DNA microstructural requirements for neocarzinostatin chromophore-induced direct strand cleavage

Seok Ho Lee, Jeanne O.Thivierge and Irving H.Goldberg

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

Received February 16, 1989; Revised and Accepted June 22, 1989

ABSTRACT

The microstructural requirements for optimal interaction of neocarzinostatin chromophore (NCS-C) with DNA have been investigated using a series of hexadeoxyribonucleotides with modified bases such as O⁶-methyl G (MeG), I, 5-methyl C (MeC), U, or 5-Bromo U (BrU) at specific sites in its preferred trinucleotide 5'GN_aN_b3':5'N_aN_bC3' (N_a = A, C, or T). Results show that MeG:C and G:MeC in place of G:C improve direct strand cleavage at the target N_b (N_b = T > A >> C > G), whereas MeC:G and C:MeG in place of N_a:N_b hinder cleavage. The optimal base target at N_b appears to be determined by its ability to form T:A type base pairing instead of C:G type. The observed differences in DNA strand cleavage patterns can be rationalized by induced changes in target site structure and are compatible with a model for NCS-C:DNA interaction in which the naphthoate moiety intercalates between 5'GN_a3', and the activated tetrahydro-*s*-indacene, lying in the minor groove, abstracts a hydrogen atom from C-5' of N_b.

INTRODUCTION

Recent studies on the mode of interaction of the chromophore (NCS-C) of the antitumor antibiotic neocarzinostatin with DNA, employing a series of synthetic oligodeoxyribonucleotides (hexanucleotides), has led us to the identification of the preferred trinucleotide 5'GN_aN_b3' (N_a = A, C, or T and N_b = T > A >> C > G), where NCS-C appears to intercalate in-between GN_a and cleave at N_b (1). NCS-C (2-5) consists of three main structural subunits: a 2-hydroxy-7-methoxy-5-methyl naphthoic acid, a *N*-methyl- α -D-fucosamine, and a bicyclo[7.3.0]dodecadienediyne epoxide (Figure 1 (A)). The substituted naphthoic acid moiety of NCS-C intercalates DNA via the minor groove, and the positively charged amino sugar (*N*-methyl- α -D-fucosamine) moiety is presumed to provide electrostatic interaction (6). The

bicyclo[7.3.0]dodecadienediyn e poxide moiety is then juxtaposed along the minor groove, where it is converted by thiol into its active form (7). NCS-C produces O₂-dependent single strand breaks by selective abstraction of a hydrogen atom from the C-5' of deoxyribose (8) of mainly T residues (T > A >> C > G) (7) which leads to the C-5' aldehyde (9,10), or it forms O₂-independent stable sugar adducts (11). NCS-C also generates abasic sites, especially at the C of 5'AGC3' (12-15). The two types of damage, direct strand cleavage and abasic lesion, occur two nucleotides apart in opposite strands (14,15). However, not all of the direct strand cleavage at the T of 5'GCT3' on 5'GCT3':5'AGC3' is accompanied by the abasic lesion at the C of 5'AGC3'.

We have now investigated a series of synthetic hexanucleotides of designed sequence, containing single modified bases at specific sites either on the strand under observation (plus strand) or on the complementary strand (minus strand), to pinpoint the exact structural requirements of the preferred trinucleotides for NCS-C interaction, binding, and cleavage. Due to the possibility of altered hydrogen-bonding interaction between bases which contain a modified functionality and of altered vertical stacking of bases induced by the incorporation of partially mismatched base pairs, which might destabilize or modify the double helix (16,17), it was expected that other parameters, such as roll, slide, twist, and rise of base pairs, would possibly be affected. Despite these potential local alterations, it has been suggested that even non-complementary base pairs can be accommodated in the double helix with little perturbation of global structure (16,17). Also, these hexanucleotides likely have a B-DNA structure under the experimental conditions employed (18-20).

The binding of NCS-C was anticipated to occur at or around its usual site of action, due not only to the overwhelming specificity of the target site (T > A >> C > G) but also to the limited number of potential intercalation sites in the hexanucleotides, making it possible to directly correlate the consequences derived from specific site modifications on the bases.

MATERIALS AND METHODS

All procedures involving NCS-C were carried out in the dark at 0 °C. NCS-C was normally stored in methanol buffered with 20 mM sodium citrate-citric acid, pH 4.5 at -70 °C in brown vials until use. NCS-C was prepared as described (1).

Oligodeoxyribonucleotides

A series of hexanucleotides containing single modified bases at specifically designated sites was synthesized as described (1). The modified mononucleotide phosphoramidites employed in the synthesis were O⁶-methyl deoxyguanosine (MeG), 5-methyl deoxycytidine (MeC), deoxyinosine (I), deoxyuridine (U) (American Bionetics), and 5-bromodeoxyuridine (BrU; MilliGen). The hexanucleotides were purified by reversed phase high performance thin layer chromatography and labelled at the 5'-end with [γ -³²P]-ATP, purified further by gel electrophoresis, and sequenced, as described (1).

Nomenclature

The plus and minus strands of the *preferred trinucleotide* are referred to as 5'GN_aN_b3' and 5'GN_aN_b3', respectively. All hexanucleotides are designated as 5'N₁-N₆3' for both strands.

