Structural Basis for Recognition of Guanosine by a Synthetic Tricyclic Cytosine Analogue: Guanidinium G-Clamp

by **Christopher J. Wilds**[a]), **Martin A. Maier**[b]), **Muthiah Manoharan***[b]), and **Martin Egli***[a])

[a]) Vanderbilt University, Department of Biological Sciences, Nashville, Tennessee 37212 USA (fax: 1-615-343-6707; email: martin.egli@vanderbilt.edu)

[b]) *Isis Pharmaceuticals, Inc.*, Department of Medicinal Chemistry, Carlsbad, California 92008 USA (email: mmanohar@isisph.com)

Dedicated to Professor *Jack D. Dunitz* on the occasion of his 80th birthday

An oligonucleotide analogue containing a novel heterocyclic analogue, the guanidinium G-clamp, was designed to allow formation of five H-bonds to guanosine. The guanidinium group was introduced postsynthetically by treatment of the deprotected oligonucleotide containing a free amino group with a solution of 1*H*-pyrazole-1-carboxamidine and purified by a combination of size-exclusion chromatography and reversed-phase HPLC. A single incorporation of this modification into an oligodeoxynucleotide sequence was found to increase duplex stability by 13° and 16° per modification to RNA and DNA, respectively. Crystals of a self-complementary decamer sequence containing this modification were grown and diffracted to 1-Å resolution. The structure was solved by molecular replacement and revealed that the modification forms additional H-bonds to O(6) and N(7) of guanosine through the amino and imino N-atoms, respectively. The origins of enhanced duplex stability are also attributed to increased stacking interactions mediated by the phenoxazine moiety of the G-clamp and formation of H-bond networks between the positively charged guanidinium group, H_2O molecules, and negatively charged O-atoms from phosphates on the adjacent strand.

Introduction. – The rational design of molecules that can sequence specifically inhibit viral protein biosynthesis has spurred a new paradigm of drug discovery: antisense [1]. For a chemically modified oligonucleotide (ON) to be effective as a therapeutic agent, it should exhibit a number of properties, such as improved binding to the intended target (RNA) as well as stability to various nucleases that could degrade it before reaching the target [2]. To address these critical properties, numerous modifications have been synthesized that contain alterations at the heterocycle, sugar, and internucleotide linkage [3].

To date, the modification that has exhibited the greatest effect on increased binding to the target is the heterocyclic modification known as the G-clamp [4]. This modification has been designed to allow for improved binding to a guanosine *via i*) improved stacking interactions due to an extended aromatic (phenoxazine) ring system and *ii*) formation of a fourth *Hoogsteen*-type H-bond between a tethered amino group and O(6) of G. Sequence-specific antisense inhibition has been observed when the modification is incorporated into a 15-mer phosphorothioate oligodeoxynucleotide targeted to the cyclin-dependent kinase inhibitor $p27^{kip1}$ [5]. In addition, incorporation at the 3′-terminus of an oligonucleotide offers protection against 3′-exonuclease degradation [6]. However, to date, no high-resolution crystal structure or NMR data has proven unequivocally the mechanism of binding for the original G-clamp modification.

Recently, we reported the synthesis of a second-generation G-clamp analogue: the guanidinium G-clamp obtained by postsynthetic conversion of the 2-aminoethoxy tether of G-clamp into a guanidinium-ethoxy moiety [7]. In the double-helical structure of nucleic acids, the 2-aminoethoxy spacer of the G-clamp protrudes into the major groove (*Fig. 1, a*). We rationalized that the extension of the 2-aminoethoxy spacer of the G-clamp to a planar guanidinium derivative would provide another *Hoogsteen* bond between the imino or amino N-atoms of the tethered guanidinium and N(7) of a complementary guanine base (*Fig. 1, b*). In addition, the guanidinium group exhibits a higher pK_a value (p$K_a \approx 12.5$) compared to the amino functionality, which provides a positive charge that is maintained over a wide pH range.

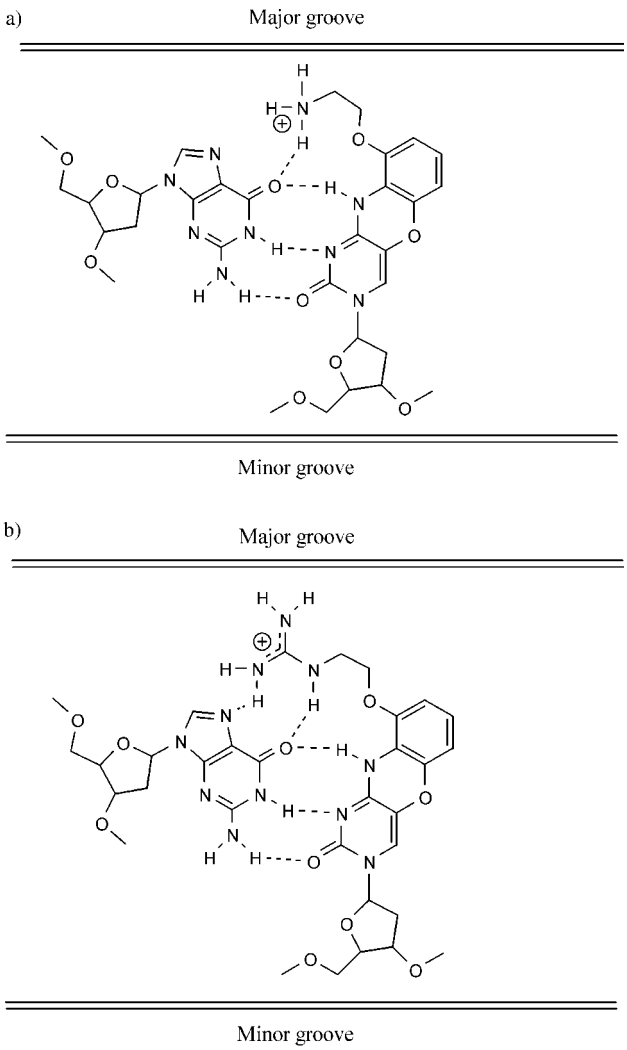
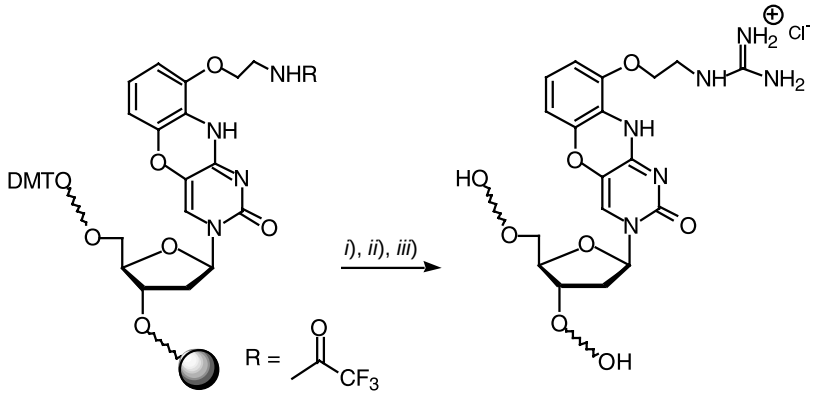


