COMMUNICATIONS

Benzophenone Photophore Flexibility and Proximity: Molecular and Crystal-State Structure of a Bpa-Containing Trichogin Dodecapeptide Analogue

Michele Saviano,*[a] Roberto Improta,[a] Ettore Benedetti, ^[a, b] Benedetta Carrozzini, ^[c] Giovanni Luca Cascarano, ^[c] Claude Didierjean, ^[d] Claudio Toniolo,^[e] and Marco Crisma^[e]

Exploitation of photolabeling compounds for obtaining covalent modifications at peptide ligand-protein receptor sites has attracted the attention of biochemists for about two decades.^[1] In particular, the introduction of the Bpa (para-benzoylphenylalanine) residue (Scheme 1),^[2] with a photoactivable aryl ketone

nOct-Aib-Gly-Leu-Aib-Gly-Gly-Leu-Aib-Gly-Ile-Lol (nOct: n-octanoyl; Lol: leucinol)

Trichogin GA IV

Scheme 1. Chemical structures of the α -amino acids Bpa and TOAC, and primary structure of the lipopeptaibol antibiotic trichogin GA IV.

side chain, has led to a number of investigations of synthetic peptide – protein interactions.^[1-10] The use of a benzophenone photoprobe has a series of distinct advantages: i) it is chemically more stable than other types of photoactivable reagents; ii) it can be manipulated in ambient light and can be activated in the

350 - 360 nm spectral region thus avoiding protein-damaging wavelengths; iii) it reacts preferentially with unreactive C-H bonds even in the presence of water and nucleophiles and in particular it shows a remarkable site specificity; iv) the efficiency of cross-linking via the photogenerated triplet state is high because only a small amount is lost to hydrolysis. In addition, useful and more specific characteristics of Bpa are its full compatibility with peptide-synthesis methodologies, including the solid-phase approach, and its ability to work with the E. coli translational machinery so that it can be incorporated into proteins with high efficiency.[10]

Beside this intermolecular application of Bpa directed at mapping the interphase and specifically amino acid-to-amino acid contact points between ligand and receptor, recently we have studied the intramolecular quenching of the Bpa photoexcited triplet state by a nitroxide free radical linked to peptide templates by time-resolved EPR with pulsed-laser excitation.[11, 12] The systems investigated have been 3 $_{10}$ - $^{[13, 14]}$ and mixed 3 $_{10}/\alpha$ helix forming peptides having in the amino acid sequence a Bpa residue along with a helicogenic, stable free radical 2,2,6,6 tetramethylpiperidine-1-oxyl-4-amino-4-carboxylic acid (TOAC) residue^[15] (Scheme 1) incorporated at different relative positions. Upon interaction with the excited triplet, the TOAC nitroxide radical spin sublevel populations assume values that differ from the Boltzmann equilibrium values.^[11, 12] This spin-polarization effect produces EPR emission lines whose time evolution reflects the triplet quenching rate. In particular, in a series of short peptides labeled with Bpa and TOAC at selected positions in the $3₁₀$ -helix, a radical-triplet interaction was observed in all cases. Interestingly, the EPR signal of the terminally protected peptide Boc-Bpa-Aib-Gly-Leu-Aib-(Gly)₂-Leu-TOAC-Gly-Ile-Leu-OMe (Boc, $tert$ -butyloxycarbonyl; Aib, α -aminoisobutyric acid), in which the Bpa and TOAC residues are separated by a large number of amino acids (including two helicogenic Aib residues), $[16, 17]$ exhibits the same features as those shown by the shortest peptides investigated, though of significantly lower intensity. This peptide is also the [Boc-Bpa0 TOAC8 Leu11-OMe] analogue of the membrane-active lipopeptaibol antibiotic trichogin GA IV^[18-25] (Scheme 1).

The ultimate goal of this novel EPR approach was to obtain information on the intramolecular distance between the two probes in bulk solution and, indirectly, on the type of secondary structure adopted by the peptide template under these conditions and to compare this to the preferred folding of the same compound in the membrane environment.