Aerobic, Single Strand Cleavage by NCS-C

A typical reaction mixture contained the followings in 50 μ l: 10 μ M duplex hexanucleotide (2000 Bq), 60 μ M NCS-C in 10 μ l methanol, thiol (either 10 mM dithiothreitol (DTT) or 20 mM glutathione (GSH)), 5 mM sodium citrate-citric acid, pH 4.5, 50 mM Tris-HCl, pH 8.0. Equimolar amounts of unlabelled corresponding complementary strands were mixed in the presence of Tris-HCl, pH 8.0, heated at 65 °C for 5 min, and then allowed to cool down slowly to 25 °C in the same bath. To ensure that all of the labelled hexanucleotide strand under observation was double stranded, it was annealed to an excess of its unlabelled complementary strand in the presence of Tris-HCl, pH 8.0, separately, as described above. The two duplexes, labelled and unlabelled, were then mixed. After the double stranded hexanucleotide was cooled to 0 °C for 10 min, a thiol, either DTT or GSH, was added. Finally, NCS-C was added to initiate the reaction at 0 °C in the dark. The reaction was allowed to proceed for 30 min. A control reaction without NCS-C was carried out under the same conditions. The reaction mixture was then lyophilized and taken up in the gel loading buffer for electrophoresis. Reactions were done in duplicates and repeated two to four times.

Sequencing gel electrophoresis and densitometry of autoradiographic films were as described (1).

Statistical Analysis

Mean values of the data obtained from densitograms were used in the Tables. Standard deviations were within 10 % for cleavage of more

Table I. Modified Bases at G:C Base Pair of Trinucleotide 5'GCT3':5'AGC3'.

Strand				Per Cent Cut			
Plus		Minus		N ₃		N ₄	
				DIT	GSH	DIT	GSH
G	GCTGC	GCAGC	C	0.6	0.9	25.5	58.7
G	ICTGC	GCAGC	C		1.5	19.0	49.5
G ^{Me}	GCTGC	GCAGC	C		1.3	27.0	74.9
G	ACTGC	GCAGT	C		1.2	18.7	55.3
G ^{Me}	GCTGC	GCAGT	C			15.2	58.4
GG	CTGC	GCAG	CC	0.6	0.9	25.5	58.7
GG	CTGC	GCAG ^{Me}	CC	1.1	0.8	30.1	72.2
GA	CTGC	GCAG	TC		1.2	18.7	55.3
GA	CTGC	GCAG	UC		1.7	18.2	55.6
GA	CTGC	GCAG ^{Br}	UC		1.5	16.0	59.8

Hexanucleotides are presented with the 5'-end at the left. Plus strands are labelled strands under observation and minus strands are unlabelled complementary strands. N_n in Per Cent Cut (mean values) indicates the n-th base as a target on the plus strand from the 5'-end. Modified bases are located on either G of 5'GCT3' or C of 5'AGC3'.

than 5 %. However, standard deviations for cleavage of less than 2 % were somewhat variable in a range of 5 to 40 %. The software applications used were Excel (Microsoft) and Data Desk Professional (Odesta).

RESULTS

A series of hexanucleotides was prepared containing single modified bases at specific locations on either the strand under investigation (plus strand) or its complementary strand (minus strand). The NCS-C-induced direct strand cleavage patterns were analyzed to identify the microscopic structural requirements within the preferred trinucleotides 5'GNT3':5'ANC3'. The hexanucleotide chosen as the unmodified reference (control) was 5'GGCTGC3':5'GCAGCC3', where 5'GCT3':5'AGC3' served as the preferred trinucleotide base pair. Other

Table II. Modified Bases at C:G Base Pair of Trinucleotide 5'GCT3':5'AGC3'.

Strand				Per Cent Cut			
Plus		Minus		N ₃		N ₄	
				DTT	GSH	DTT	GSH
GG	CTGC	GCAG	CC	0.6	0.9	25.5	58.7
GG ^{Me}	CTGC	GCAG	CC		0.4	25.3	48.5
GG	TTGC	GCAA	CC	6.1	10.8	37.2	58.8
GG	UTGC	GCAA	CC	4.0	7.7	28.2	60.1
GG ^{Br}	UTGC	GCAA	CC	5.2	12.8	25.4	53.9
GGC	TGC	GCA	GCC	0.6	0.9	25.5	58.7
GGC	TGC	GCA ^{Me}	GCC	2.3	1.2	14.8	28.7
GGC	TGC	GCA	ICC	1.2	1.4	49.9	82.3
GGT	TGC	GCA	ACC	6.1	10.8	37.2	58.8
GGT	TGC	GCA ^{Me}	GCC	2.1	3.8	23.3	37.4

Analysis is as described in Table I. Modified bases are located on either C of 5'GCT3' or G of 5'AGC3'.

hexanucleotides contained a modified nucleotide at a specific position on either the plus or minus strand. Both strands of the duplexes have been tested, although the strand 5'GCAGCC3' with or without modified bases did not in general possess reasonably good target sites for NCS-C-induced direct strand cleavage.

