Direct Observation of a Cytosine Analogue that Forms Five Hydrogen Bonds to Guanosine: Guanidino G-Clamp**

Christopher J. Wilds, Martin A. Maier, Valentina Tereshko, Muthiah Manoharan,* and Martin Egli*

The two main factors that stabilize pairing between nucleic acid strands are stacking interactions and hydrogen bonding,[1] properties that form the basis for biomolecular recognition. Our understanding of recognition has spurred the development of antisense molecules (molecules complementary to a coding oligonucleotide sequence) as potential therapeutics.[2] Although modified oligonucleotides offer unparalleled potential in terms of binding selectivity, there are still numerous obstacles to overcome in terms of nuclease resistance, optimum induction of RNase H activity, uptake into cells, and tissue distribution. While the phosphodiester linkage and sugar moiety have been modified extensively, modifications to the heterocyclic base have been relatively limited as it is necessary to maintain the specific hydrogen-bonding motifs required for base-pair specificity.[3]

Some heterocyclic modifications have been shown to enhance the binding affinity of nucleic acids through increased hydrogen bonding and/or base-stacking interactions. Examples include 2,6-diaminopurine, that allows for a third hydrogen bond with thymidine, and replacement of the hydrogen atom at the C5 position of pyrimidine bases with a propynyl group that allows increased stacking interactions.[4, 5] Matteucci and co-workers recently reported the synthesis and binding properties of a cytosine analogue termed G-clamp (Figure 1 A).[6, 7] For a decamer DNA sequence, a single incorporation of a G-clamp was found to increase the melting temperature T_m of the DNA duplex by 18 °C. Antisense inhibition by RNase H cleavage has been observed for a phosphorothioate oligonucleotide containing a single G-clamp, which illustrates its potential for sequence-specific downregulaton of mRNA.[6, 7] This modification is assumed to stabilize the duplex through the formation of an additional hydrogen bond to O6 from the Hoogsteen side. It may also enhance stacking as a result of a larger surface area. Unfortunately, no high-resolution structure is available to date to confirm the existence of the presumed binding motif.

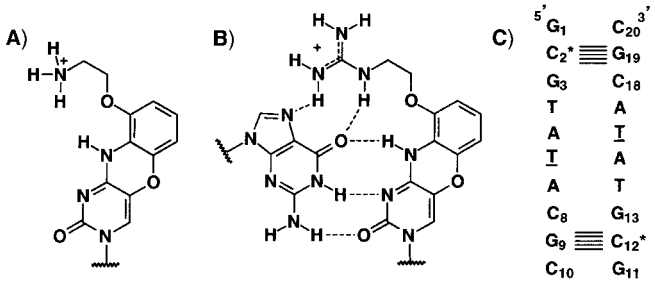

Figure 1. Structure of the tricyclic cytosine analogue G-clamp[5] (A), of its extended analogue guanidino G-clamp hybridized to complementary guanosine (B), and of the palindromic decamer duplex crystallized for this study (C). The five hydrogen bonds formed between C* and G are indicated by horizontal lines (C* = guanidino G-clamp, T = 2′-*O*-MOE-T).

The guanidino G-clamp modification 9-(2-guanidinoethoxy)-phenoxazine was designed to allow for additional hydrogen bonds to the O6 and N7 Hoogsteen binding sites of guanosine (Figure 1 B). Binding studies of DNA oligomers containing a single guanidino G-clamp unit to an RNA target revealed an increase in T_m of 16 °C relative to the unmodified DNA, slightly lower than the ΔT_m observed for the original G-clamp modification (see Supporting Information). To investigate the structural properties of this modification we determined the X-ray crystal structure of a modified decamer duplex with the sequence GC*GTAT$_{MOE}$ACGC, where C* is the guanidino G-clamp and T$_{MOE}$ is a 2′-*O*-methoxyethyl thymine (Figure 1 C).[8–10] Crystals of this decamer duplex were grown by the hanging-drop vapor-diffusion method. Data collection was performed at a synchrotron source and data collection and refinement statistics are listed in Table 1.