Fig. 1. *H-Bonding scheme of* a) *G-clamp and* b) *guanidinium G-clamp to guanosine*

We also reported preliminary structural data of the guanidinium G-clamp modification within a DNA duplex [8]. Here, we report the full structural details of the duplex formed by a self-complementary decamer sequence containing a single guanidinium G-clamp modification that has crystallized in the A-form. The high-resolution structure, based on X-ray-diffraction data collected up to 1-Å resolution, gives insight as to the origin of the enhanced stability for this family of heterocyclic modifications.

Results and Discussion. – *Oligonucleotide Synthesis and Postsynthetic Guanidinylation.* The G-clamp containing oligonucleotides (ONs) were synthesized by standard phosphoramidite methodology. Prior to the guanidinylation reaction, the oligonucleotides were purified by reversed-phase HPLC. As outlined in the *Scheme*, a postsynthetic strategy has been employed to selectively convert the amino tether of the G-clamp to its corresponding guanidinium derivative. To achieve this, the fully deprotected ON is reacted with 1*H*-pyrazole-1-carboxamidine hydrochloride in aqueous Na_2CO_3 solution [9]. The guanidinium-modified ONs synthesized for this study are summarized in *Table 1*. Interestingly, in the case of the self-complementary sequences ON-1 and ON-3, complete guanidinylation of the amino groups could only be achieved by performing the reaction at elevated temperature (55°) and for an extended reaction time (12). This confirms that, within the double-stranded structure of these self-complementary ONs, the primary amino group is involved in base-pairing interactions and is not susceptible to the guanidinylation reagent. After desalting and purification of the compounds by another reversed-phase HPLC step, the guanidinylated compounds were analyzed by electrospray mass spectrometry (ES-MS) and capillary gel electrophoresis (CGE) analysis. There was no evidence for residual starting material bearing the primary amino function or any other side product in the purified samples.

Scheme. *Postsynthetic Guanidinylation of G-Clamp Oligonucleotides*



i) 40% aq. $MeNH_2$/conc. aq. NH_3 1:1, 55°, 1 h. *ii*) HPLC purification and DMT deprotection. *iii*) 1.0M soln. of 1*H*-pyrazole-1-carboxamidine hydrochloride in 1.0M aq. Na_2CO_3, 3 h at r.t., for ON-1, ON-3: 12 h at 55°, isolated yield after 2nd HPLC purification: 50–60%.

Table 1. *Sequence and* T_m *Data*[a]) *for the G-Clamp and Guanidinium G-Clamp Oligonucleotides* (ONs)

ON	Sequence 5′ → 3′	C* Modification	MW_{calc}	MW_{found}	Target	T_m [°]	ΔT_m/mod
ON-1	GC*G TAT_{MOE}[b]) ACG C	G-Clamp	3251.3	3251.0	DNA	> 95	–
ON-2	GC*G TA T ACG C	Guanidinium G-clamp	3293.3	3292.8	DNA	> 95	–
ON-3	GCG TA T AC*GC	G-Clamp	3251.3	3251.0	–	N/A	–
ON-4	GCG TA T AC*GC	Guanidinium G-clamp	3293.3	3293.0	–	N/A	–
ON-5	TCT CC*C TCT C	G-Clamp	3039.1	3039.4	DNA	59.2	22.1
					RNA	70.8	18.4
ON-6	TCT CC*C TCT C	Guanidinium G-clamp	3081.1	3080.8	DNA	53.5	16.4
					RNA	64.9	12.5

[a]) Buffer: 100 mM NaCl, 10 mM phosphate, and 0.1 mM EDTA, pH 7.0; [b]) 2′-*O*-(2-methoxyethy)thymidine. N/A: not available.

The change in binding affinity towards target RNA and DNA as a consequence of conversion of the G-clamp to a guanidinium G-clamp was evaluated by thermal-melting analysis with ON-1, ON-5, and their guanidinylated analogues ON-2 and ON-6. The melting experiments with the self-complementary sequences ON-1 and ON-2 were performed without an additional target strand, and the transition temperature for the helix to coil conversion could not be determined ($T_m > 95°$). However, this observation certainly confirmed the intrinsic high-affinity nature of the G-clamp modification and its guanidinium analogue. The corresponding unmodified DNA shows a $T_m < 50°$. However, the T_m analysis of guanidinylated ON-6 in comparison with ON-5 did not show the increase in affinity expected for the formation of an additional (5th) H-bond. In this particular sequence context (the G-clamp is flanked by two C-atoms, and the target strand has a five-base overhang at its 5′-end), guanidinylation resulted in a decreased T_m for pairing with both the RNA (64.9° vs. 70.8°) and the DNA (53.5° vs. 59.2°) target strands.

Crystallization and Structure Determination of the Decamer Sequence Containing the Guanidinium G-Clamp. To obtain crystals of DNA containing the guanidinium G-clamp, several oligonucleotide sequences that are known to readily crystallize were synthesized as analogues containing guanidinium G-clamps in place of selected cytosines (*Table 1*). The sequence that produced the best quality crystals was a decamer containing a guanidinium G-clamp (C*) at C(2) and a 2′-*O*-(2-methoxyethyl)-modified (MOE) residue at position T6 (ON-1; *Table 1*). We previously investigated the structure of the decamer containing only the 2′-MOE modification and found it to crystallize in the A-form [10]. Setups performed with commercially available screens gave diffraction quality crystals in a week. We also synthesized sequences that contained the 'original' G-clamp modification. However, none of these have yielded crystals.

Crystals containing the guanidinium G-clamp modification were cryofrozen in liquid N_2, and X-ray-diffraction data to 1-Å resolution were collected at a synchrotron source. The structure was determined by the molecular replacement technique with another A-form decamer duplex as the search model [11]. Selected crystal and structure refinement data are shown in *Table 2*.