Thiols

Previously, it was determined that GSH was generally more efficient than DTT, although they showed the same strand selectivity and sequence specificity (1). To ensure that this property also holds for hexanucleotides containing modified bases, both thiols were employed in this study. As shown in parallel in the Tables, results indicate a similar trend for both thiols. Therefore, only the data on GSH are discussed.

Effect of Modified Bases at the Base Pair G:C of the Preferred Trinucleotide 5'GCT3':5'AGC3'

This group of hexanucleotides (Table I) contains modified (substituted) bases at the base pairs N₂:N₅; specifically, the modified

bases are placed at G:C base pairs, either at G (^{Me}G, I, or A) of 5'*GCT*3' or at C (^{Me}C, T, U, or BrU) of 5'*AGC*3'. The most noticeable results are observed for the hexanucleotides 5'*G*^{Me}*GCTGC*3':5'*GCAGCC*3' and 5'*GGCTGC*3':5'*GCAG*^{Me}*CC*3', as these hexanucleotides showed distinctly elevated direct strand cleavage at T compared with the unmodified hexanucleotide 5'*GGCTGC*3':5'*GCAGCC*3'. On the other hand, the hexanucleotide 5'*GICTGC*3':5'*GCAGCC*3' showed somewhat depressed cleavage. The hexanucleotide 5'*GACTGC*3':5'*GCAGNC*3' whose A on the plus strand was paired with T, U, or BrU, however, did not show recognizable differences in cleavage regardless of the nature of modified bases at the N on the minus strand (Table I).

Effect of Modified Bases at the Base Pair C:G of the Preferred Trinucleotide 5'*GCT*3':5'*AGC*3'

This group of hexanucleotides contains modified base pairs N₃:N₄; specifically, this group possesses modified bases at either C (^{Me}C, T, U, or BrU) of 5'*GCT*3' on the plus strand or G (^{Me}G, I, or A) of 5'*AGC*3' on the minus strand (Table II). The results were quite dramatic. The hexanucleotide 5'*GGCTGC*3':5'*GCAICC*3' containing C:I in place of C:G at 5'*GCT*3':5'*AGC*3' to give 5'*GCT*3':5'*AIC*3' turned out to be the best within the series tested in NCS-C-induced direct strand cleavage at T. On the other hand, the hexanucleotide 5'*GGCTGC*3':5'*GCA*^{Me}*GCC*3' containing C:^{Me}G was by far the worst within the series tested. However, the hexanucleotide 5'*GG*^{Me}*C**TGC*3':5'*GCAGCC*3' containing ^{Me}C:G produced only moderately decreased direct strand cleavage. Also, the substitution of C of 5'*GCT*3' with T, U, or BrU, did not appear to cause much perturbation, when compared with the result with the unmodified hexanucleotide counterpart 5'*GGCTGC*3':5'*GCAGCC*3'.

Therefore, the order of preferred base pair at N₃:N₄ (C:G of 5'*GCT*3':5'*AGC*3'), which was found as C:I (82.3 %) > T:A (58.8 %) = U:A (60.1 %) = C:G (58.7 %) > BrU:A (53.9 %) = ^{Me}C:G (48.5 %) > C:^{Me}G (28.7 %), clearly indicates that steric factors dominate at this position rather than electronic factors that seem to be rather subtle.

Effect of Modified Bases at the Target Base Pair on Either Strand

The hexanucleotides containing modified bases at the base pair N₄:N₃, i.e., T of 5'*GCT*3' (T, U, BrU, C, or ^{Me}C) on the plus strand or A of 5'*AGC*3' (A, G, ^{Me}G, or I) on the minus strand also revealed interesting results (Table III). The noticeably enhanced direct strand cleavage at U on 5'*GGCUGC*3':5'*GCAGCC*3' and BrU on 5'*GGCBrUGC*3':5'*GCAGCC*3', compared

Table III. Modified Bases at T:A Base Pair of Trinucleotide 5'GCT3':5'AGC3'

Strand				Per Cent Cut			
Plus		Minus		N ₃		N ₄	
				DTT	GSH	DTT	GSH
GGC	TGC	GCA	GCC	0.6	0.9	25.5	58.7
GGC	UGC	GCA	GCC	0.0	0.8	22.8	71.1
GGC	^{Br} UGC	GCA	GCC		1.1	35.3	73.1
GGC	CGC	GCG	GCC		1.2	0.5	21.0
GGC	^{Me} CGC	GCG	GCC		1.2	2.1	25.7
GGCT	GC	GC	AGCC	0.6	0.9	25.5	58.7
GGCT	GC	GC	^{Me} GGCC	1.3	2.0	17.0	28.3
GGCC	GC	GC	GGCC		1.2	0.5	21.0
GGCC	GC	GC	^{Me} GGCC		0.7	1.5	10.7
GGCC	GC	GC	IGCC		0.5	7.3	40.8

Analysis is as described in Table I. Modified bases are located on either T of 5'GCT3' or A of 5'AGC3'.

with T on 5'GGCTGC3':5'GCAGCC3', seems to indicate that electronic factors may play a role here. The difference among the three bases, T, ^{Br}U, and U, is the presence of an electron-donating substituent (a methyl group on T), an electron-withdrawing substituent (a bromine atom on ^{Br}U), or no such substituent (U). The C of 5'GCC3' is a poor target as shown in both 5'GGCCGC3':5'GCGGCC3' and 5'GGC^{Me}CGC3':5'GCGGCC3'. Although the mechanism of the slight increase in cleavage in the latter upon the substitution of the target C with ^{Me}C is not clear, it could be due to its resemblance with T, the best target.

