1.21	
TFE	Phe[8]	7.60	4.47	3.02, 3.17	CδH[c]; CεH[c]; CζH[c]
Water		7.93	4.51	3.01, 3.12	CδH 7.23; CεH 7.33; CζH 7.28
TFE	Gly[9]	7.88	3.78, 3.92		
Water		8.18	3.80, 3.89		
TFE	Leu[10]	7.59	4.31	1.60	1.60; CδH 0.84, 0.89
Water		7.96	4.28	1.59	1.57; CδH 0.83, 0.89
TFE	NH				6.66, 7.25
Water					6.98, 7.54

[a]50% trifluoroethanol-d_3/50% H_2O; [b]90% H_2O/10% D_2O; [c]no data

Table 3. ^{1}H chemical shifts (ppm) of Dp-AST 8

Medium	Residue	Chemical shift			
		NH	CαH	CβH	CγH and Others
TFE[a]	Gly[1]	–[c]	–[c]		
Water[b]		–[c]	–[c]		
TFE	Gly[2]	8.39	4.01		
Water		8.08	3.66, 3.71		
TFE	Ser[3]	8.05	4.48	3.77, 3.85	
Water		8.17	4.30	3.66	
TFE	Leu[4]	8.07	4.21	1.42, 1.46	1.50; CδH 0.79, 0.84
Water		8.14	4.10	1.41	1.41; CδH 0.66, 0.72
TFE	Tyr[5]	7.69	4.45	2.83, 2.98	CδH 7.03; CεH 6.78
Water		7.94	4.42	2.70, 2.83	CδH 6.92; CεH 6.66
TFE	Ser[6]	7.72	4.31	3.70, 3.78	
Water		7.84	4.20	3.58	
TFE	Phe[7]	7.74	4.51	3.03, 3.15	CδH 7.21; CεH 7.30; CζH 7.22
Water		7.91	4.43	2.89, 3.01	CδH 7.12; CεH 7.22; CζH 7.16
TFE	Gly[8]	7.95	3.77, 3.92		
Water		8.08	3.75, 3.80		
TFE	Leu[9]	7.61	4.31	1.58, 1.63	1.59; CδH 0.84, 0.89
Water		7.86	4.16	1.46	1.46; CδH 0.72, 0.77
TFE	NH				6.71, 7.27
Water					6.87, 7.42

[a]50% trifluoroethanol-d_3/50% H_2O; [b]90% H_2O/10% D_2O; [c]no data

Table 4. ^{1}H chemical shifts (ppm) of Dp-AST 5

Medium	Residue	Chemical shift			
		NH	CαH	CβH	CγH and Others
TFE[a]	Asp[1]	–[c]	4.28	2.86, 2.94	–[c]
Water[b]		–[c]	4.25	2.76, 2.88	–[c]
TFE	Arg[2]	8.54	4.31	1.74	1.56; CδH 3.13; δNH 7.03
Water		8.62	4.27	1.69	1.49; CδH 3.10; δNH 7.04
TFE	Leu[3]	7.87	4.29	1.50	1.50; CδH 0.82, 0.88
Water		8.21	4.26	1.40	1.48; CδH 0.78, 0.85
TFE	Tyr[4]	7.76	4.49	2.88, 2.95	CδH 7.04; CεH 6.78
Water		8.17	4.50	2.83, 2.92	CδH 7.06; CεH 6.76
TFE	Ser[5]	7.75	4.33	3.69, 3.77	
Water		8.00	4.32	3.69	
TFE	Phe[6]	7.74	4.53	3.04, 3.14	CδH[c]; CεH[c]; CζH[c]
Water		8.04	4.55	3.01, 3.12	CδH 7.23; CεH 7.33; CζH 7.28
TFE	Gly[7]	7.94	3.75, 3.94		
Water		8.21	3.79, 3.88		
TFE	Leu[8]	7.61	4.32	1.60	1.60; CδH 0.85, 0.90
Water		7.99	4.27	1.58	1.58; CδH 0.82, 0.89
TFE	NH				6.68, 7.27
Water					6.98, 7.54