Similar to the duplex containing only the 2′-MOE modification, the duplex with incorporated guanidinium G-clamps crystallized in the A-form, illustrating that the

Table 2. *Selected Crystal Data and Refinement Parameters*

Space group	orthorhombic $P2_12_12_1$
Cell constants [Å]	$a = 25.24$, $b = 43.02$, $c = 46.68$
Resolution [Å]	1.00
No. of unique reflections (10.0 – 1.0 Å)	27320
Completeness [%]	99.6
R_{merge} (%)	4.7
R Factor ([%], no σ cutoff)	12.8
No. of refinement parameters	5528
No. of restraints	26595
No. of DNA atoms (incl. H-atoms)	708
No. of H_2O	155, 1 spermine
R.m.s. bonds [Å]	0.015
R.m.s. angles ($1\cdots3$ dist., Å)	0.021

base modification has little effect on the global structure (*Fig. 2*). All sugars adopt a C(3′)-*endo* type pucker with very little deviation from the reference duplex in terms of the backbone and glycosyl torsion angles (*Table 3*).

Mechanism of Binding of Guanidinium G-Clamp to Guanosine. The final model clearly shows formation of *Hoogsteen* H-bonds to O(6) and N(7) of guanosine by the amino and imino N-atoms of the tethered guanidinum group (*Fig. 3*). The excellent quality of the electron-density map around the tethered groups (not shown)

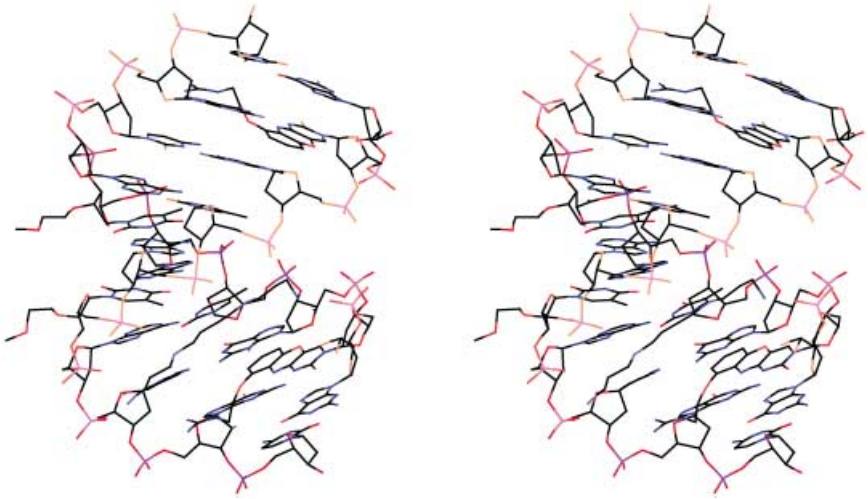


Fig. 2. *Stereo illustration of the DNA duplex [d(GC*GTAT$_{MOE}$ACGC)]$_2$ with single guanidinium G-clamp* (C*) *and 2′-*O*-methoxyethylthymine* (T_{MOE}) *modifications per strand.* The view is across the major and minor grooves. The two guanidinium G-clamps are visible on the right-hand side (upper and lower halfs) and the two MOE substituents are jutting into the minor groove on the left. Each duplex harbors a spermine molecule in the major groove; the polyamine adopts an extended conformation with multiple H-bonds to phosphate groups from both backbones and is visible in the lower half of the drawing.

Table 3. *Backbone Torsion Angles*[a]), *Glycosyl Angles, and Sugar Puckers* (deviations from reference structure [11] in parentheses)