The two hexanucleotides, 5'GGCCGC3':5'GC^{Me}GGCC3' and 5'GGCCGC3':5'GCIGCC3', provide the best examples of what the optimal target site configuration for NCS-C-induced direct strand cleavage should be. The hexanucleotide 5'GGCCGC3':5'GC^{Me}GGCC3', whose target C remained at the base pair C:^{Me}G, was cleaved very poorly, whereas 5'GGCCGC3':5'GCIGCC3', whose target C was located on the base pair C:I, was attacked almost as well as 5'GGCTGC3':5'GCAGCC3' with the best usual target T.

Direct Strand Cleavage on Minus Strands

Direct cleavage of the strand containing 5'AGC3' has also been studied. In the hexanucleotides tested, target A (A, G, MeG, or I) belongs to the trinucleotide 5'GCA3', where NCS-C intercalates in between GC and cleaves at A. However, the C (C, MeC, T, U, BrU) of the trinucleotide 5'AGC3' becomes the target site where NCS-C is supposed to intercalate in between AG and cleave at C. This is of particular interest because the C of 5'AGC3' is also the most specific target site for the NCS-C-induced alkali-labile, abasic lesion.

Although independent intercalative binding of two NCS-C molecules, through the 2-hydroxy-7-methoxy-5-methyl naphthoic acid moiety, at the 5'GC3':5'GC3' of 5'GCT3':5'AGC3' (5'GGCTGC3':5'GCAGCC3') and at the 5'GC3':5'GC3' of 5'GCA3':5'TGC3' (5'GCAGCC3':5'GGCTGC3') in the same hexanucleotide duplex should not be mutually exclusive (the two sites are three base pairs apart), the two NCS-C molecules would have to share the same target base pair T:A, where either T or A, respectively, is the target. This means that the participation of two tetrahydro-s-indacene diradicals (activated bicyclo[7.3.0]dodecadienediyl epoxide) on two separate NCS-C molecules, oriented in opposite directions, in direct strand cleavage could become mutually exclusive to a degree, if not absolutely. This argument is supported by the results showing the inverse relationships between the two targets, T and A, on respective complementary strands. This holds particularly well in hexanucleotides with modified base pairs at the C:G of 5'GCT3':5'AGC3' (Table II and IV (B)) or the T:A of 5'GCT3':5'AGC3' (Table III and IV (A)). For example, the T of 5'GGCTGC3':5'GCAMeGCC3' was the worst target (Table II) within its group, whereas the A of 5'GCAMeGCC3':5'GGCTGC3' was the best target within its group (Table IV (B)). Also, the T of 5'GGCTGC3':5'GCAICC3' was the best target within its group (Table II), whereas the A of 5'GCAICC3':5'GGCTGC3' was one of the poorest within its group (Table IV (B)).

Table IV (A) shows hexanucleotides with modified bases at the base pair N₃:N₄; i.e., A of 5'GCA3' or 5'AGC3' (A, G, MeG, or I) on plus strands or T of 5'TGC3' or 5'GCT3' (T, U, BrU, C, or MeC) on minus strands. The MeG of 5'GCMeGGCC3':5'GGCCGC3' was a far inferior target, as expected. However, the I of 5'GCIGCC3':5'GGCCGC3' was a better target than the A of 5'GCAGCC3':5'GGCTGC3'.

Table IV (B) illustrates a group of hexanucleotides containing

Table IV. Direct Strand Cleavage at AGC Containing Strands.

Strand				Per Cent Cut					
Plus		Minus		N ₃		N ₄		N ₅	
				DTT	GSH	DTT	GSH	DTT	GSH
A	GC	AGCC	GGCT	GC	3.2	9.0		1.0	9.0
	GC ^{Me}	GGCC	GGCT	GC	1.6	1.9		2.9	6.3
	GC ^{Me}	GGCC	GGCC	GC	1.2	2.6			3.2
	GC	IGCC	GGCC	GC	4.0	17.0		0.7	4.6
	GCA	GCC	GGC	TGC	3.2	9.0		1.0	9.0
	GCA	GCC	GGC	UGC	2.3	6.8		1.4	10.2
	GCA	GCC	GGC ^{Br}	UGC	2.3	6.5		1.7	2.3
	GCG	GCC	GGC ^{Me}	CGC					6.7
	GCG	GCC	GGC	CGC					4.9
	GCA	GCC	GGC	TGC	3.2	9.0		1.0	9.0
	GCA	ICC	GGC	TGC	5.1	10.6	1.7	9.2	0.4
	GCA ^{Me}	GCC	GGC	TGC	6.0	21.5		3.1	1.1
	GCA ^{Me}	GCC	GGT	TGC	0.9	3.9	1.3	3.8	14.3
	GCAA	CC	GG	TTGC	1.5	2.4	2.5	7.1	1.3
B	GCAA	CC	GG	UTGC	1.2	2.7	2.1	4.6	1.2
	GCAA	CC	GG ^{Br}	UTGC	2.6	2.2	2.3	5.1	0.2
	GCAG	CC	GG	CTGC	3.2	9.0		1.0	9.0
	GCAG	CC	GG ^{Me}	CTGC	3.3	7.1		1.8	1.0
	GCAG	CC	GG	CTGC	3.2	9.0		1.0	9.0
	GCAG ^{Me}	CC	GG	CTGC	5.8	17.9		2.6	1.9
	GCAG	TC	GA	CTGC	2.1	6.5	1.0	1.9	12.0
	GCAG	UC	GA	CTGC	2.1	8.6			11.6
	GCAG ^{Br}	UC	GA	CTGC	1.6	6.4			14.1
	GCAGT	C	G	ACTGC	2.1	6.5	1.0	1.9	12.0
	GCAGT	C	G ^{Me}	GCTGC	2.6	9.1			2.6
	GCAGC	C	G	GCTGC	3.2	9.0		1.0	9.0
	GCAGC	C	G ^{Me}	GCTGC	3.3	15.8		1.0	0.6
	GCAGC	C	G	ICTGC	2.8	12.5		0.9	5.8
C									31.4