However, it is evident that, for meaningful information on the intermolecular regioselectivity of the ligand - receptor interactions and on the assessment of the intramolecular interprobe distances as well, the rigidity (or, at least, a very limited flexibility) of the photoprobe is crucial. It is also clear that, in the Bpa residue, this property is governed significantly by the amount of conformational space that can be explored by the side-chain - $CH₂$ - benzophenone moiety (namely, by the amino acid $\chi¹$ and χ^2 torsion angles).^[26] In this context, here we describe the 3D structure and the conformational behavior of the abovementioned trichogin analogue as determined by an X-ray diffraction analysis (part of this work has been reported in a

HEMBIOCHEM

preliminary form in ref. [27]). The molecular structures of the two independent molecules A and B in the asymmetric unit are illustrated in Figure 1.

Figure 1. Side view of the X-ray diffraction structures of the two crystallographically independent molecules A and B in the asymmetric unit of the [Boc-Bpa0, TOAC8, Leu11-OMe] trichogin GA IV analogue with residue numbering.

Helical molecules with right-handed screws are observed for both conformers A and B (Table 1). The general backbone conformational features of A and B differ only for some parameters. At the N terminus of molecule A, the carbonyl oxygen of the Boc protecting group (O0'A) is involved in a $1 \leftarrow 4$ intramolecular H-bond with the peptide nitrogen (N2A) of Gly2; this gives rise to a type-III C₁₀-ring structure (β -turn) $^{[28]}$ (Table 2). In addition, the O0'A atom accepts a H-bond from the peptide nitrogen (N3A) of Leu3; this leads to the formation of a C_{13} -ring structure (α -turn).^[29] Thus, the structure starts as a 3₁₀/ α -helix, but switches to a pure α -helix that consists of five consecutive α turns which include the H-bonds between Bpa0 (O0A) and Aib4 (N4A), Aib1 (O1A) and Gly5 (N5A), Gly2 (O2A) and Gly6 (N6A), Leu3 (O3A) and Leu7 (N7A), and Aib4 (O4A) and TOAC8 (N8A). The carbonyl O5A atom of Gly5 is not involved in the intramolecular H-bonding scheme, but rather in an intermolecular H-bond with the Ow_1 water molecule. This H-bond changes the peptide folding into a type-III β -turn (Gly6 O6A \leftarrow Gly9 N9A) followed by a type-I β -turn (Leu7 O7A \leftarrow Ile10 N10A). At the C terminus, a H-bonded α -turn further stabilizes the helical structure (Leu7 O7A \leftarrow Leu11 N11A). The carbonyl O7A atom, as discussed above for the O0'A atom, is a double acceptor of intramolecular H-bonding. The C-terminal Leu11 residue is in a semiextended conformation.

The backbone conformation of molecule B differs from that of molecule A for the intramolecular H-bonding network. At the N

terminus of molecule B, the carbonyl oxygen of the Boc group (O0'B) does not participate in either the intra- or the intermolecular H-bonding scheme. The carbonyl O0B and O1B atoms of Bpa0 and Aib1 residues are involved in two consecutive $1 \leftarrow 4$ intramolecular H-bonds with the peptide nitrogens (N3B and N4B) of Leu3 and Aib4, respectively. These conformations are a type-III (I) and a type-I β -turn, respectively. The peptide nitrogen N5B atom of Gly5 is not directly involved in the intramolecular H-bonding scheme. Rather, it donates a H-bond to the cocrystallized Ow₁ water molecule, which, in turn, is H-bonded to the O2B carbonyl oxygen atom of Gly2. As a result, the water molecule acts as a bridge between the N5B and O2B atoms (watermediated, hydrogen-bonded β -turn^[30]). Here too, as in molecule A, a perturbation of the regularity in the H-bonding scheme alters the peptide folding. Thus, the O3B atom is a double acceptor of intramolecular H-bonding, with formation of a type-III β -turn (Leu3 O3B \leftarrow Gly6 N6B) and an α -turn (Leu3 O3B \leftarrow Leu7 N7B). Then this 310/ α -helical structure switches to an α helix that consists of four consecutive α -turns, including the

H-bonds between Aib4 (O4B) and TOAC8 (N8B), Gly5 (O5B) and Gly9 (N9B), Gly6 (O6B) and Ile10 (N10B), and Leu7 (O7B) and Leu11 (N11B). The peptide nitrogen N10B atom of Ile10 also participates in a type-III β -turn with the O7B atom of the Leu7 residue (bifurcated H-bond). The Leu11 residue is found in the H (high-energy) region of the conformational space, $[31]$ with the O11B atom involved in an intermolecular H-bond with the OW_2 water molecule.