[a]50% trifluoroethanol-d_3/50% H_2O; [b]90% H_2O/10% D_2O; [c]no data

Like the NOESY data of Dp-AST7, any meaningful nOe cross peaks were not observed in the NOESY experiments of three Dp-ASTs. Therefore, their structures are considered flexible in aqueous solution. In addition, any meaningful nOe cross peaks were not observed in their NOESY spectra obtained in TFE/water binary solution too. As a result, the NMR experiments of Dp-AST9, Dp-AST8, and Dp-AST5 showed that they are flexible in both aqueous solution and TFE/water binary solution. The molecular modeling calculations were carried out in three peptides. While the RMSD of the refined structures of Dp-AST7 determined from nOe experiments was less than 1 Å, the RMSD values

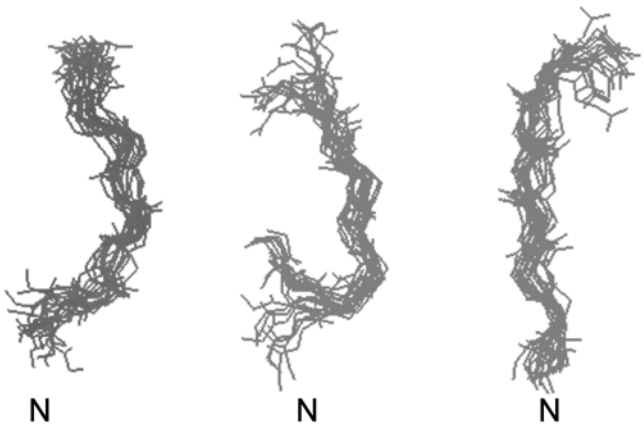

Figure 5. The superimposed structures of the 20 conformers of Dp-AST9 (left), Dp-AST8 (center), and Dp-AST5 (right) obtained from molecular modeling calculation carried out in an aqueous condition.

of the superimposed structures of Dp-AST9, Dp-AST8, and Dp-AST5 were 1.73 Å, 1.55 Å, and 1.41 Å, respectively. As shown in Figure 5, the structures of Dp-AST9, Dp-AST8, and Dp-AST5 are flexible.

Among six C-terminal residues, Leu-Tyr-Xaa-Phe-Gly-Leu-NH_2, in four Dp-ASTs, only Xaa is different with each other. Dp-AST7 in TFE/water solution shows the hinged structure at Tyr9 which disappears in the structure obtained from the MD calculation carried out in an aqueous condition. Even though the other three peptides include Tyr at the same position, they do not show any fixed structures in both aqueous solution and TFE/water binary solution. As a result, three more residues contained in the N-terminal of Dp-AST7 may contribute to making the structure. As mentioned before, allatostatins contain a common C-terminal sequence, Tyr-Xaa-Phe-Gly-Leu-NH_2. Allatostatin analogues incorporating pseudopeptide moieties such as cyclopropylalanine and amino indane-2-carboxylic acid were synthesized and their conformational studies were carried out by Nachman *et al*.,[6,7] where pseudopeptide moieties were inserted in the C-terminal sequence. The presence of turn was observed in aqueous solution. Their allatostatic effects were similar to those of natural allatostains used for this work. In conclusion, allatostatic effects are not caused by the presence of firm structure.

Materials and Methods

Sample preparation. Four peptides were synthesized by American Peptides Co. INC. (Sunnyvale, CA, USA). The primary sequence of Dp-AST7 is Ala-Pro-Ser-Gly-Ala-Gln-Arg-Leu-Tyr-Gly-Phe-Gly-Leu-NH_2 whose molecular weight is 1337.4 Da. The mass data obtained by MALDI/TOF was 1337.36 and the purity of the peptide determined by Reversed Phase (RP) –HPLC was 95.47%. The primary sequence of Dp-AST9 is Gly-Asp-Gly-Arg-Leu-Tyr-Ala-Phe-Gly-Leu-NH_2 whose molecular weight is 1067.5 Da. The experimental mass was 1066.99 and the purity was 99.28%. The primary sequence of Dp-AST8 is Gly-Gly-Ser-Leu-Tyr-Ser-Phe-Gly-Leu-NH_2 whose molecular weight is 899.3 Da. The experimental mass was 899.49 and the purity was 99.39%. The primary sequence of Dp-AST5 is Asp-Arg-Leu-Tyr-Ser-Phe-Gly-Leu-NH_2 whose molecular weight is 969.1 Da. The experimental mass was 969.19 and the purity was 100.0%.

