

Solution structure and stability of a disulfide cross-linked nucleopeptide duplex†

Irene Gómez-Pinto,^a Vicente Marchán,^b Federico Gago,^c Anna Grandas^{*b} and Carlos González^{*a}

^a Instituto de Química Física "Rocasolano", C/Serrano 119, 28006 Madrid, Spain.

E-mail: cgonzalez@iqfr.csic.es; Fax: +34-915642431; Tel: +34-915619400

^b Departament de Química Orgànica, Universitat de Barcelona, C/ Martí i Franquès 1, 08028 Barcelona, Spain. E-mail: anna.grandas@ub.edu; Fax: +34-933397878; Tel: +34-934021263

^c Departamento de Farmacología, Universidad de Alcalá, 28871 Madrid, Spain

Received (in Cambridge, UK) 27th June 2003, Accepted 20th August 2003

First published as an Advance Article on the web 17th September 2003

NMR methods are used to study the structure and stability of the duplex formed by the nucleopeptide [Ac-Cys-Gly-Ala-Hse(p^{3'}dGCATGC)-Ala-OH]₂[S-S], in which the oligonucleotide is self-complementary and the cysteine residues of the two peptide chains form a disulfide bridge; thermal transitions and NMR-derived structural calculations are consistent with a 3-D structure in which the oligonucleotide forms a standard B-DNA helix without significant distortions; the peptide chains are relatively disordered in solution and lie in the minor groove of the DNA helix; this nucleopeptide duplex exhibits a high melting temperature, indicating that peptide-oligonucleotide conjugates containing cysteines are suitable molecules to establish cross-links between DNA strands and stabilize the duplex.

Conformationally constrained oligonucleotides are interesting models for a variety of DNA structures, which include hairpins, cruciforms, triplexes, and so on. In many occasions, these non-canonical DNA structures are difficult to study in short oligonucleotide fragments because of their inherent flexibility. In these cases, conformationally constrained oligonucleotides are more suitable models than the pure DNA fragments.¹ Cyclic oligonucleotides are a particular case of constrained oligonucleotides that have been used as models for studying DNA hairpins² and quadruplex structures,³ and have been very useful in many thermodynamic studies. A more general approach to stabilize oligonucleotides is to connect the different strands that make up the structure with a cross-link. Cross-linked oligonucleotides can be obtained in a variety of ways, but one of the most often used methods is through incorporation of thiol groups and formation of a disulfide bridge. Such an approach has been used to study different constrained DNA⁴ and RNA structures.⁵

Nucleopeptides are peptide-oligonucleotide conjugates with a covalent phosphodiester bond between the hydroxyl group of an amino acid side chain and the 3'- or 5'-terminal hydroxyl group of an oligonucleotide chain (Fig. 1A). This kind of bond

occurs in some covalently-linked protein-DNA complexes. In addition, the fact that the peptide chain is linked to the 3'-end of the oligonucleotide chain makes these molecules more resistant to exonucleases.⁶ Peptide-oligonucleotide conjugates are also interesting because of their pharmaceutical relevance as antisense or antigene compounds. These and other potential applications of these conjugates have been reviewed recently.⁷

The aim of this study is to explore the possibility of using peptides to establish a cross-link between opposing DNA strands. In this communication we report on the structure and thermodynamics of the duplex formed by the nucleopeptide [Ac-Cys-Gly-Ala-Hse(p^{3'}dGCATGC)-Ala-OH]₂[S-S] (see Fig. 1), in which the two cysteines form a disulfide bond that covalently links the two 3' ends of the oligonucleotide moiety. The DNA sequence is self-complementary, and the peptide is long enough to allow for the crosslinking of the two cysteines without disrupting the structure of the duplex. Homoserine (Hse) was chosen as the linking amino acid to prevent base-promoted degradation in the final steps of the synthesis.⁸ The thermal stability of this compound was compared with that of the pure DNA duplex with the same sequence.