In addition to the backbone conformational differences, molecules A and B show diverging side-chain torsion angles (x) for the Bpa0 residue. In molecule A, the Bpa0 side chain has a $\chi^{\scriptscriptstyle 1}$ value close to 180 $^{\circ}$ (t conformation), whereas the $\chi^{\scriptscriptstyle 2,1}$ angle is close to -90° , while in molecule **B** the side chain has a χ^1 value close to -60° (g⁻ conformation), whereas the $\chi^{2,1}$ angle is close to 90 $^{\circ}$. In both molecules, the TOAC piperidine ring adopts a twist conformation with a planar geometry around the nitroxide moiety.[15]

The packing mode of the terminally protected dodecapeptide is characterized by two N-H \cdots O=C' intermolecular H-bonds $(NOA-H \cdots O10B=C'10B$ and $NOB-H \cdots O10A=C'10A)$ and three intermolecular H-bonds involving separately the two independent peptide molecules and the Ow₂ water molecule (N1A-H \cdots Ow₂, Ow₂-H \cdots O9B=C'9B, and Ow₂-H \cdots O11B=C'11B). These intermolecular interactions, occurring along the b direction, link together A and B molecules in a head-to-tail fashion, thus producing rows of A-to-B H-bonded peptide molecules. These rows are linked in the ac plane by H-bond interactions with the other water molecule (N5B-H \cdots Ow₁, Ow₁-H \cdots O5A=C'5A and Ow₁-H \cdots O2B=C'2B). This crystal structure is further stabilized by van der Waals interactions between the hydrophobic groups.

In summary, we have reported here the solution by the direct method of X-ray crystallography of a large peptide structure containing as many as 208 non-H atoms, without the help of any heavy atom and in a noncentrosymmetric space group. This trichogin GA IV analogue is folded in mixed $3_{10}/\alpha$ -helical conformations closely related to those already published for a different analogue^[22] and trichogin itself,^[21] thus emphasizing the limited plasticity of this lipopeptaibol.

However, in our view, the most important finding of the crystallographic work described in this communication is the remarkable flexibility of the Bpa side chain. Although the backbone conformation would be quite similar (highly helical) in the two independent molecules(the ϕ, ψ torsion angles diverge by less than 10°, and the (Bpa) $C^{\alpha} \cdots C^{\alpha}$ (TOAC) distance is 12.16 Å in molecule A and 12.40 Å in B), the separation between the Bpa side-chain carbonyl carbon and the midpoint of the N -O bond in the almost-rigid TOAC residue differs by as much as 10.96 ä in the two molecules (from 7.59 ä in molecule A to 18.55 ä in B). Thus, this significant variation should be almost entirely ascribed to the markedly different Bpa side-chain torsion angles χ^1 and χ^2 that characterize the -CH₂-phenyl group of the benzophenone moiety. It is worth noting, however, that this finding is not surprising in view of the well-known side-chain flexibility exhibited by the closely related Phe residue in peptides and proteins.[32, 33]

In conclusion, our results strongly support the view that, despite the unquestionable usefulness of the Bpa residue as a photoprobe, caution should be exercised in extrapolating the analytical results of photo cross-linking experiments and photophysical data to Bpa proximity, that is, to protein mapping and intramolecular distances, respectively. However, in this context, it is also very important to remember that peptide ligands quite often show a much reduced conformational flexibility when bound to a protein receptor site. In particular, this advantageous situation has been recently assessed in a work closely related to that reported in the present communication, namely the X-ray diffraction structure of a Bpa-containing peptide-substrate mimic in complex with the enzyme tyrosine phosphatase B.^[34] In this crystal structure, the only one published for a Bpa-bound peptide prior to that of the trichogin analogue discussed here, the Bpa residue adopts a backbone conformation close to the β sheet structure, and, more importantly, its side chain is rigidified to the q^- conformation because the bulky aromatic residue occupies a largely hydrophobic patch in the interior of the enzyme.