The overall structure of this duplex is A-form as a result of 2′-*O*-methoxyethyl thymine units at positions 6 and 16 in the duplex. An A-form environment is desirable to study the structure of nucleic acid modifications for antisense purposes. As illustrated in the case of base pair C12* – G9 (Figure 2), electron density around the heterocycles clearly shows the two Hoogsteen-type hydrogen bonds formed between the amino and imino nitrogen atoms of the tethered guanidinium

Table 1. Reflection data and refinement statistics.

Resolution [Å]	*N* (unique)	Mean $[I/\sigma(I)]$	[%] Complete	*R*-factor[a]
10.00 – 3.00	1073	26.90	98.8	0.168
3.00 – 2.50	768	31.51	99.9	0.142
2.50 – 2.00	1722	34.38	100.0	0.133
2.00 – 1.80	1288	36.70	100.0	0.116
1.80 – 1.60	2005	30.33	99.9	0.102
1.60 – 1.40	3314	27.90	100.0	0.115
1.40 – 1.20	5804	24.68	100.0	0.115
1.20 – 1.10	4680	20.08	100.0	0.116
1.10 – 1.00	6666	14.35	99.6	0.130
All data	27320	23.63	99.6	0.128

[a] $R\text{-factor} = \Sigma_{hkl} \mid F(hkl)_o - F(hkl)_c \mid / \Sigma_{hkl} F(hkl)_o$; no σ cutoff was used.

[*] Prof. M. Egli, Dr. C. J. Wilds, Dr. V. Tereshko
Department of Biological Sciences
Vanderbilt University
Nashville, TN 37235 (USA)
Fax: (+1) 615-343-6707
E-mail: martin.egli@vanderbilt.edu
Dr. M. Manoharan, Dr. M. A. Maier
Department of Medicinal Chemistry
Isis Pharmaceuticals
Carlsbad, CA 92008 (USA)
Fax: (+1) 760-929-0036
E-mail: mmanohar@isisph.com

[**] Financial support from the National Institutes of Health (Grant: GM-55237 to M.E.) is gratefully acknowledged. C.J.W. is the recipient of a post-graduate scholarship from the National Science and Engineering Research Council of Canada (NSERC). We thank Dr. George Minasov for his assistance with data collection and Guity Balow and Robert H. Springer for synthetic intermediates. Use of the 5-ID beamline of the DuPont-Northwestern-Dow CAT Synchrotron Research Center at the Advanced Photon Source (Argonne, IL) was supported by E. I. DuPont de Nemours & Co., The Dow Chemical Company, the NSF, and the State of Illinois.

Supporting information for this article is available on the WWW under http://www.angewandte.com or from the author.

  1433-7851/02/4101-0115 $ 17.50+.50/0

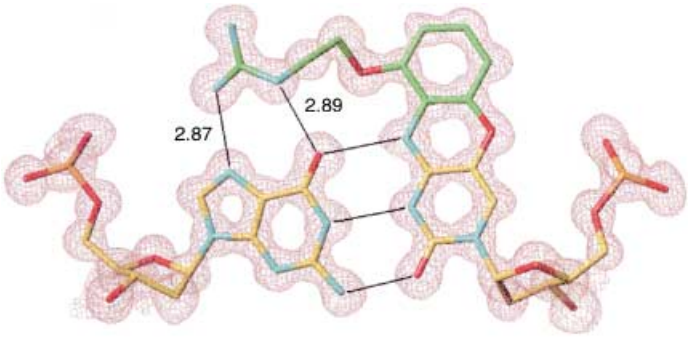

Figure 2. Fourier $(2F_o - F_c)$ sum electron-density map (contoured at 1.25σ) around C12* and G9 confirming formation of five hydrogen bonds (indicated by thin solid lines, selected distances shown [Å]). Red: oxygen, blue: nitrogen, orange: phosphorus, yellow: carbon, green: carbons of the modification.