Residue	α [deg]	β [deg]	γ [deg]	δ [deg]	ε [deg]	ξ [deg]	χ [deg]	P [deg]	Amplitude	Pucker
G(1)	–	–	67.7 (– 5.1)	86.2 (– 1.8)	– 142.5 (2.3)	– 72.9 (– 1.5)	– 178.7 (4.3)	14.2 (– 3.3)	37.1 (2.4)	C(3′)-*endo*
C(2)	*– 65.7 (5.8)*	*176.1 (– 17.4)*	*51.7 (5.6)*	*90.3 (– 8.8)*	*– 166.3 (0.7)*	*– 67.4 (13.6)*	*– 148.2 (-14.0)*	*10.2 (9.8)*	*39.0 (2.6)*	C(3′)-*endo*
G(3)	151.7 (56.0)	– 173.5 (– 10.0)	– 179.4 (– 44.8)	85.7 (– 7.5)	– 152.2 (– 3.2)	– 66.9 (– 5.9)	– 175.4 (– 3.4)	12.0 (15.0)	35.5 (4.3)	C(3′)-*endo*
T(4)	– 79.6 (5.5)	– 161.7 (– 5.9)	44.5 (4.7)	83.4 (0.1)	– 148.6 (– 3.9)	– 67.0 (6.4)	– 147.1 (0.6)	12.6 (2.1)	39.6 (-1.0)	C(3′)-*endo*
A(5)	139.6 (– 8.1)	– 170.6 (– 4.5)	– 175.9 (7.1)	80.2 (6.2)	– 146.7 (– 0.5)	– 65.8 (– 2.7)	– 165.6 (2.9)	8.9 (– 4.0)	43.0 (-1.8)	C(3′)-*endo*
T(6)	– 65.2 (0.2)	172.3 (– 3.5)	50.7 (1.6)	76.6 (3.3)	– 147.5 (– 2.3)	– 71.3 (– 4.5)	– 156.9 (3.4)	10.9 (2.8)	46.2 (3.3)	C(3′)-*endo*
A(7)	– 68.2 (– 2.9)	174.9 (– 0.4)	53.7 (– 0.4)	81.8 (1.6)	– 171.0 (12.4)	– 58.2 (– 11.5)	– 158.6 (– 0.4)	20.4 (– 5.4)	38.1 (4.9)	C(3′)-*endo*
C(8)	139.5 (155.7)	– 176.3 (– 15.4)	– 173.4 (– 129.3)	84.9 (– 4.6)	– 150.9 (– 5.9)	– 68.9 (– 5.7)	– 173.0 (13.7)	6.9 (12.9)	37.8 (5.5)	C(3′)-*endo*
G(9)	– 74.3 (7.7)	– 178.7 (– 10.2)	51.9 (5.5)	76.8 (4.6)	– 148.3 (– 9.9)	– 69.1 (6.9)	– 163.1 (– 5.1)	14.4 (8.0)	44.0 (– 2.8)	C(3′)-*endo*
C(10)	– 62.2 (– 11.3)	177.4 (9.1)	55.7 (– 2.2)	80.9 (– 4.4)	–	–	– 158.8 (1.3)	18.9 (5.4)	40.3 (0.5)	C(3′)-*endo*
G(11)	–	–	56.8 (106.7)	90.2 (– 1.9)	– 149.7 (– 1.1)	– 71.3 (4.4)	– 167.6 (– 4.3)	8.5 (351.0)	33.0 (7.0)	C(3′)-*endo*
C(12)	*– 70.0 (2.3)*	*174.8 (0.9)*	*51.6 (5.7)*	*76.9 (6.4)*	*– 149.4 (1.6)*	*– 60.0 (– 9.3)*	*– 152.7 (– 13.2)*	*13.1 (2.6)*	*42.9 (-0.6)*	C(3′)-*endo*
G(13)	– 69.7 (0.5)	172.2 (– 4.6)	52.9 (7.4)	74.9 (2.6)	– 163.9 (1.0)	– 64.2 (– 6.2)	– 170.2 (2.5)	22.6 (6.4)	47.5 (-2.4)	C(3′)-*endo*
T(14)	– 70.8 (– 2.6)	– 173.1 (– 8.0)	52.2 (2.6)	73.8 (7.1)	– 150.5 (– 6.3)	– 71.2 (0.7)	– 161.3 (– 0.1)	17.8 (3.9)	45.7 (-0.8)	C(3′)-*endo*
A(15)	– 66.0 (– 0.3)	170.6 (1.6)	60.3 (– 3.2)	80.1 (1.2)	– 157.3 (– 3.6)	– 71.7 (0.6)	– 161.3 (– 1.8)	12.7 (4.2)	41.1 (0.5)	C(3′)-*endo*
T(16)	– 70.9 (– 3.6)	175.3 (4.5)	52.5 (– 0.9)	76.2 (5.3)	– 147.9 (– 0.2)	– 67.1 (– 11.9)	– 153.5 (0.0)	11.9 (6.9)	44.4 (0.7)	C(3′)-*endo*
A(17)	– 74.0 (6.5)	175.4 (– 3.2)	52.3 (2.6)	75.9 (5.8)	– 152.7 (11.8)	– 65.2 (– 10.5)	– 160.6 (2.0)	16.0 (– 1.6)	44.2 (– 1.3)	C(3′)-*endo*
C(18)	– 63.1 (6.8)	170.5 (– 11.0)	57.1 (1.9)	84.4 (– 2.8)	– 170.0 (6.1)	– 65.1 (– 11.9)	– 158.5 (1.9)	18.8 (15.7)	35.7 (7.8)	C(3′)-*endo*
G(19)	150.8 (142.7)	– 172.4 (– 21.1)	-179.1 (-115)	88.3 (– 6.3)	– 144.4 (– 16.8)	– 62.5 (– 9.9)	– 177.2 (7.2)	8.2 (15.5)	33.6 (7.8)	C(3′)-*endo*
C(20)	– 70.3 (– 9.2)	– 178.3 (4.9)	52.2 (8.4)	87.9 (1.8)	–	–	– 154.8 (1.5)	32.6 (– 10.9)	33.1 (5.7)	C(3′)-*endo*

[a]) Backbone torsion angles defined as: O(3′)-P-α-O(5′)-β-C(5′)-γ-C(4′)-δ-C(3′)-ε-O(3′)-ξ-P-O(5′).

demonstrates that the modification is well-ordered and does not assume a random orientation. The H-bond distances between the amino and imino N-atoms to the O(6) and N(7) atoms of guanosine are 2.88 and 2.92 Å, respectively, for the C2*/G19 base pair (*Fig. 3, a*) and 2.89 and 2.87 Å, respectively, for the C12*/G9 base pair. Overall, the conformations of the two guanidinium G-clamp/G pairs in the structure of the modified DNA duplex are rather similar (*Fig. 3, b*). The root mean-square (r.m.s.) deviation between the two base pairs is 0.1 Å.

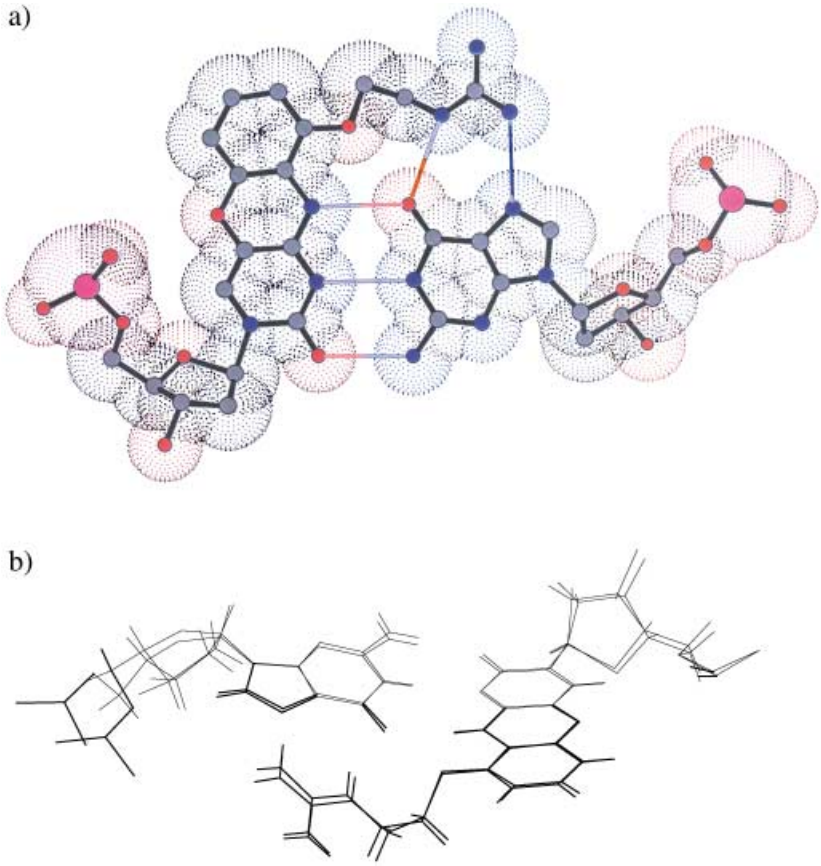


Fig. 3. a) *Diagram of the C2*/G19 base pair, illustrating the formation of* Hoogsteen*-type H-bonds to O(6) and N(7) of G by the amino and imino N-atoms of the guanidinium group in C* with distances of 2.89 and 2.87 Å, respectively* (colored lines). The view is roughly perpendicular to the phenoxazine ring and the space-filling representation is based on approximate *Van der Waals* radii of individual atoms. b) *Superposition of the C2*/G19 and G9/C12* base pairs* (positions of H-atoms are calculated).