In any case, it is our contention that the availability of benzophenone-based, side-chain conformationally restricted α -amino acids would be extremely useful to biochemists and biophotophysicists. Work is in progress in our laboratories to this end.

Experimental Section

The synthesis and chemical characterization of Boc-Bpa-Aib-Gly-Leu-Aib-(Gly)₂-Leu-TOAC-Gly-Ile-Leu-OMe have already been reported.^[12]

HEMBIOCHEM

Colorless single crystals of this terminally protected dodecapeptide were grown by slow evaporation at room temperature from a methanol solution. Data collection was carried out on a Nonius MACH3 X-ray diffractometer at the Laboratoire de Cristallographie et de Modélisation des Matériaux Minéraux et Biologiques at the University of Nancy (France) with a Nonius FR591 rotating anode generator. Crystallographic data are listed in Table 3. The peptide

crystallizes as a dihydrate with two independent peptide molecules in the asymmetric unit, for a total of 208 non-H atoms, in the orthorhombic space group $P2_12_12_1$. In spite of several efforts to solve the structure by using a variety of programs, only the SIR 2002 program,^[35] with the default procedure was successful (i.e., only space-group symbol, cell parameters, unit cell content, and experimental structure factors were supplied). The program was able to find the correct solution in few trials (trial 34), spending 26 min of cpu time (Compaq XP1000 Unix Workstation). For each trial, SIR 2002 computes a figure of merit able to discriminate the correct solution among the different trials.

Refinement of the structure was performed by full-matrix leastsquares procedures with the program SHELXL 97.^[36] All non-H atoms were refined anisotropically. H-atoms were calculated and during the refinement they were allowed to ride on their carrying atoms, with U_{iso} set equal to 1.2 times the U_{eq} of the attached atom.

CCDC-223140 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/conts/retrieving.html (or from the Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, $UK;$ fax: $(+ 44)$ 1223-336033; or deposit@ccdc.cam.ac.uk).

Acknowledgements

This work was supported by HUIR-PRIN-2002. Antonietta Gargiulo is acknowledged for technical assistance.

Keywords: conformation analysis \cdot donor - acceptor systems \cdot helical structures \cdot lipopeptaibol antibiotic \cdot X-ray diffraction