and O6 and N7 of guanosine, respectively. The hydrogen bond lengths are 2.89 and 2.87 Å and the lengths of the corresponding hydrogen bonds in base pair C2* – G19 are 2.88 Å and 2.92 Å, respectively. The quality of the electron density data around individual atoms of the phenoxazine ring and tethered group demonstrates that this modification is well ordered and does not assume random conformations. There is some buckling of modified base pairs relative to the other base pairs in the duplex (not shown). This out-of-plane distortion of the base pair between the guanidino G-clamp and G may be a consequence of the requirement to optimize the geometry of both the Watson – Crick and Hoogsteen-type hydrogen bonds within the geometric boundaries provided by the guanidinium ethoxy moiety. In addition, the observed arrangements help to avoid a steric contact between O6 of G and the ethoxy-linker oxygen of the G-clamp (Figure 1 and 2).

The presence of the tricyclic G-clamp moiety results in a considerable improvement of intra-strand stacking at the GpC* step compared with stacking between cytosine and the 5′-adjacent base (G1 and G11, respectively). The overlap between G1 and C2* is depicted in Figure 3. While the "cytosine core" displays relatively little stacking to the guanosine base, the remainder of the phenoxazine ring system virtually covers the entire guanosine base. However, while stacking between G-clamp and the base to the 5′-side is improved, stacking to the 3′-adjacent base is not affected by incorporation of the modified base.

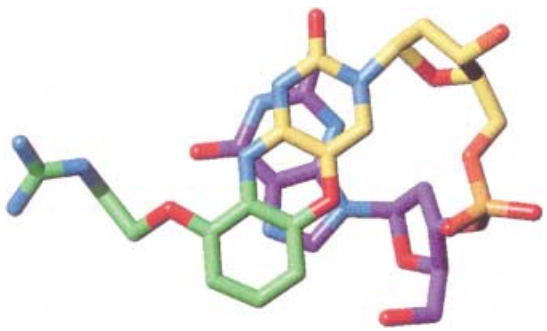

Figure 3. Stacking between G1 and C2*, viewed approximately along the vertical to the phenoxazine rings. Carbon atoms of G1 are shown in magenta, carbon atoms of the cytosine core of C2* are shown in yellow and the remainder of the carbon atoms are in green.

Placement of the positively charged guanidinium moiety in the center of the major groove, a site of strong negative potential, results in a significant electrostatic contribution to stability. Moreover, the guanidinium group and phosphates from opposite strands are relatively closely spaced. The average distance between the imino nitrogens of C* and O2P oxygens of phosphate units is 5.8 Å. Although too long for direct salt bridges, water molecules link guanidinium and phosphate groups. For C12*, single water molecules mediate contacts between guanidinium imino nitrogens and O2P oxygens of residues C8 and G9.

Interactions between positively charged amines and the Hoogsteen binding site of guanosine are well known. For example, X-ray crystallographic studies of the λ repressor bound to duplex DNA revealed specific contacts between a lysine and the O6 position of G.[11, 12] The interactions in the structure of the guanidino G-clamp here are similar to the bidentate hydrogen bonding of the arginine fork with the N7 and O6 positions of guanine observed in zinc fingers from Zif268 (a mouse immediate early protein).[13]

Two crucial stabilizing factors of this modification are an increase in the number of hydrogen bonds and improved stacking interactions. Additional contributions to stability are favorable electrostatic interactions and well-ordered water networks. It is difficult to discern if one of these contributions plays a more important role than the others. Binding studies of oligomers with the phenoxazine moiety alone showed moderate increases in T_m of 2 – 7 °C.[14] Stability was increased most when several phenoxazine groups were clustered together on the same strand, allowing for tricyclic – tricyclic stacking interactions. In the case of an acyclic G-clamp modification, no enhancement in binding was observed. Only when both the phenoxazine and tethered amino group were present was a drastic improvement in binding observed.[6] Clearly, hydrogen bonds from the guanidinium group maintain the guanidino G-clamp modification in a position that allows stacking interactions and formation of stable water networks. In the absence of an analogous crystal structure of the duplex containing the original G-clamp modification, it is difficult to hypothesize what the origin of the slight destabilization of the guanidino G-clamp relative to the original amino G-clamp might be.

