Structure of the Insect Cytokine Peptide Plasmatocyte-spreading Peptide 1 from Pseudoplusia includens

(Received for publication, November 30, 1998) Brian F. Volkman‡, Mark E. Anderson‡, Kevin D. Clark‡, Yoichi Hayakawa‡, Michael R. Strand‡, and John L. Markley‡

From the ‡National Magnetic Resonance Facility at Madison, Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706-1544, the ‡Department of Entomology, University of Wisconsin-Madison, Madison, Wisconsin 53706, and the †Biochemical Laboratory, Institute of Low Temperature Science, Hokkaido University, Sapporo 060-0819, Japan

The structure of the recently identified plasmatocyte spreading peptide from the moth Pseudoplusia includens (PSP1) has been determined by NMR spectroscopy. This novel insect cytokine consists of 23 amino acid residues and a single disulfide bond. Torsion angle dynamics calculations utilizing a total of 337 distance constraints yielded an ensemble of 30 structures with an average backbone root mean square deviation for residues 7–22 of 0.18 Å from the mean structure. The structure consists of a disordered N-terminal region and a well defined core that is stabilized by numerous hydrophobic interactions and a short О-hairpin. Structural comparisons confirm that PSP1 adopts an epidermal growth factor (EGF)-like fold with close similarity to the C-terminal subdomain of EGF-like module 5 of human thrombomodulin. The combination of the three-dimensional structure of PSP1 and the extensive literature on EGF-receptor interactions should accelerate the process of identifying the specific residues responsible for receptor binding activity of this family of immunoregulatory peptides.

The insect immune response to parasites and pathogens involves the action of different classes of blood cells (hemocytes), which adhere to and spread across the surface of foreign targets. In phylogenetically advanced insects like the Lepidoptera (moths and butterflies), plasmatocytes and granular cells are the two most important classes of hemocytes involved in cellular defense responses like encapsulation and clotting (3). However, relatively little is known about the factors mediating the movement and action of these immune cells (3). The identification of a peptide that induces the adhesion and spreading of plasmatocytes on foreign surfaces was recently reported (1). This plasmatocyte-spreading peptide (PSP1) of Pseudoplusia includens is expressed as a preproprotein of 142 residues (2), but the mature peptide consists of only the C-terminal 23 amino acids containing one disulfide bond. Although the existence of cytokine-like factors to regulate hemocyte activity has been proposed previously (3), PSP1 is among the first soluble mediators of the insect cellular immune response to be purified and functionally characterized.

The amino acid sequence of PSP1 shows no significant similarity to vertebrate cytokines, but is related to two classes of insect peptides previously identified (1). As shown in Fig. 1, the growth-blocking peptide (GBP) of Pseudaletia separata and the paralytic peptides from Manduca sexta, Heliothis virescens, and Spodoptera exigua have high (>70%) sequence identity with PSP1, including the two cysteine residues and a highly conserved glycine. High sequence variability within this group is found only at positions 4 and 8. Some similarity between the sequences of GBP and human epidermal growth factor has been recently suggested (4), and this would extend to PSP1 and the paralytic peptides.

Chemically synthesized PSP1 was shown to induce an identical plasmatocyte spreading and adhesion response to the peptide purified from the plasma of P. includens larvae and is therefore suitable for detailed structure-function analysis (1). Here we report the three-dimensional structure of this novel insect cytokine determined by NMR spectroscopy and torsion angle dynamics calculations. The structure of PSP1 shows clear homology to the C-terminal subdomain of the EGF domain family. The backbone conformation of PSP1 is very similar to that of the fifth EGF-like domain of the human anticoagulant protein thrombomodulin (hTM5), despite sharing sequence identity at only four positions, including the two cysteines of the conserved disulfide.

EXPERIMENTAL PROCEDURES

Sample Preparation—PSP1 was synthesized using tert-butyloxycarbonyl chemistry and purified by reversed-phase high performance liquid chromatography. Its identity was confirmed by matrix-assisted laser desorption/ionization time of flight mass spectrometry and NMR. The NMR sample consisted of 2 mg of pure lyophilized peptide dissolved in 500 μl of a buffer containing 90% H₂O/10% D₂O, 20 mM sodium phosphate, pH 6.0.