The H-bonds between the guanidinium moiety of the G-clamp and the *Hoogsteen* face of guanine are reminiscent of a common, sequence-specific interaction between arginine and guanine in protein–DNA complexes. For example, in the complex between DNA and Zif 268, a mouse immediate early protein, arginines in all three Zn fingers engage in bidentate interactions to Gs (*Fig. 4*) [12]. However, the recognition of

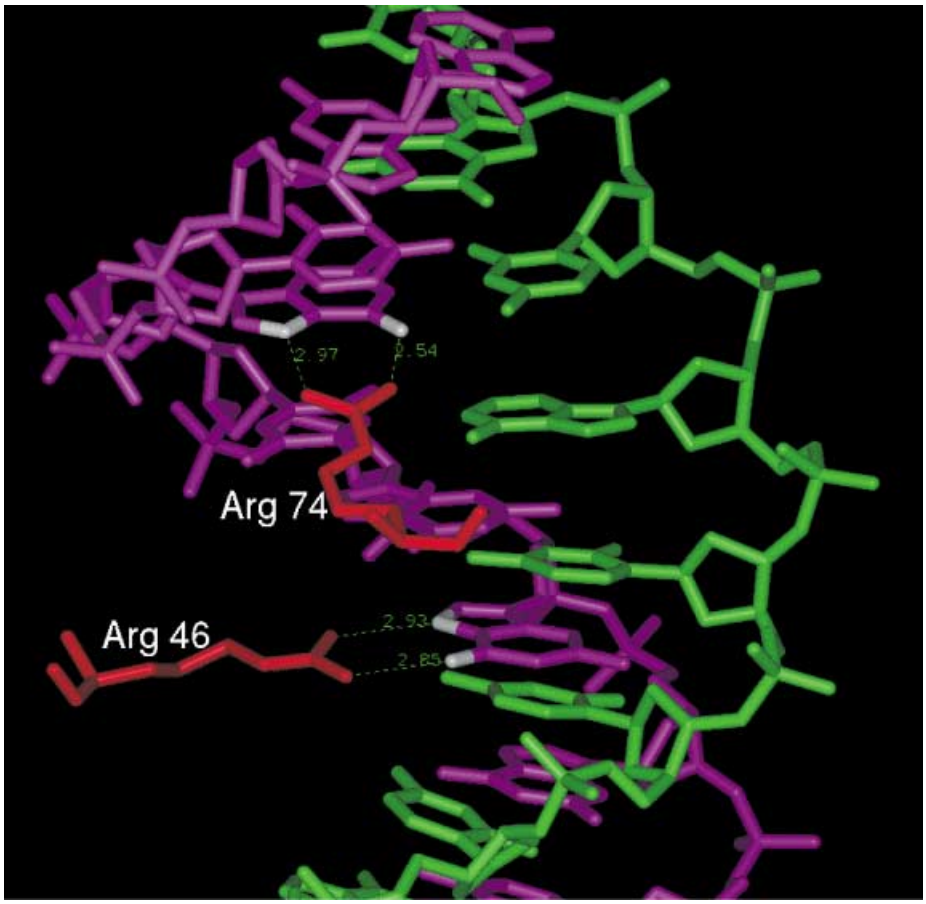


Fig. 4. *Examples of sequence-specific interactions between arginine and guanosine in the Zif268-DNA complex* [12]. The view is into the major groove of the B-form DNA, the N(7)- and O(6)-atoms of G are highlighted in gray, and H-bonds are dashed lines with distances in Å.

guanine by a single arginine *via* two H-bonds in the DNA major groove usually involves both imino N-atoms of the guanidinium group rather than amino and imino N-atoms as observed in our structure (*Fig. 3*).

Stacking Interactions of the Guanidinium G-Clamp. The presence of the tricyclic guanidinium G-clamp analogue enhances stacking to the 5′-adjacent base (a G in the present case; *Fig. 5*). The 'cytosine' core of the guanidinium G-clamp exhibits very little stacking to the guanine moiety on the 5′-side, while the remainder of the phenoxazine ring system covers the entire base. However, as is apparent from *Fig. 5*, stacking to the 3′-adjacent base is not affected by incorporation of the modified heterocycle. Incorporation of a phenoxazine ring alone into an oligonucleotide has been shown to increase duplex stability by 2–7° per modification. The effect is more prevalent when several phenoxazine rings are clustered together on the same strand, clearly illustrating the contribution of the extended ring system alone [13].