To summarize, this is the first report of a single base pair within a nucleic acid duplex combining Watson – Crick and Hoogsteen binding to a total number of five hydrogen bonds. Further evaluation of the antisense therapeutic and diagnostic applications of this novel guanidino G-clamp modification is in progress.

Experimental Section

Crystallization by the hanging-drop vapor-diffusion method using commercially available screens (Hampton Research, Laguna Niguel, CA): A 2 μL droplet (1.2 mM DNA, 5 % 2-methyl-2,4-pentanediol (MPD), 20 mM Na cacodylate pH 6.0, 6 mM spermine · 4 HCl, 40 mM NaCl, 6 mM KCl, 10 mM $MgCl_2$) was equilibrated against a reservoir of 1 mL 35 % v/v MPD. Crystals suitable for X-ray diffraction grew in one week.

Data Collection and Processing: A crystal ($0.7 \times 0.2 \times 0.2$ mm) was picked up from a droplet with a nylon loop and transferred into a cold N_2 stream (120 K). High- and low- resolution data sets were collected on the 5-ID

  1433-7851/02/4101-0116 $ 17.50+.50/0

beam line ($\lambda = 0.978$ Å) of the DND-CAT at the Advanced Photon Source, Argonne, IL, using a MARCCD detector. Data were integrated and merged with DENZO/SCALEPACK.[15] Space group $P2_12_12_1$; cell dimensions $a = 24.52$, $b = 43.02$, $c = 46.68$ Å. The overall R_{merge} for all reflections between 20 and 1 Å was 4.7%.

Structure Determination and Refinement: The structure was solved by molecular replacement using the DNA decamer described in ref. [16] as the initial model and refined with the programs CNS[17] and SHELX-97.[18] After monitoring the R_{free} using 10% of the reflections, all reflections were included in the final rounds of isotropic refinement. Hydrogen atoms were added in SHELX-97 and all atoms including the spermine molecule and solvent water molecules were treated anisotropically. Coordinates and structure factors have been deposited in the Protein Data Bank (accession code 1KGK).

Received: July 18, 2001
Revised: October 8, 2001 [Z17542]

[1] W. Saenger, *Principles of Nucleic Acid Structure*, Springer, New York, **1984**.
[2] S. T. Crooke, *Methods Enzymol.* **2000**, *313*, 3–45.
[3] P. Herdewijn, *Antisense Nucleic Acid Drug Dev.* **2000**, *10*, 297–310.
[4] A. Chollet, A. Chollet-Damerius, E. H. Kawashima, *Chem. Scr.* **1986**, *26*, 37–40.
[5] R. W. Wagner, M. D. Matteucci, J. G. Lewis, A. J. Guttierrez, C. Moulds, B. C. Froehler, *Science* **1993**, *260*, 1510–1513.
[6] K.-Y. Lin, M. D. Matteucci, *J. Am. Chem. Soc.* **1998**, *120*, 8531–8532.
[7] W. M. Flanagan, J. J. Wolf, P. Olson, D. Grant, K. Lin, R. W. Wagner, M. D. Matteucci, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3513–3518.
[8] K.-H. Altmann, N. M. Dean, D. Fabbro, S. M. Freier, T. Geiger, R. Häner, D. Husken, P. Martin, B. P. Monia, M. Müller, F. Natt, P. Nicklin, J. Phillips, U. Pieles, H. Sasmor, H. E. Moser, *Chimia* **1996**, *50*, 168–176.
[9] M. Teplova, G. Minasov, V. Tereshko, G. B. Inamati, P. D. Cook, M. Manoharan, M. Egli, *Nat. Struct. Biol.* **1999**, *6*, 535–539.
[10] The synthesis and purification of the oligonucleotides containing the guanidino G-clamp phosphoramidite will be published elsewhere.
[11] N. D. Clarke, L. J. Beamer, H. R. Goldberg, C. Berkower, C. O. Pabo, *Science* **1991**, *254*, 267–270.
[12] A. Rich in *The Chemical Bond: Structure and Dynamics* (Ed.: A Zewail), Academic Press, San Diego, CA, **1992**, pp. 31–86.
[13] N. P. Pavletich, C. O. Pabo, *Science* **1991**, *252*, 809–816.
[14] K.-Y. Lin, R. J. Jones, M. D. Matteucci, *J. Am. Chem. Soc.* **1995**, *117*, 3873–3874.
[15] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307–326.
[16] M. Egli, V. Tereshko, M. Teplova, G. Minasov, A. Joachimiak, R. Sanishvili, C. M. Weeks, R. Miller, M. A. Maier, H. Y. An, P. D. Cook, M. Manoharan, *Biopolym. Nucleic Acid Sci.* **1998**, *48*, 234–252.
[17] A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Niges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, K. L Warren, *Acta Crystallogr. Sect. D* **1998**, *54*, 905–921.
[18] G. M. Sheldrick, T. R. Schneider, *Methods Enzymol.* **1997**, *277*, 319–343.