NMR Spectroscopy—All NMR spectra were recorded at 10°C on Bruker DMX700 and DMX500 spectrometers equipped with triple-resonance (1H/13C/15N) probes and with Z− or three-axis pulsed field gradient capabilities. Quadrature detection in the indirectly detected dimensions was obtained with the States-TPPI method (5). Water suppression was achieved using a watergate sequence with a 3-9-19 selective inversion pulse (6). A two-dimensional total correlation spectroscopy spectrum with an isotropic mixing period of 67 ms was acquired at 499.84 MHz, with spectral widths of 6009.6 Hz in both dimensions and a root mean square deviation.

* This work was supported by National Institutes of Health Grant AI32917 and U.S. Department of Agriculture Grant 95-37302-1811. Equipment in the National Magnetic Resonance Facility at Madison, Department of Biochemistry, University of Wisconsin-Madison was purchased with funds from the University of Wisconsin, National Science Foundation Biological Instrumentation Program Grant DMB-8415048, National Institutes of Health Biomedical Research Technology Program Grant RR02301, National Institutes of Health Shared Instrumentation Program Grant RR02781, and funds from the U.S. Department of Agriculture. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 1b1v and 1b5n) have been deposited in the Protein Data Bank, Brookhaven National Laboratory, Upton, NY.

‡ To whom correspondence should be addressed: National Magnetic Resonance Facility at Madison, Dept. of Biochemistry, University of Wisconsin-Madison, 433 Babcock Dr., Madison, WI 53706-1544. Tel: 608-262-0459; Fax: 608-262-3759; E-mail: volkman@nmrfam.wisc.edu.

† The abbreviations used are: PSP1, plasmatocyte-spreading peptide 1; GBP, growth blocking peptide; EGF, epidermal growth factor; hEGF, human EGF; hTM5, fifth EGF-like domain of human thrombomodulin; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; RMSD, root mean square deviation.
Structure of Insect Cytokine PSP1

RESULTS AND DISCUSSION

Structure Determination—Two-dimensional $^1$H NMR spectroscopy was used to obtain complete resonance assignments for chemically synthesized plasmatocyte-spreading peptide 1 (PSP1). As shown in Fig. 2, the chemical shift dispersion and large number of NOEs observed indicate the presence of well defined secondary and tertiary structure. A total of 699 cross-peaks were assigned in the two-dimensional NOESY spectrum of PSP1, from which 333 nontrivial NOE distance constraints were produced after discarding duplicate and structurally redundant NOEs. Two additional pairs of upper and lower bound distance constraints were included between the $S$ and $C^\beta$ atoms of Cys$^1$ and Cys$^{19}$ to produce novel disulfide bond geometry ($2.0 \, \text{Å} < d_{S-S} < 2.1 \, \text{Å}, 3.72 \, \text{Å} < d_{C^\beta-C^\beta} < 3.99 \, \text{Å}$). A summary of all 337 distance constraints by class is given in Table I, and the distribution of constraints by residue is shown in Fig. 3. Stereospecific assignments were obtained from analysis of structures and NOE intensities for the $H^\beta$ protons of Gly$^7$, the $H^\beta$ protons of Tyr$^{11}$ and Pro$^{21}$, and the $H^\gamma$ protons of Pro$^{21}$.

The ensemble of 30 PSP1 conformers, which resulted from the final stage of refinement, is shown in Fig. 4. The low average target function (0.12 ± 0.02 Å$^2$) indicates good agreement between the ensemble of calculated conformers and the experimental constraints. Superposition of the well ordered residues (7–22) of this ensemble onto the corresponding residues of the mean structure produced an average RMSD to the mean of 0.18 Å for the backbone atoms and 0.84 Å for all nonhydrogen atoms. After minimization of the mean structure in DYANA to a final target function of 0.13, the average RMSD for residues 7–22 became 0.23 Å for the backbone and 0.98 Å for the heavy atoms. A summary of residual violations and RMSDs is given in Table I.

The tertiary structure of PSP1 is stabilized by a combination of the covalent disulfide linkage, hydrogen bonding within the $\beta$-hairpin structure, and hydrophobic side chain packing. All 30 torsion angle dynamics conformers were consistent with a pair of hydrogen bonds involving the backbone amide and carbonyl groups of Met$^{15}$ and Lys$^{20}$. Evidence for two additional hydrogen bonds (Asp$^{16}$ HN–Thr$^{14}$ O$^\gamma$, Thr$^{14}$ HN–Arg$^{19}$ O) was observed in a subset of the family of structures. Within the ordered portion of the molecule, a subset of side chains forms a well defined hydrophobic core centered around the side chain of Tyr$^{11}$. This residue is conserved throughout the family of EGF-like domains as a phenylalanine or tyrosine (14), has NOE contacts to the side chains of Cys$^7$, Leu$^6$, Ala$^8$, Cys$^{19}$, Pro$^{21}$, and Phe$^{23}$.