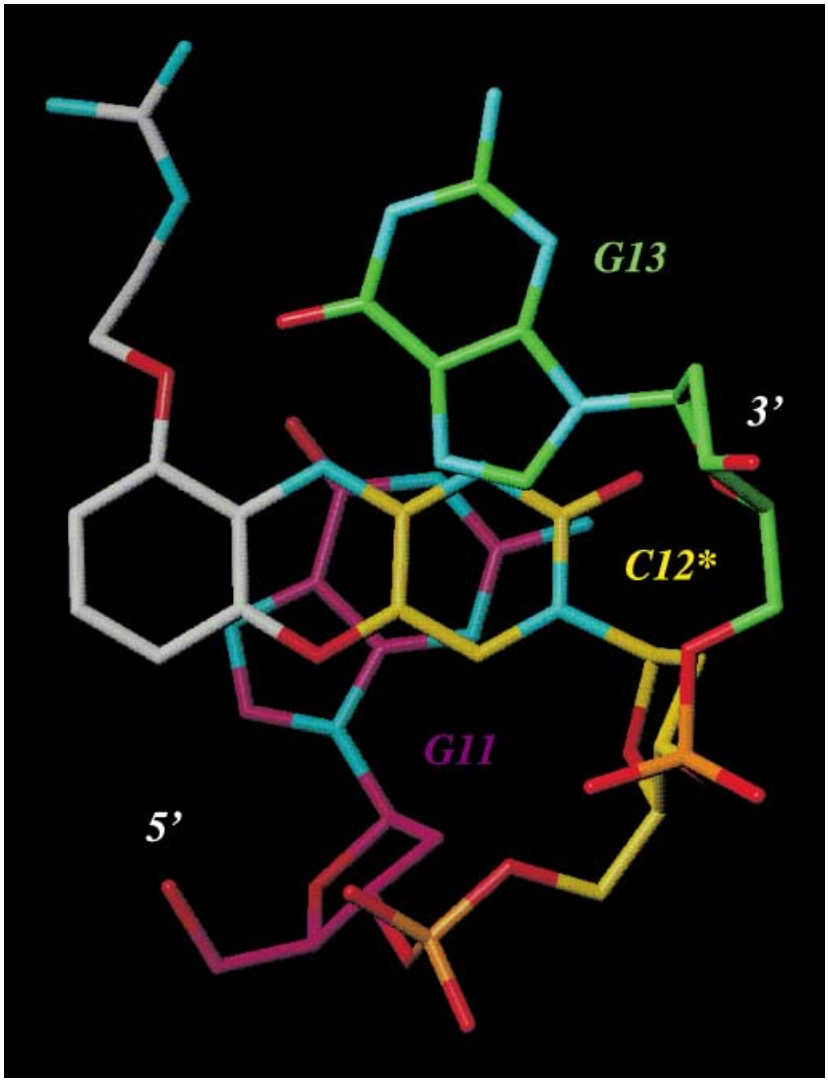


Fig. 5. *The guanidinium G-clamp C12* viewed approximately along the normal to the phenoxazine plane, illustrating the improved stacking interactions with the 5′-adjacent residue* (G11). However, stacking between C12* and the 3′-adjacent residue (G13) remains unaltered. C-Atoms of the 'cytosine' core of the guanidinium G-clamp are colored in yellow.

The various helical parameters reveal little difference from the reference structure, except for notable deviations in the buckle, propeller twist, and base-opening angles for base pairs C2*/G19 and C12*/G9 (*Table 4*). A possible reason for this could be the relatively short contact between the O-atom linking the phenoxazine ring to the guanidinium tether of the guanidinium G-clamp and the O(6) atom of G (*Figs. 3* and *6*). The O ··· O distances observed in the C2*/G19 and C12*/G9 base pairs are 2.98 and 3.12 Å, respectively.

Table 4. *Selected Helical Parameters and Deviations from the Reference Duplex* [11] *in Parentheses*

(Global Base-Base)	Shear S_x [Å]	Stretch S_y [Å]	Stagger S_z [Å]	Buckle κ [deg]	Propeller ω [deg]	Opening σ [deg]
G(1)-C(20)	0.1 (– 0.5)	0.3 (– 0.4)	0.3 (– 0.2)	8.6 (– 8.4)	– 0.9 (– 12.8)	2.5 (– 2.4)
C(2)-G(19)*	*0.6 (– 0.6)*	*0.8 (– 0.6)*	*– 0.3 (0.4)*	*12.1 (– 9.3)*	*– 8.9 (– 9.3)*	*9.8 (– 7.4)*
G(3)-C(18)	– 0.2 (– 0.2)	– 0.4 (– 0.5)	0.1 (– 0.2)	– 4.4 (– 6.2)	– 13.6 (1.9)	4.9 (– 3.8)
T(4)-A(17)	– 0.1 (– 0.0)	0.1 (0.1)	– 0.2 (0.1)	– 3.7 (0.5)	– 14.7 (– 5.5)	1.5 (1.7)
A(5)-T(16)	0.2 (– 0.1)	– 0.1 (0.1)	0.0 (0.1)	– 3.4 (1.6)	– 15.5 (– 1.8)	– 1.5 (1.7)
T(6)-A(15)	– 0.3 (0.3)	0.7 (– 0.1)	0.2 (– 0.1)	1.3 (5.5)	– 15.3 (– 4.6)	10.3 (– 1.1)
A(7)-T(14)	0.0 (0.2)	0.5 (0.0)	– 0.1 (0.1)	3.4 (0.4)	– 6.7 (– 5.1)	6.0 (0.5)
C(8)-G(13)	– 0.1 (0.4)	0.5 (– 0.3)	– 0.1 (– 0.1)	3.6 (3.2)	– 9.3 (– 4.1)	5.9 (– 2.0)
G(9)-C(12)*	*– 0.4 (0.2)*	*0.7 (– 0.6)*	*– 0.3 (0.3)*	*– 10.1 (7.9)*	*– 13.4 (2.3)*	*7.9 (– 5.9)*
C(10)-G(11)	– 0.3 (0.6)	0.6 (– 0.7)	0.3 (0.0)	– 6.5 (4.4)	4.1 (– 4.8)	7.2 (– 6.5)
(Local Inter-Base Pair)	Shift D_x [Å]	Slide D_y [Å]	Rise D_z [Å]	Tilt τ [deg]	Roll ρ [deg]	Twist π [deg]
G(1)/C(20)-C(2)/G(19)*	*– 0.1 (0.5)*	*– 1.9 (0.0)*	*3.2 (0.0)*	*0.6 (– 0.5)*	*1.7 (– 2.0)*	*40.2 (– 3.4)*
C(2)/G(19)-G(3)/C(18)*	*– 0.5 (0.7)*	*– 2.4 (– 0.1)*	*3.7 (0.0)*	*– 4.1 (5.7)*	*3.7 (1.2)*	*24.9 (3.6)*
G(3)/C(18)-T(4)/A(17)	– 1.2 (0.1)	– 1.9 (– 0.1)	3.2 (0.0)	– 1.1 (0.2)	6.2 (0.3)	36.7 (– 3.2)
T(4)/A(17)-A(5)/T(16)	0.6 (0.1)	– 1.8 (0.0)	3.4 (0.1)	2.0 (– 0.8)	19.6 (0.8)	27.1 (– 0.8)
A(5)/T(16)-T(6)/A(15)	1.0 (– 0.2)	– 1.7 (0.1)	3.2 (– 0.1)	0.4 (0.8)	5.0 (– 0.4)	31.6 (– 0.1)
T(6)/A(15)-A(7)/T(14)	– 0.3 (0.0)	– 1.7 (0.0)	3.2 (0.1)	0.5 (– 0.5)	13.9 (0.4)	31.5 (1.5)
A(7)/T(14)-C(8)/G(13)	0.6 (– 0.4)	– 2.7 (0.7)	3.4 (– 0.1)	0.6 (1.1)	3.0 (0.7)	25.8 (5.6)
C(8)/G(13)-G(9)/C(12)*	*– 0.5 (0.2)*	*– 2.1 (– 0.1)*	*3.6 (– 0.1)*	*3.1 (– 5.4)*	*10.3 (0.5)*	*32.0 (– 2.8)*
G(9)/C(12)-C(10)/G(11)*	*0.7 (– 0.7)*	*– 1.8 (– 0.6)*	*3.3 (0.2)*	*0.6 (– 0.4)*	*4.0 (– 1.8)*	*32.2 (1.1)*