Analysis is as described in Table I.

5'GCA^{3'}:5'TGC^{3'} in 5'GCAGCC^{3'}:5'GGCTGC^{3'} with modified base pair N₄:N₃; i.e., G of 5'GCAGCC^{3'} or C of 5'GGCTGC^{3'}, that are the nearest neighbor bases in the 3'-direction from the target site A of 5'GCA^{3'}:5'TGC^{3'}. Therefore, it was possible to concentrate on the effect of perturbation at this particular location in the performance of the target site A. As shown previously, nearest neighbor bases around the preferred trinucleotides were expected to have appreciable effects on the target site cleavage (1). Since these modified bases are naturally located at the base pair C:G of 5'GCT^{3'}:5'AGC^{3'} on 5'GCAGCC^{3'}:5'GGCTGC^{3'}, where T is a far superior target, the effect of these modified bases on the cleavage at T should have a significant effect on the cleavage at A on the complementary strand as shown with 5'GCAMeGCC^{3'}:5'GGCTGC^{3'} and 5'GCAGCC^{3'}:5'GGCTGC^{3'}.

Table IV (C) contains hexanucleotides where modified bases are at the base pair N₅:N₂; i.e., C of 5'AGC^{3'} (C, MeC, T, U or BrU) on the plus strand or G of 5'GCT^{3'} (G, MeG, A, or I) on the minus strand. In general, the order of preference for the target base was the same in 5'AGC^{3'} as in 5'GCT^{3'}, although the effects by modified bases were rather small within the groups of related bases; for example, C and MeC or T, U, and BrU showed quite insignificant differences within the group. However, the target C of 5'AGC^{3'}:5'GCT^{3'}, where the G of 5'GCT^{3'} on the minus strand was the site of replacement by modified bases such as G, MeG, or I, exhibited a remarkable difference in the degree of direct strand cleavage. The C became almost as good a target as the T of 5'AGT^{3'}:5'ACT^{3'} (5'GCAGTC^{3'}:5'GACTGC^{3'}) when it was paired with I to give 5'AGC^{3'}:5'ICT^{3'} (5'GCAGCC^{3'}:5'GICTGC^{3'}) in place of G as of 5'AGC^{3'}:5'GCT^{3'} (5'GCAGCC^{3'}:5'GGCTGC^{3'}).

Target Base Specificity

The order of target base preference at N₄ on the plus strand, paired with either normal or substituted bases, is as follows (Table V): BrU (73.1 %; BrU:A) ≥ U (71.1 %; U:A) > T (58.7 %; T:A) > C (40.8 %; C:I) > MeC (25.7 %; MeC:G) ≥ C (21.0 %; C:G) > C (10.7 %; C:MeG) >> G (not detectable; G:C).

Also, the data clearly show that I, when paired with C (17.0 %; I:C of 5'GCIGCC^{3'}:5'GGCCGC^{3'}), was a better target than A (9.0 %; A:T of 5'GCAGCC^{3'}:5'GGCTGC^{3'}) at N₃. Notably, C, when paired with I (40.8 %; C:I of 5'GGCCGC^{3'}:5'GCIGCC^{3'}), became almost as good a target as T (58.7 %; T:A of 5'GGCTGC^{3'}:5'GCAGCC^{3'}).

Table V. Target Base Specificity.