As seen in Fig. 1, the sequences of PSP1, GBP from Pseudaleutia separata, and a paralytic peptide from M. sexta differ in...
Structure of Insect Cytokine PSP1

Table I

| Structural statistics for 30 PSP1 conformers |
|---------------------------------------------|
| (TAD), the ensemble of 30 DYANA torsion angle dynamics (TAD) conformers; (TAD), the average coordinates obtained from a least squares superposition of residues 7–22; (TAD)_min, the structure obtained by applying 1000 steps of conjugate gradient minimization to the mean structure, (TAD). |

| NOE constraint totals by type | Number |
|------------------------------|--------|
| Long                         | 92     |
| Medium                       | 73     |
| Short                        | 109    |
| Intraresidue                 | 59     |
| Disulfide                    | 4      |
| Total                        | 337    |
| Constraints/residue          | 14.6   |

| DYANA parameters |
|------------------|
| Target function (Å^2) | 0.12 ± 0.02 | 0.13 |
| Upper limit violations |
| Number > 0.1 Å | 4 ± 1 | 3 |
| Sum of violations (Å) | 1.3 ± 0.1 | 1.2 |
| Maximum violation (Å) | 0.19 ± 0.00 | 0.19 |
| Van der Waals' violations |
| Number > 0.1 Å | 0 ± 0 | 0 |
| Sum of violations (Å) | 0.2 ± 0.1 | 0.3 |
| Maximum violation (Å) | 0.05 ± 0.01 | 0.05 |

| Average atomic RMSD (Å) |
|-------------------------|
| N, C^\text{a}, C' | 0.18 ± 0.08 | 0.81 ± 0.20 |
| All nonhydrogen         |
| (TAD) vs. (TAD) | 0.23 ± 0.10 | 0.94 ± 0.24 |
| (TAD) pairwise | 0.27 ± 0.11 | 1.15 ± 0.30 |

Note: Atomic RMSDs were calculated by superposition of either backbone N, C^\text{a}, C' atoms or all nonhydrogen atoms for residues 7–22.

Fig. 3. Distribution by residue of NOE constraints used in the structure determination of PSP1. For each residue, the number of intraresidue, sequential, medium-range (1 < |i − j| < 4) and long range NOEs are indicated by bars of white, gray, blue, red, and yellow, respectively.

very few positions, and all may be expected to adopt the same three-dimensional structures. Although the structures of these related peptides have not been reported, preliminary analysis of the NOESY spectrum of GBP indicates that its structure will be very similar to that reported here for PSP1.\(^3\)\(^4\) Consistent with this result are functional assays using both PSP1 and GBP that show no significant differences in activity.\(^4\)

Comparison with Epidermal Growth Factor Modules—EGF-like domains are peptides of approximately 40 residues in length, containing six highly conserved cysteine residues. Three loops, designated A, B, and C, are defined by the pattern of disulfide bonding. EGF modules are alternatively described in terms of two subdomains, each containing a \(\beta\)-hairpin structure. The N-terminal EGF-like subdomain consists of the first two disulfide-linked loops, and the smaller C-terminal subdomain consists of the third disulfide and the C loop.

The structures of human epidermal growth factor (hEGF) (15) and the hTM5 (16), which have been determined by NMR spectroscopy, share the general structural features of the family of EGF-like domains. The variability of the EGF family is illustrated by the differences between these two structures, including alternative pairing for the first two disulfides and a longer C loop in hTM5 than hEGF (16). Sequence identity between PSP1 and these representatives of the family of EGF-like domains is limited to the essential cysteine, glycine, and tyrosine residues of the C-terminal EGF-like subdomain (Fig. 1), but PSP1 has an extended C loop as found in hTM5. The backbone conformations of these three peptides are compared in Fig. 5. Except for the displacement of the \(\beta\)-turn in hEGF, resulting from its shorter C loop, the structures of all three molecules are quite similar. PSP1 and hTM5 align to a remarkably high degree, with an average backbone atomic RMSD of \(-1\) Å for the 13 residues within the disulfide loop.