The nucleopeptide was assembled following described stepwise solid-phase procedures.⁹ Samples for NMR experiments were prepared in a 10 mM phosphate, 100 mM NaCl (pH = 7) buffer solution, and had a nucleopeptide concentration of about 1 mM. NMR spectra were acquired in a Bruker DMX spectrometer operating at 600 MHz. All 2D experiments were carried out at 5 °C. ¹H NMR assignment was conducted using established techniques.¹⁰ The assignment pathways for the peptide and oligonucleotide chains are shown in Fig. 2. A complete assignment list can be found in the ESI.† The cross-

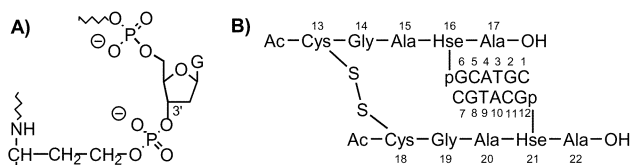


Fig. 1 A: Structure of the covalent linkage between the peptide and the oligonucleotide. B: Numbers assigned to the residues of the nucleopeptide disulfide dimer.

† Electronic supplementary information (ESI) available: the imino region of the NOESY spectrum of [Ac-Cys-Gly-Ala-Hse(p^{3'}dGCATGC)-Ala-OH]₂[S-S] in H₂O; NMR and UV melting curves of [Ac-Cys-Gly-Ala-Hse(p^{3'}dGCATGC)-Ala-OH]₂[S-S]; table of ¹H NMR assignments. See <http://www.rsc.org/suppdata/cc/b3/307300a/>

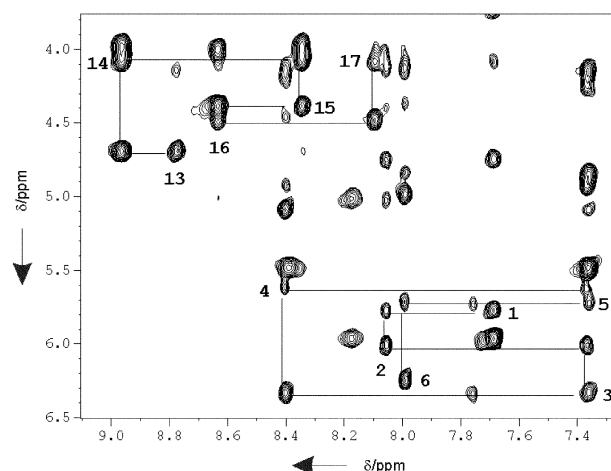


Fig. 2 Region of the NOESY spectra in H₂O ($\tau_m = 200$ ms). Sequential assignment pathways are shown for the peptide and oligonucleotide strands. Intra-residual H α -HN, and H1'-H6/H8 cross peaks are labeled with the residue number (for numbering see Fig. 1).

peak patterns observed in the NOESY experiment for the DNA exchangeable protons indicate that all bases form Watson–Crick pairs throughout the duplex (see Fig. S2 in the ESI).†

The melting behavior of [Ac-Cys-Gly-Ala-Hse(p³dGCATGC)-Ala-OH]₂[S-S] was monitored by following the changes in some well-resolved proton resonances. NMR and UV¹¹ melting experiments were carried out at different nucleopeptide concentrations. In all conditions the nucleopeptide duplex melted at 50 °C, indicating that the transition is unimolecular‡. It is interesting to compare the thermodynamic parameters for the melting transition of this molecule with those of the pure parent oligonucleotide duplex, reported in a previous publication.¹² In the oligonucleotide duplex, the thermal transition is bimolecular and, therefore, the melting temperature depends on the overall oligonucleotide concentration. Analysis of the thermodynamic data obtained from the melting transitions indicates that the higher *T_m* observed in this nucleopeptide arises from the change in the molecularity of the reaction. Thus, the disulfide cross-link provokes a strong entropic stabilization, $\Delta\Delta S^0 = 114 \text{ J mol}^{-1}$. However, the change in enthalpy is not favorable, $\Delta\Delta H^0 = -49 \text{ kJ mol}^{-1}$. This enthalpic destabilization might be due either to distortions in the DNA structure or to unfavorable contacts with the peptide.