H_2O Structure around Guanidinium G-Clamps. In the A-form duplex, positively charged guanidinium groups and nonbridging O(2)P phosphate O-atoms from opposite strands are relatively closely spaced. The distances between imino N-atoms and O-atoms range from 4.5 to 7.6 Å. Although these distances are too long for formation of salt bridges, single H_2O or pairs of H_2O molecules can bridge the two moieties (*Fig. 6*). In the case of base pair G9/C12*, the H_2O structure is more symmetrical (*Fig. 6, a*). There, a H_2O molecule is H-bonded to both imino N-atoms of the guanidinium group. Single H_2O molecules then link it to the phosphate groups of C8 and G9. With base pair C2*/G19, a single H_2O molecule bridges the inner imino N-atom of the guanidinium moiety to the phosphate group of C18 (*Fig. 6, b*). This H_2O is part of a tandem of solvent molecules that link the same imino N-atom to the phosphate group of G19.

In conclusion, the guanidinium G-clamp modification is a heterocyclic modification that has shown specificity for guanosine and forms duplexes of significantly enhanced thermal stability. The presence of the positively charged guanidinium group may also be very useful for improving the solubility of certain oligonucleotide analogues such as PNA [14][15]. Furthermore, the guanidinium group may also improve the nuclease resistance of oligonucleotides like its amino precursor [6][16] and cellular permeation of antisense oligonucleotides and PNA [15].

Financial support from the *National Institutes of Health* (Grant GM55237 to *M. E.*) is gratefully acknowledged. *C. J. W.* was the recipient of a *NSERC* (*Natural Sciences and Engineering Research Council of Canada*) post-doctoral fellowship. We thank Drs. *George Minasov* and *Valentina Tereshko* for assistance with data collection, and *Guity Balow* and *Robert H. Springer* for synthetic intermediates. Portions of this work were performed at the *DuPont-Northwestern-Dow Collaborative Access Team* (DND-CAT) *Synchrotron Research Center*, located at Sector 5 of the *Advanced Photon Source*. DND-CAT is supported by the *E.I. DuPont de Nemours & Co.*, *The Dow Chemical Company*, the *U.S. National Science Foundation* through Grant DMR-9304725, and the State of Illinois through the *Department of Commerce* and the *Board of Higher Education*

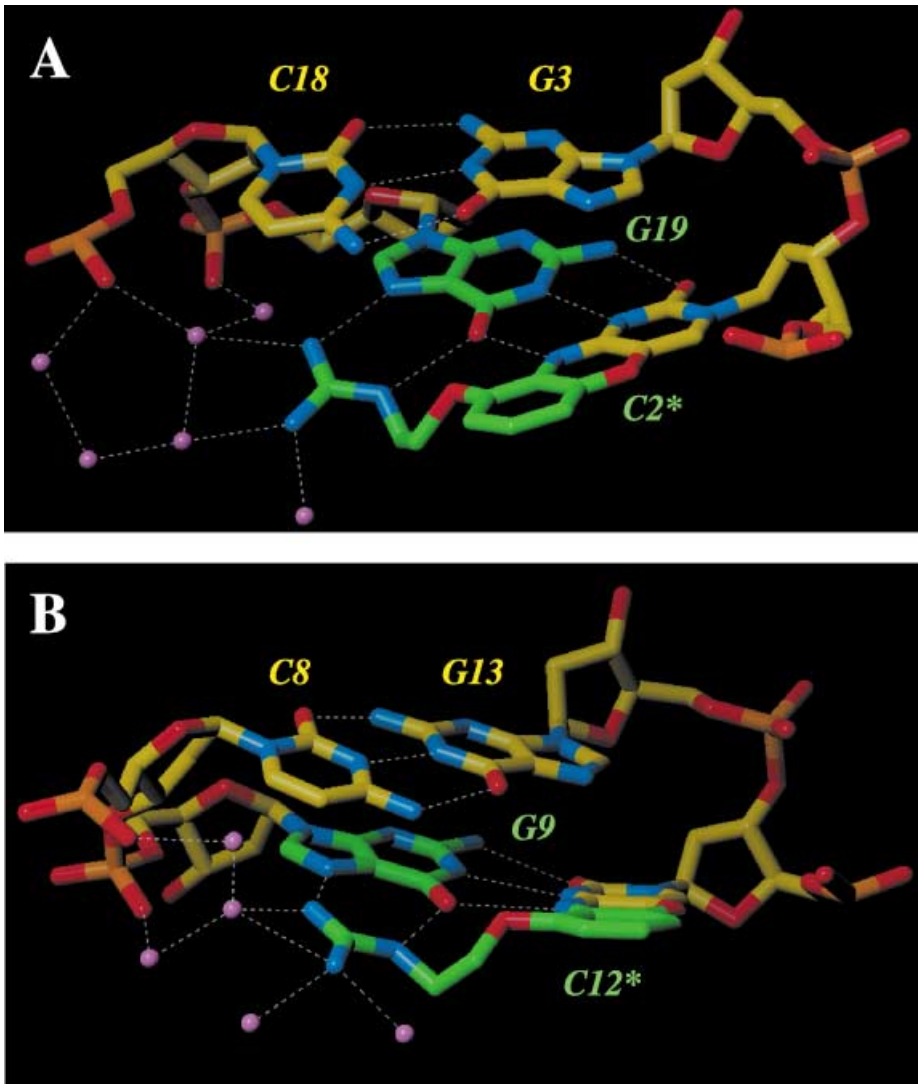


Fig. 6. *H_2O Molecules mediating H-bonding networks between the guanidinium G-clamp and phosphate groups from the opposite strand at residue C2** (A) *and C12** (B). H_2O Molecules are depicted as purple spheres and H-bonds are indicated as dashed lines.