NMR spectroscopy. All NMR measurements were performed on a Bruker Avance 400 spectrometer system (9.4 T, Karlsruhe, Germany) at 298K.[8] The NMR spectra of ^{1}H-NMR, correlated spectroscopy (COSY), total correlated spectroscopy (TOCSY), and nuclear Overhauser and exchange spectroscopy (NOESY) were collected in 90%H_2O/10%D_2O or 50%trifluoroethanol(TFE)-d_3/50%H_2O. The concentration of the samples was approximately 1 mM. For ^{1}H-NMR analysis, 128 transients were acquired with a 1 sec relaxation delay using 32 K data points. The 90° pulse was 9.7 μsec with a spectral width of 4,000 Hz. Two-dimensional spectra were acquired with 2,048 data points for t2 and 256 for t1 using time proportional phase increments except COSY experiments where magnitude mode was applied. NOESY experiments of allatostains were performed at the various mixing time of 50 msec, 75 msec, 100 msec, 200 msec, 300 msec, 400 msec, 500 msec, 600 msec, 700 msec, 1,000 msec, and 1,500 msec. The mixing times for TOCSY with MLEV17 spin-lock pulse program of Dp-AST7, Dp-AST9, Dp-AST8, and Dp-AST5 were 126 msec, 126 msec, 100 msec, and 100 msec, respectively. In all experiments water peaks were suppressed by presaturation. Prior to fourier transformation, zero filling of 2 K and sine squared bell window function were applied using XWIN-NMR (Bruker, Karlsruhe, Germany).

Structure calculations. Data analysis was carried out using Sparky.[9] The molecular modeling calculations were carried out on an O2 R12,000 Silicon Graphics workstation. The initial structures of the molecules were generated using InsightII (Accelrys, San Diego, CA, USA.). The forcefield used for molecular dynamics (MD) and energy minimization was consistent valence force field (cvff) provided by Accelrys. The simulated annealing followed the method performed by Taura *et al*.[10] The molecules were subjected to energy minimization by Discover module included in InsightII. Steepest descents were carried out until maximum derivative of 1.0 kcal/molÅ, and conjugate gradients were followed until maximum derivative of 0.1 kcal/molÅ. After energy minimization, MD was performed at 300 K, 1 atm for 500 psec with 1 fsec each step. The output conformers were collected at every 2.5 psec and 200 conformers were saved in the history file. The energy profile was analyzed using the Analysis module included in InsightII. Among 200 conformers, 20 conformers showing the low total energy were selected and they were superimposed. Of these, the conformer with the lowest energy was chosen and PROCHECK was applied for a statistical evaluation.

Acknowledgement. This work was supported by Korea Research Foundation Grant (KRF-2001-041-G00029).

References

1. Pratt, G. E.; Farnsworth, D. E.; Siegel, N. R.; Fok, K. F.; Feyereisen, R. *Biochem. Biophy. Res. Comm.* **1989**, *163*, 1243.
2. Donly, B. C.; Ding, Q.; Tobe, S. S. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 8807.
3. Pratt, G. E.; Farnsworth, D. E.; Fok, K. F.; Siegel, N. R.; McCormack, A. L.; Shabanowitz, J.; Hunt, D. F.; Feyereisen, R. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2412.
4. Bendena, W. G.; Garside, C. S.; Yu, C. G.; Tobe, S. S. *Annals. NY Acad. Sci.* **1997**, *814*, 53.
5. Lee, S.; Kim, D.; Kim, H.; Lee, Y.; Won, H. *Bull. Korean Chem. Soc.* **2004**, *25*, 1845.
6. Nachman, R. J.; Moyna, G.; Williams, H. J.; Tobe, S. S.; Scott, A. I. *Bioorg. Med. Chem.* **1998**, *6*, 1379.
7. Nachman, R. J.; Garside, C. S.; Tobe, S. S. *Peptides* **1999**, *20*, 23.
8. Park, H.; Kim, E.; Kang, M. R.; Chung, I. K.; Cheong, C.; Lee, W. *Bull. Korean Chem. Soc.* **2004**, *25*, 1559.
9. Goddard, T. D.; Kneller, D. G. *Sparky 3*; University of California, San Francisco, 2004.
10. Taura, S.; Coutinho, E.; Srivastava, S. *Magn. Reson. Chem.* **2002**, *40*, 211.