Analysis of the NMR spectra suggests that the distortions in the DNA structure are small. The nucleotides affected by the peptide chain in the nucleopeptide can be identified by comparing their proton chemical shifts with those of the control duplex. Significant differences are only observed in residue G6, indicating that the peptide only affects the 3'-terminus of the oligonucleotide chain.

To get further insight into the effect of the cross-link on the structure of the oligonucleotide, a structure calculation was carried out based on 246 experimental distance restraints derived from NOE intensities. Peptide–DNA NOEs are only observed between the Hse and the terminal G6. Torsion angle restraints derived from the *J*-coupling constants for deoxyribose were also used. Calculations were carried out by using restrained molecular dynamics with the program DYANA 1.5¹³ and further refined with the molecular dynamics package AMBER 5.1,¹⁴ following standard protocols used in our laboratory.^{12,15} The structures were refined including explicit solvent, and using the Particle Mesh Ewald method to evaluate long-range electrostatic interactions. In the resulting structures, final distance restraint violations and *E*_{NOE} terms were small, showing the ability of all of these structures to fulfill the experimental restraints.

The ten structures resulting from the final refinement are displayed in Fig. 3. The conformation of the DNA in the ensemble of structures is well-defined (average pairwise RMSD of 0.8 Å) and belongs to the B-family of DNA structures. All glycosidic angles are *anti* and the pseudorotation phase angles of the deoxyriboses are in the general S-domain. The bases are better defined than the sugar–phosphate backbone, as is typically observed in NMR-derived structures. The cross-linked peptide does not induce significant distortions in the structure of

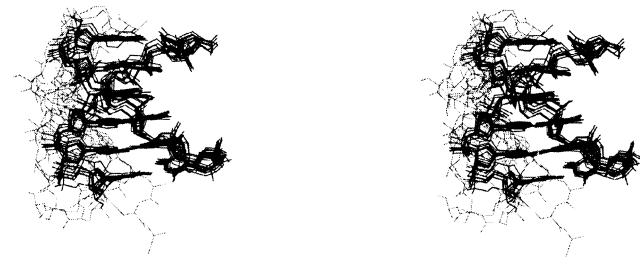


Fig. 3 Stereoscopic view of the superposition of the ten refined structures. The oligonucleotide strands are shown as thick lines. The peptide moiety is shown as thin lines.

the DNA. Structure calculations do not converge to a well-defined structure for the peptide chain. This is due to the lack of long range experimental distance constraints in this part of the molecule. All the observed NOEs are intra-residual and sequential. We conclude that the peptide chain is quite flexible in solution. However, inspection of the ensemble of structures shows that in all cases the peptide chain lies in the minor groove of the DNA (see Fig. 3). The fact that no hydrogen bonds have been detected in this case suggests that the peptide chain is not optimally suited for snugly fitting into the minor groove in the way that lexitropsins or pyrrole–imidazole hairpin polyamides do. This might account for the lower enthalpy of formation of the nucleopeptide duplex compared to that of the pure DNA complex. On the other hand, the favorable entropic effect is probably the result of a combination of factors including restricting the breathing of the duplex by covalently linking the two 3' ends, alteration of the water structure in the vicinity of the minor groove, changes in the counterion distribution around the DNA molecule, or some combination of these factors.

This work was supported by the DGICYT grant BQU-2001-3693 and the Generalitat de Catalunya (2001SGR49 and Centre de Referència de Biotecnologia).