Strand				Per Cent Cut					
Plus		Minus		N ₃		N ₄		N ₅	
				DTT	GSH	DTT	GSH	DTT	GSH
GC	AGCC	GGCT	GC	3.2	9.0				
GC ^{Me}	GGCC	GGCT	GC	1.6	1.9				
GC ^{Me}	GGCC	GGCC	GC	1.2	2.6				
GC	IGCC	GGCC	GC	4.0	17.0				
GGC	TGC	GCA	GCC			25.5	58.7		
GGC	UGC	GCA	GCC			22.8	71.1		
GGC ^{Br}	UGC	GCA	GCC			35.3	73.1		
GGC	CGC	GCG	GCC			0.5	21.0		
GGC ^{Me}	CGC	GCG	GCC			2.1	25.7		
GGCT	GC	GC	AGCC			25.5	58.7		
GGCT	GC	GC ^{Me}	GGCC			17.0	28.3		
GGCC	GC	GC	GGCC			0.5	21.0		
GGCC	GC	GC ^{Me}	GGCC			1.5	10.7		
GGCC	GC	GC	IGCC			7.3	40.8		
GCAGT	C	G	ACTGC					12.0	35.9
GCAGT	C	G ^{Me}	GCTGC					2.6	7.2
GCAGC	C	G	GCTGC						9.0
GCAGC	C	G ^{Me}	GCTGC					0.6	12.0
GCAGC	C	G	ICTGC					6.0	31.4

Analysis is as described in Table I. Modified bases fall on the target site base pairs, either on the plus or minus strand.

DISCUSSION

A series of synthetic hexanucleotides of specific sequence containing single modified bases at designated locations on either the plus or minus strand has been studied to elucidate the optimum structural requirements of the preferred trinucleotides in providing

preferential binding sites as well as target sites for NCS-C. Attention has been focused on each of the base pairs G:C, C:G, or T:A of the preferred trinucleotide 5'GCT3':5'AGC3' to determine how NCS-C reacts to the altered environment, both steric and electronic, provided by a particular substitution. Like actinomycin D, which intercalates in between 5'GC3' (21), NCS-C shows distinctive sequence-specific interactions with DNA at purine-(3',5')-pyrimidine sequences, in particular, at its preferred trinucleotides 5'GN_aN_b3' (N_a = A, C, or T), where NCS-C appears to intercalate in between the dinucleotide GN_a and cleaves the C-5' of the nucleotide N_b (N_b = T > A >> C > G) (1).

Structural studies on the base pairing interaction of MeG:C and MeG:T (22,23) indicate that oligonucleotides containing these abnormal bases have a relatively small conformational change from a regular duplex, although the incorporation of MeG creates a region of localized instability in the double helix (24). A proposed base pairing for the C:MeG shows that the MeG base slides somewhat toward the major groove while the C base slides toward the minor groove, resulting in two amino-ring nitrogen hydrogen bonds (22). Even with T:MeG base pairing, perturbations from a regular helix are relatively minor (23). The T and MeG bases in the anti conformation can pair through an amino-carbonyl hydrogen bond at the minor groove edge while the bases open out in the major groove edge to avoid unfavorable methoxy-carbonyl interactions. It has also been reported that G:C base pairs are stable at low temperature when flanked by either C:MeG or T:MeG modification sites (23). Although, the methoxy group at the modification site stacks better with adjacent base pairs in the oligonucleotide duplex containing C:MeG compared to the oligonucleotide duplex containing T:MeG, MeG incorporation into the duplex results in a destabilization of the entire helix independent of the pyrimidine base opposite the modification site. In the case of G:T mismatched base pairs, the mismatches adopt a wobble type structure, and the double helix is able to accommodate this G:T pairing with little distortion of the overall helical conformation (16).

Activated NCS-C, believed to be the thiol-adduct at C-12 of a tetrahydro-*s*-indacene diradical (Figure 1 (B)) (2,5,25), as a result of a concerted bimolecular nucleophilic substitution (*S_N2*), is asymmetric in that the two radicals are situated on opposite sides of the indene ring of the tetrahydro-*s*-indacene diradical. Results indicate that modifications such as MeG:C and G:MeC instead of G:C for the base pair at the 5'-side (on

Figure 1. Neocarzinostatin chromophore (NCS-C) (A) and activated NCS-C containing a thiol-bound tetrahydro-*s*-indacene diradical (B). R¹ and R² represent 2-hydroxy-7-methoxy-5-methyl naphthoic acid and *N*-methyl- α -D-fucosamine, respectively, and GS represents glutathione.

the plus strand) of the intercalation site 5'GC3':5'GC3' of 5'GCT3':5'AGC3' improve the action of NCS-C. Therefore, these analogues appear to enhance NCS-C action either because overall binding becomes firmer or because intercalation through its 2-hydroxy-7-methoxy-5-methyl naphthoic acid moiety at the 3'-side (on the plus strand) of the base pairs MeG:C and G:MeC more efficiently orients the activated, thiol-bound bicyclo[7.3.0]dodecadienediyne epoxide moiety towards the 3'-direction (on the plus strand) along the minor groove, placing one of its two radicals in juxtaposition to the target. On the other hand, any substitution such as MeC:G or C:MeG instead of C:G for the base pair at the 3'-side (on the plus strand) at the intercalation site 5'GC3':5'GC3' of 5'GCT3':5'AGC3' results in a negative impact on the action of NCS-C at the target T of 5'GCT3' (N_b of 5'GN_aN_b3'), due likely to an inappropriate positioning of the tetrahydro-*s*-indacene diradical. The data also imply that the target base specificity of NCS-C, T > A >> C > G, stems from the nature of bases on both strands at N_b of the preferred trinucleotide 5'GN_aN_b3'. More specifically, the target base specificity arises from the nature of hydrogen bonding between the base pair at the target site as well as the entity and spatial arrangements of the functional groups on