Implications for Receptor Binding—PSP1 was isolated as the peptide responsible for spreading of insect cell plasmatocytes on foreign surfaces, an important process in cellular defense. Hence, it will be of interest to determine the cellular receptor for this peptide. Given the striking similarity between the structure of PSP1 and the C-terminal subdomain of EGF-like modules, it is tempting to speculate that the ligand-receptor interactions of PSP1 may be analogous to those of the array of EGF domain-receptor complexes that have been investigated (17).

Fig. 6 shows the minimized mean structure of PSP1, includ-
ing all side chains. The charged residues Arg, Asp, Arg, and Lys are clearly clustered on one side of the molecule, and the patch of hydrophobic residues surrounding the Tyr side chain are located on the opposing side. No consensus receptor-binding site has been identified yet for EGF receptor ligands, despite intense efforts stemming from the relevance of this system to the design of potential drugs for treating human cancers. However, the studies have determined that an arginine conserved in EGF receptor binding domains (Arg in hEGF) is essential for binding. Interestingly, Arg is the analogous residue in PSP1.

Significant effort has been spent investigating the details of thrombin recognition by thrombomodulin, including alanine-scanning mutagenesis, NMR spectroscopy, and x-ray crystallography of the complexes between small peptide fragments of hTM5 and thrombin (18–20). Although the distinction between residues critical for proper folding and those directly involved in binding may not be entirely clear from mutagenesis alone, a structural interpretation clearly implicates the C loop of hTM5, including the many acidic side chains on either side of Cys (Fig. 1). None of these residues has a direct correlation, but the cluster of charged side chains occurs in the same vicinity of the PSP1 structure. In the absence of other information on the activity of PSP1, this concentration of acidic and basic residues is an obvious starting point for investigations of the receptor-binding determinants of this new class of insect cytokine peptides.

Acknowledgments—We thank Dr. Milo Westler for assistance in acquiring and processing NMR data and Dr. Dorothee Kern for advice on using Garant and XEASY. We also thank Dr. Ed Mooberry and other members of the National Magnetic Resonance Facility at Madison staff for gracious help.

REFERENCES

1. Clark, K. D., Pech, L. L., and Strand, M. R. (1997) J. Biol. Chem. 272, 23440–23447
2. Clark, K. D., Witherell, A., and Strand, M. R. (1998) Biochem. Biophys. Res. Commun. 250, 194–199
3. Strand, M. R., and Pech, L. L. (1995) Annu. Rev. Entomol. 40, 31–56
4. Hayakawa, Y., and Ohnishi, A. (1998) Biochem. Biophys. Res. Commun. 250, 967–974
5. Marion, D., and Wuthrich, K. (1983) Biochem. Biophys. Res. Commun. 113, 967–974
6. Pietto, M., Saudek, V., and Sklenar, V. (1992) J. Biomol. NMR 2, 661–665
7. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
8. Bax, A., Ikura, M., Kay, L. E., and Zhu, G. (1991) J. Magn. Reson. 91, 174–178
9. Wuthrich, K. (1986) NMR of Proteins and Nucleic Acids, John Wiley & Sons, Inc., New York
10. Bartels, C., Billeter, M., Güntert, P., and Wuthrich, K. (1996) J. Biomol. NMR 7, 207–213
11. Bartels, C., Xia, T.-H., Billeter, M., Güntert, P., and Wuthrich, K. (1995) J. Biomol. NMR 5, 1–10
12. Güntert, P., Mumenthaler, C., and Wuthrich, K. (1997) J. Mol. Biol. 273, 283–298
13. Koradi, R., Billeter, M., and Wuthrich, K. (1996) J. Mol. Graph. 14, 51–55
14. Appella, E., Weber, I. T., and Blasi, F. (1988) FEBS Lett. 231, 1–4
15. Hombach, U., Harvey, T. S., Driscoll, P. C., and Campbell, I. C. (1992) J. Mol. Biol. 227, 271–282
16. Sampoli-Benitez, B. A., Hunter, M. J., Meineminger, D. P., and Komives, E. A. (1997) J. Mol. Biol. 273, 913–920
17. McNees, C., and Sykes, B. D. (1998) Biopolymers 43, 339–366
18. Nagashima, M., Lundh, E., Leonard, J. C., Morser, J., and Parkinson, J. F. (1993) J. Biol. Chem. 268, 2888–2892
19. Mathews, I. I., Padmanabhan, K. P., Tulinisky, A., and Sadler, J. E. (1994) Biochemistry 33, 13547–13552
20. Hrabal, R., Komives, E. A., and Ni, F. (1996) Protein Sci. 5, 195–203