Grant IBHE HECA NWU 96. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Energy Research under Contract No. W-31-102-Eng-38.

Experimental Part

Synthesis and Purification of G-Clamp and Guanidinium G-Clamp-Containing Oligonucleotides. Solid-phase synthesis of G-clamp-containing oligonucleotides was performed on an *Applied Biosystems* (*Perkin-Elmer Corp.*) DNA/RNA synthesizer *380B* with *N*(4)-Ac-protected dC monomers and standard phosphor-

amidite chemistry. Cleavage and deprotection of the oligonucleotides was carried out with a soln. of 40% aq. $MeNH_2$ / 28–30% aq. NH_3 1:1) at r.t. for 4 h. The oligonucleotides were purified DMT-on by reversed-phase HPLC with a *306 Piston Pump System*, an *811C Dynamic Mixer*, a *170 Diode-Array Detector*, and a *215 Liquid Handler* together with the *Unipoint Software* from *Gilson* (Middleton, Wi). After chromatographic purification, the oligonucleotides were detritylated (aq. AcOH, standard procedure), lyophilized, and stored at $-20°$.

For postsynthetic guanidinylation, a 1.0M soln. of 1*H*-pyrazole-1-carboxamidine in 1.0M aq. Na_2CO_3 soln. was prepared; 1 ml of the soln. was added per 0.5 μmol of the purified G-clamp oligonucleotide, and the mixture was kept for 3 h at r.t. For the self-complementary sequences ON-1 and ON-3, the reaction temp. was increased to 55°, and the reaction times were extended to 12 h. Subsequently, the samples were purified by size-exclusion chromatography (SEC) on a *Sephadex G25* column, with H_2O as the eluent, followed by reversed-phase HPLC. The oligonucleotides were analyzed by capillary gel chromatography and electrospray ionization mass spectrometry.

X-Ray Crystal-Structure Determination of the Guanidinium G-Clamp-Modified Decamer. Crystals of the decamer duplex were grown with the hanging-drop-vapor-diffusion technique, with the *Nucleic Acid Miniscreen* from *Hampton Research* (Laguna Niguel, CA). Each drop consisted of 1 μl of oligonucleotide (2.4 mM) and 1 μl of the appropriate screening soln., and was equilibrated against a reservoir of 1 ml of 35% 2-methylpentane-2,4-diol (MPD). A number of conditions yielded perfect, diffraction-quality crystals within a week. The crystal used in this study was grown according to condition 18 of the miniscreen (5% MPD, 20 mM Na cacodylate pH 6.0, 6 mM spermine·4 HCl, 40 mM NaCl, 6 mM KCl, and 10 mM $MgCl_2$). The crystal (dimensions: $0.7 \times 0.2 \times 0.2$ mm) was picked up from the 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 beam line ($\lambda = 0.978$ Å) of the *DND-CAT* at the *Advanced Photon Source*, Argonne, IL, with a MARCCD detector. Data were integrated and merged with DENZO/SCALEPACK [17]. The structure was determined by the molecular-replacement method with the DNA decamer described in [11] as the initial model and refined with the programs CNS [18] and SHELXL-97 [19]. Helical parameters and backbone torsion angles were calculated with the program CURVES [20]. Final coordinates and structure factors have been deposited in the *Protein Data Bank* [21] (PDB code 1KGK) and the *Nucleic Acid Database* [22] (NDB code AD0025).

REFERENCES

[1] S. T. Crooke, *Methods Enzymol.* **2000**, *313*, 3.
[2] M. Manoharan, *Antisense Nucleic Acid Drug Development* **2002**, *12*, 103.
[3] S. M. Freier, K.-H. Altmann, *Nucleic Acids Res.* **1997**, *25*, 4429.
[4] K.-Y. Lin, M. D. Matteucci, *J. Am. Chem. Soc.* **1998**, *120*, 8531.
[5] W. M. Flanagan, J. J. Wolf, P. Olson, D. Grant, K. Lin, R. W. Wagner, M. D. Matteucci, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3513.
[6] M. A. Maier, J. M. Leeds, G. Balow, R. H. Springer, R. Bharadwaj, M. Manoharan, *Biochemistry* **2002**, *41*, 1323.
[7] M. A. Maier, I. Barber-Peoc'h, M. Manoharan, *Tetrahedron Lett.* **2002**, *43*, 7613.
[8] C. J. Wilds, M. A. Maier, V. Tereshko, M. Manoharan, M. Egli, *Angew. Chem.* **2002**, *41*, 115.
[9] M. S. Bernatowicz, Y. Wu, G. R. Matsueda, *J. Org. Chem.* **1992**, *57*, 2497.
[10] V. Tereshko, S. Portmann, E. C. Tay, P. Martin, F. Natt, K.-H. Altmann, M. Egli, *Biochemistry* **1998**, *37*, 10626.
[11] 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. Nucl. Acid Sci. Sect.* **1998**, *48*, 234.
[12] N. P. Pavletich, C. O. Pabo, *Science* **1991**, *252*, 809.
[13] K.-Y. Lin, R. J. Jones, M. D. Matteucci, *J. Am. Chem. Soc.* **1995**, *117*, 3873.
[14] K. G. Rajeev, M. A. Maier, E. A. Lesnik, M. Manoharan, *Org. Lett.* **2002**, *4*, 4395.
[15] a) J. G. Karras, M. A. Maier, T. Lu, A. Watt, M. Manoharan, *Biochemistry* **2001**, *40*, 7853; b) P. Sazani, S. H. Kang, M. A. Maier, C. Wei, J. Dillman, J. Summerton, M. Manoharan, R. Kole, *Nucleic Acids Res.* **2001**, *29*, 3965; c) P. Sazani, F. Gemignani, S. H. Kang, M. A. Maier, M. Manoharan, M. Persmark, D. Bortner, R. Kole, *Nat. Biotech.* **2002**, *20*, 1228.
[16] M. Teplova, S. T. Wallace, V. Tereshko, G. Minasov, A. M. Symons, P. D. Cook, M. Manoharan, M. Egli, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14240.
[17] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307.

[18] 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* **1988**, *54*, 905.
[19] G. M. Sheldrick, T. R. Schneider, *Methods Enzymol.* **1997**, *277*, 319.
[20] a) R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1988**, *6*, 63; b) R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *7*, 655.
[21] Web address: http://www.rcsb.org.
[22] Web address: http://ndbserver.rutgers.edu.

Received January 17, 2003