the base at the target site. The effectiveness of the target base on the plus strand seems to be mostly determined by its ability to form T:A like rather than C:G like base pair with the base on the minus strand, as shown in the order of target base preference found as ${}^{\text{Br}}\text{U} ({}^{\text{Br}}\text{U}:\text{A}) \geq \text{U} (\text{U}:\text{A}) > \text{T} (\text{T}:\text{A}) > \text{C} (\text{C}:\text{I}) > \text{I} (\text{I}:\text{C}) > \text{A} (\text{A}:\text{T}) > {}^{\text{Me}}\text{C} ({}^{\text{Me}}\text{C}:\text{G}) \geq \text{C} (\text{C}:\text{G}) > \text{C} (\text{C}:\text{MeG}) \gg {}^{\text{Me}}\text{G} ({}^{\text{Me}}\text{G}:\text{C}) \geq \text{G} (\text{G}:\text{C})$. This is because a proper placement of the thiol-bound tetrahydro-*s*-indacene diradical at the target site in the minor groove is crucial to produce predominantly direct single strand breaks at the C-5' of the N_2 of ${}^5\text{GN}_1\text{N}_2{}^3'$. In general, a pyrimidine base, rather than a purine base, is the preferred target on the plus strand.

A striking result is the markedly improved status of C as a target nucleotide (eg., C of ${}^5\text{GCC}{}^3'$) when it is paired with I (C:I in ${}^5\text{GGCCGC}{}^3':{}^5\text{GC/IGCC}{}^3'$), in place of G (C:G in ${}^5\text{GGCCGC}{}^3':{}^5\text{GCGGCC}{}^3'$). The reason why C of C:G and G of G:C are very unfavorable targets appears to be the extra hydrogen bonding on these base pairs between the C-2 amino group on G and the C-2 carbonyl group on C. This is supported by the results with the base pair C:I which lacks these the groups that can participate in the third hydrogen bonding. In addition to these steric factors at the target base pair, there might also be electronic factors which have to be taken into consideration. In fact, U was a better target (71.1 % in ${}^5\text{GGCUGC}{}^3':{}^5\text{GCAGCC}{}^3'$; Table III) than T (58.7 % in ${}^5\text{GGCTGC}{}^3':{}^5\text{GCAGCC}{}^3'$; Table III). Nevertheless, result with ${}^{\text{Br}}\text{U}$ (73.1 % in ${}^5\text{GGC}{}^{\text{Br}}\text{UGC}{}^3':{}^5\text{GCAGCC}{}^3'$; Table III) was quite similar to that with U, although ${}^{\text{Br}}\text{U}$ was closer to T in terms of its sizable substituent. In addition, I as the target nucleotide (I:C in ${}^5\text{GC/IGCC}{}^3':{}^5\text{GGCCGC}{}^3'$) is even better than A (A:T in ${}^5\text{GCAGCC}{}^3':{}^5\text{GGCTGC}{}^3'$).

It can be appreciated that these results should come from the remarkable similarity in base pairing between T:A and and C:I or between A:T and I:C at the target site, given the second base as the target base. The only clear dissimilarity among these base pairs, which could have an effect on base stacking in the duplex structure is the C-5 methyl group. The additional space occupied by the 6-membered ring of a purine base on the plus strand, in addition to the C-2 amino group on G on the target base of the plus strand, might simply hinder proper positioning of the thiol-bound tetrahydro-*s*-indacene diradical.

In spite of the fact that ${}^5\text{GGCTGC}{}^3':{}^5\text{GCA/CC}{}^3'$ (82.3 %), ${}^5\text{GGCTGC}{}^3':{}^5\text{GCAG}{}^{\text{Me}}\text{CC}{}^3'$ (72.2 %), and ${}^5\text{GGCTGC}{}^3':{}^5\text{GCAGCC}{}^3'$ (58.7 %) showed markedly different extents of direct strand cleavage at the T of

Figure 2. Efficiency of the cleavage reaction with different minus strands. 5'GGCTGC3' was used as the plus strand in all three duplexes. The minus strands used were 5'GCAICCC3', 5'GCAGMeCC3', and 5'GCAGCC3', respectively, from top to bottom of the curves. *Per Cent Cut* indicates direct strand cleavage at C-5' of T on 5'GGCTGC3'. *Mole Ratio* indicates the relative mole ratio of NCS-C to hexanucleotide base pairs. A series of reaction mixtures contained GSH and varying amounts of NCS-C with a constant amount (10 μ M) of hexanucleotide, 5'GGCTGC3':5'GCAICCC3', 5'GGCTGC3':5'GCAGMeCC3', or 5'GGCTGC3':5'GCAGCC3'. The ratio of NCS-C to each base pair of hexanucleotide used was varied as 0, 0.01, 0.1, 0.3, 0.6, or 1.0. The direct strand cleavage at the T of 5'GGCTGC3' was then measured by autoradiography.

their common plus strand 5'GGCTGC3', the mole ratio of NCS-C molecules to hexanucleotide base pairs required to provide the most efficient strand cleavage (as indicated by the cross-over point of the extension of the lines linking the cleavage at the lowest mole ratio (0.1) and the plateau of the final maximal extent of cleavage - Figure 2) remains constant. This result clearly indicates that it is not the efficiency of binding but that of the cleavage reaction that has been altered. The appropriate architecture not only of the potential intercalation site but also of the preferred trinucleotide 5'GCT3':5'AGC3' as a whole is critical to guarantee the most favorable orientation of the thiol-bound tetrahydro-s-indacene ring in the minor groove. Any positive adjustment on the trinucleotide improving the interaction of NCS-C with DNA enhances the efficiency of the radical-based reaction, without altering the nature of the interaction.