De Novo Protein Surface Design: Use of Cation–π Interactions to Enhance Binding between an α-Helical Peptide and a Cationic Molecule in 50% Aqueous Solution**

Brendan P. Orner, Xavier Salvatella, Jorge Sánchez Quesada, Javier de Mendoza, Ernest Giralt, and Andrew D. Hamilton*

We have previously shown that synthetic guanidinium-based receptors (such as **1**; Scheme 1) can recognize aspartate- and glutamate-containing peptides in 10% aqueous methanol.[1–5] The principal driving force for binding was ion pairing and hydrogen bonding between the carboxylate and guanidinium groups that leads, with optimal matching of the sequence, to peptide stabilization in a helical conformation. In order to extend this approach to targets of more biological relevance we sought to identify peptide sequences that could recognize the receptor with increased affinity and in a more aqueous solvent system. One strategy to increase the strength of the interaction would be to provide additional binding contacts between the receptor and the helical peptide. Herein we report the combination of ion pairing and cation–π interactions projected from a helical surface to enhance the affinity of the synthetic receptor for the peptide. This approach represents an example of de novo design of a protein surface where the properties of the surface of a protein (in this case the helical peptide) are modified to optimize binding to a target molecule.

The design of the peptide sequences exploited a "molecular hinge" strategy that we had used earlier in synthetic receptors for nucleotide bases.[6, 7] In this approach acylaminopyridine and naphthalene groups were linked to provide binding through hydrogen-bonding and π-stacking interactions (Scheme 1A). In this current research our target peptides incorporate four aspartate groups, each flanked above or below by aromatic residues on the helix surface. In this way the tetraguanidinium receptor can potentially make both hydrogen-bonding and cation–π contacts[8] with side-chain groups on the peptide (Scheme 1B).[9, 10]

Three peptides were synthesized with four aspartic acid residues in the *i*, *i+3* relationship that has been shown to be

[*] Prof. A. D. Hamilton, B. P. Orner
Yale University
New Haven, Connecticut 06520 (USA)
Fax: (+1) 203-432-3221
E-mail: andrew.hamilton@yale.edu
X. Salvatella, Prof. E. Giralt
Departament de Química Orgànica
Universitat de Barcelona
Martí í Franqués 1-11, 08028-Barcelona (Spain)
Dr. J. Sánchez Quesada, Prof. J. de Mendoza
Departmento de Química Orgánica
Universidad Autónoma de Madrid
Cantoblanco 28049-Madrid (Spain)

[**] The authors thank Lara Estroff for critical reading of the manuscript. This work was supported by the National Institute of Health (GM35208) and CYCIT (BIO99-484).

Supporting information for this article is available on the WWW under http://www.angewandte.com or from the author.

  1433-7851/02/4101-0117 $ 17.50+.50/0