- [1] G. Dorman, G. B. Prestwich, Biochemistry 1994, 33, 5661 5673.
- [2] J.C. Kauer, S. Erickson-Viitanen, H.R. Wolfe, Jr., W.F. De Grado, J. Biol. $Chem.$ 1986, 261, 10695 - 10700.
- [3] E. Peggion, S. Mammi, E. Schievano, Biopolymers (Pept. Sci.) 1997, 43, $419 - 431$
- [4] R. Drozdz, E. Hintermann, H. Tanner, U. Zumsteg, A. N. Eberle, J. Pept. Sci. 1999, 5, 234 - 242.
- [5] A. Bettio, A. G. Beck-Sickinger, Biopolymers (Pept. Sci.) 2001, 60, 420 437.
- [6] L. K. Henry, S. Khare, C. Son, V. V. S. Babu, F. Naider, J. M. Becker, Biochemistry 2002, 41, 6128-6139.
- [7] D. Yahalom, A. Wittelsberger, D. F. Mierke, M. Rosenblatt, J. M. Alexander, M. Chorev, Biochemistry 2002, 41, 8321 - 8331.
- [8] J. Pérodin, M. Deraët, M. Auger-Messier, A. A. Boucard, L. Rihakova, M.-E. Beaulieu, P. Lavigne, J.-L. Parent, G. Guillemette, R. Leduc, E. Escher, Biochemistry 2002, 41, 14348 - 14356.
- [9] E. Sachon, G. Bolbach, G. Chassaing, S. Lavielle, S. Sagan, J. Biol. Chem. $2002, 277, 50409 - 50414.$
- [10] J.W. Chin, P.G. Schultz, ChemBioChem 2002, 1135 1137.
- [11] C. Corvaja, E. Sartori, A. Toffoletti, F. Formaggio, M. Crisma, C. Toniolo, Biopolymers (Pept. Sci.) 2000, 55, 486 - 495.
- [12] E. Sartori, A. Toffoletti, F. Rastrelli, C. Corvaja, A. Bettio, F. Formaggio, S. Oancea, C. Toniolo, J. Phys. Chem. A 2003, 107, 6905 - 6912.
- [13] C. Toniolo, E. Benedetti, Trends Biochem. Sci. 1991, 16, 350 353.
- [14] K. A. Bolin, G. L. Millhauser, Acc. Chem. Res. 1999, 32, 1027 1033.
- [15] C. Toniolo, M. Crisma, F. Formaggio, Biopolymers (Pept.Sci.) 1998, 47, 153 -158.
- [16] I. L. Karle, P. Balaram, Biochemistry 1990, 29, 6747 6756.
- [17] C. Toniolo, M. Crisma, F. Formaggio, C. Peggion, Biopolymers (Pept.Sci.) 2001, 60, 396 - 419.
- [18] C. Auvin-Guette, S. Rebuffat, Y. Prigent, B. Bodo, J. Am. Chem. Soc. 1992, 114, 2170 - 2174.
- [19] C. Toniolo, M. Crisma, F. Formaggio, C. Peggion, V. Monaco, C. Goulard, S. Rebuffat, B. Bodo, J. Am. Chem. Soc., 1996, 118, 4952-4958.
- [20] V. Monaco, F. Formaggio, M. Crisma, C. Toniolo, P. Hanson, G. L. Millhauser, Biopolymers 1999, 50, 239 - 253.
- [21] C. Toniolo, C. Peggion, M. Crisma, F. Formaggio, X. Shui, D. S. Eggleston, Nat. Struct. Biol. 1994, 1, 908 - 914.
- [22] M. Crisma, V. Monaco, F. Formaggio, C. Toniolo, C. George, J. L. Flippen-Anderson, Lett. Pept. Sci. 1997, 4, 213 - 218.
- [23] V. Monaco, F. Formaggio, M. Crisma, C. Toniolo, P. Hanson, G. Millhauser, C. George, J. R. Deschamps, J. L. Flippen-Anderson, Bioorg. Med. Chem. 1999, $7.119 - 131.$
- [24] C. Toniolo, M. Crisma, F. Formaggio, R. F. Epand, R. M. Epand, Cell. Mol. Life Sci. 2001, 58, 1179 - 1188.
- [25] C. Peggion, F. Formaggio, M. Crisma, R. F. Epand, R. M. Epand, C. Toniolo, J. Pept. Sci. 2003, 9, 679-689.
- [26] IUPAC-IUB Commission on Biochemical Nomenclature, Biochemistry 1970, 9, 3471 - 3479
- [27] M. Saviano, R. Improta, E. Benedetti, B. Carrozzini, G. L. Cascarano, C. Toniolo, M. Crisma in Peptides 2002 (Eds.: E. Benedetti, C. Pedone), Ziino, Naples, 2002, pp. 882 - 883.
- [28] C. M. Venkatachalam, Biopolymers 1968, 6, 1425 1436.
- [29] V. Pavone, G. Gaeta, A. Lombardi, F. Nastri, O. Maglio, C. Isernia, M. Saviano, Biopolymers 1996, 38, 705-721.
- [30] C.-H. Yang, J. N. Brown, K. D. Kopple, Int. J. Pept. Protein Res. 1979, 14, 12-20.
- [31] S. S. Zimmerman, M. S. Pottle, G. Némethy, H. A. Scheraga, Macromolecules 1977, 10, $1 - 9$.
- [32] E. Benedetti, G. Morelli, G. Némethy, H. A. Scheraga, Int. J. Pept. Protein Res. 1983, $22, 1 - 15$.
- [33] P. Chakrabarti, D. Pal, Protein Eng. 1998, 11, 631 647.
- [34] M. Sarmiento, Y. A. Puius, S. W. Vetter, Y.-F. Keng, L. Wu, Y. Zhao, D. S. Lawrence, S. C. Almo, Z.-Y. Zhang, Biochemistry 2000, 39, 8171 - 8179.
- [35] M. C. Burla, M. Camalli, B. Carrozzini, G. L. Cascarano, C. Giacovazzo, G. Polidori, R. Spagna, J. Appl. Crystalloar, 2003, 36, 1103.
- [36] G.M. Sheldrick, SHELXL 97. Program for Crystal Structure Refinement, University of Göttingen, (Germany), 1997.

Received: November 3, 2003 [Z 811]