The data presented in this paper with hexanucleotides enable us to reach conclusions as to some of the critical factors that are involved in generating the optimal local geometry in DNA for NCS-C interaction. As indicated earlier (1), neighboring nucleotides of the *preferred trinucleotide*, can still be influential in determining the specificity of interaction. It has also been reported that they may be critical in determining highly sensitive interaction sites in high molecular weight DNA (26). Therefore, the nucleotide environment on each side of the target trinucleotide sequences has been kept constant throughout the series of hexanucleotides used in this study, to avoid such complications.

ACKNOWLEDGEMENTS

This work was supported by United States Public Health Service Research Grants CA 44257 and GM 12573 from the National Institutes of Health.

REFERENCES

1. Lee, S. H. and Goldberg, I. H. (1989) *Biochemistry* **28**, 1019-1026.
2. Hensens, O. D., Dewey, R. S., Liesch, J. M., Napier, M. A., Reamer, R. A., Smith, J. L., Albers-Schönberg, G., and Goldberg, I. H. (1983) *Biochem. Biophys. Res. Commun.* **113**, 538-547.
3. Edo, K., Mizugaki, M., Koide, Y., Seto, H., Furihata, K., Otake, N., and Ishida, N. (1985) *Tetrahedron Lett.* **26**, 331-334.
4. Edo, K., Akiyama, Y., Saito, K., Mizugaki, M., Koide, Y., and Ishida, N. (1986) *J. Antibiotics* **39**, 1615-1619.
5. Myers, A. G., Proteau, P. J., and Handel, T. M. (1988) *J. Am. Chem. Soc.* **110**, 7212-7214.
6. Napier, M. A. and Goldberg, I. H. (1983) *Mol. Pharmacol.* **23**, 500-510.
7. Goldberg, I. H. (1986) *Mechanisms of DNA Damage and Repair*, Simic, M. G., Grossman, L., and Upton, A. C. ed. Plenum Press 231-244.
8. Kappen, L. S. and Goldberg, I. H. (1985) *Nucleic Acids Res.* **13**, 1637-1648.
9. Hatayama, T. and Goldberg, I. H. (1980) *Biochemistry* **19**, 5890-5898.
10. Kappen, L. S., Goldberg, I. H., and Liesch, J. M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 744-748.
11. Povirk, L. F. and Goldberg, I. H. (1984) *Biochemistry* **23**, 6304-6311.
12. Povirk, L. F. and Goldberg, I. H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3182-3186.

13. Povirk, L. F. and Houlgrave, C. W. (1988) *Biochemistry* **27**, 3850-3857.
14. Kappen, L. S., Chen, C., and Goldberg, I. H. (1988) *Biochemistry* **27**, 4331-4340.
15. Povirk, L. F., Houlgrave, C. W., and Han, Y. (1988) *J. Biol. Chem.* **263**, 19263-19266.
16. Hunter, W. N., Kneale, G., Brown, T., Rabinovich, D., and Kennard, O. (1986) *J. Mol. Biol.* **190**, 605-618.
17. Sarai, A., Mazur, J., Nussinov, R., and Jernigan, R. L. (1988) *Biochemistry* **27**, 8498-8502.
18. Nilges, M., Clore, G. M., Gronenborn, A. M., Brunger, A. T., Karplus, M., and Nilsson, L. (1987) *Biochemistry* **26**, 3718-3733.
19. Nilges, M., Clore, G. M., Gronenborn, A. M., Piel, N., and McLaughlin, L. W. (1987) *Biochemistry* **26**, 3734-3744.
20. Aboul-ela, F., Varani, G., Walker, G. T., and Tinoco, I., Jr. (1988) *Nucleic Acids Res.* **16**, 3559-3572.
21. Scott, E. V., Zon, G., Marzilli, L. G., and Wilson, W. D. (1988) *Biochemistry* **27**, 7940-7951.
22. Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L., and Jones, R. A. (1986) *Biochemistry* **25**, 1027-1036.
23. Patel, D. J., Shapiro, L., Kozlowski, S. A., Gaffney, B. L., and Jones, R. A. (1986) *Biochemistry* **25**, 1036-1042.
24. Gaffney, B. L., Marky, L. A., and Jones, R. A. (1984) *Biochemistry* **23**, 5686-5691.
25. Myers, A. G. (1987) *Tetrahedron Lett.* **28**, 4493-4496.
26. Grimwade, J. E., Cullinan, E. B., and Beerman, T. A. (1988) *Biochim. Biophys. Acta* **950**, 102-112