Notes and references

‡ Note that the disulfide bridge is not broken upon duplex melting. Thermodynamic parameters for duplex formation of [Ac-Cys-Gly-Ala-Hse(p³dGCATGC)-Ala-OH]₂[S-S] are $\Delta H^0 = -136 \pm 12 \text{ kJ mol}^{-1}$, and $\Delta S^0 = -421 \pm 40 \text{ J mol}^{-1}$. In the case of (p³dGCATGC)₂ these values are $\Delta H^0 = -185 \pm 2 \text{ kJ mol}^{-1}$, and $\Delta S^0 = -535 \pm 10 \text{ J mol}^{-1}$ (The melting temperature is 28 °C at a 0.1 mM strand concentration).¹² These data were obtained according to the method of Breslauer.¹⁶

- G. D. Glick, *Biopolymers*, 1998, **48**, 83.
- J. H. Ippel, V. Lanzotti, A. Galeone, L. Mayol, J. E. van den Boogaart, J. A. Pikkemaat and C. Altona, *J. Biomol. NMR*, 1995, **6**, 403; N. Escaja, I. Gómez-Pinto, M. Rico, E. Pedrosa and C. González, *Chembiochem*, 2003, **4**, 623.
- C. González, N. Escaja, M. Rico and E. Pedrosa, *J. Am. Chem. Soc.*, 1998, **120**, 2176; N. Escaja, E. Pedrosa, M. Rico and C. González, *J. Am. Chem. Soc.*, 2000, **122**, 12732; N. Escaja, J. L. Gelpi, M. Orozco, M. Rico, E. Pedrosa and C. González, *J. Am. Chem. Soc.*, 2003, **125**, 5654.
- S. A. Wolfe and G. L. Verdine, *J. Am. Chem. Soc.*, 1993, **115**, 12585; R. J. Cain, E. R. P. Zuiderweg and G. D. Glick, *Nucleic Acids Res.*, 1995, **23**, 2153; S. A. Wolfe, A. E. Ferentz, V. Grantcharova, M. E. A. Churchill and G. L. Verdine, *Chem. Biol.*, 1995, **2**, 213; S. E. Osborne, J. Völker, S. Y. Stevens, K. Breslauer and G. D. Glick, *J. Am. Chem. Soc.*, 1996, **118**, 11993.
- S. B. Cohen and T. R. Cech, *J. Am. Chem. Soc.*, 1997, **119**, 6259.
- C. D. Juby, C. D. Richardson and R. Brousseau, *Tetrahedron Lett.*, 1991, **32**, 879; D. A. Stetsenko and M. J. Gait, *Bioconjugate Chem.*, 2001, **12**, 576; J. Robles, M. Maseda, M. Beltrán, M. Concernau, E. Pedrosa and A. Grandas, *Bioconjugate Chem.*, 1997, **8**, 785.
- C. H. Tung and S. Stein, *Bioconjugate Chem.*, 2000, **11**, 605.
- M. Beltrán, M. Maseda, J. Robles, E. Pedrosa and A. Grandas, *Lett. Pept. Sci.*, 1997, **4**, 147.
- L. Debéthune, V. Marchán, G. Fábregas, E. Pedrosa and A. Grandas, *Tetrahedron*, 2002, **58**, 6965.
- K. Wüthrich, *NMR of proteins and nucleic acids*, John Wiley & Sons, New York, 1986.
- L. Debéthune, V. Marchán, E. Pedrosa and A. Grandas, submitted.
- I. Gómez-Pinto, V. Marchán, F. Gago, A. Grandas and C. González, *Chembiochem*, 2003, **4**, 40.
- P. Guntert, C. Mumenthaler and K. Wüthrich, *J. Mol. Biol.*, 1997, **273**, 283.
- D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. C. III, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, D. M. Ferguson, R. J. Radmer, G. L. Seigel, U. C. Singh, P. K. Weiner and P. A. Kollman, in *AMBER*, 5th edition, San Francisco, 1997.
- R. Soliva, V. Monaco, I. Gómez-Pinto, N. J. Meeuwenoord, G. A. Marel, J. H. Boom, C. González and M. Orozco, *Nucleic Acids Res.*, 2001, **29**, 2973.
- K. J. Breslauer, *Methods Enzymol.*, 1995, **259**, 221.