DNA sequence recognition by under-methylated analogues of triostin A

C.M.Loretta Low, Keith R.Fox, R.K.Olsen[1] and Michael J.Waring

Department of Pharmacology, Medical School, University of Cambridge, Hills Road, Cambridge CB2 2QD, UK and [1]Department of Chemistry and Biochemistry, Utah State University, Logan, UT 84322, USA

ABSTRACT

　　Two new analogues of TANDEM (des-N-tetramethyl triostin A) have been synthesised in an effort to elucidate the molecular basis of DNA nucleotide sequence recognition in this series of compounds.　Their binding preferences have been investigated by DNAase I footprinting and differential inhibition of restriction nuclease attack.　The presence of a single N-methyl group on only one valine residue (in [N-MeVal⁴] TANDEM) abolishes the ability to recognise DNA, presumably because this antibiotic analogue has suffered an unfavourable conformational change in the depsipeptide ring.　A bis-methylated analogue, [N-MeCys³, N-MeCys⁷]TANDEM, was found to interact quite strongly with DNA and afforded binding sites, rich in AT residues, identical to those of TANDEM. Footprinting with various DNA fragments of known sequence showed that this analogue recognises sequences containing the dinucleotide TpA, although we cannot exclude the possibility that it binds to ApT as well. [N-MeCys³, N-MeCys⁷]TANDEM inhibits cutting by RsaI, a restriction enzyme that recognises GTAC but not by Sau3AI which recognises GATC.　This provides further supportive evidence that the ligand (and, by extension, TANDEM itself) prefers binding to sequences containing the dinucleotide step TpA.

INTRODUCTION

　　Quinoxaline antibiotics are potent inhibitors of nucleic acid synthesis which act by forming tight, though reversible, sequence-specific complexes with DNA [1,2].　The antibiotics are divided into two families depending upon the sulphur-containing cross-bridge which links the two rotationally equivalent halves of their octadepsipeptide ring; quinomycins have a thioacetal cross-bridge whereas triostins contain a disulphide linkage. Echinomycin (quinomycin A) and triostin A are representative members of the respective families.　They have been shown to recognise and bind to sites containing the dinucleotide sequence CpG [3-5].　By contrast, TANDEM (des-N-tetramethyl triostin A), a synthetic analogue of triostin A which lacks the methyl substituents normally present on the Ala-Cys and Cys-Val peptide bonds (arrowed in Fig 1), exhibits a pronounced preference for DNAs rich in A+T residues [6,7].　The recognition sequence for this ligand has not previously
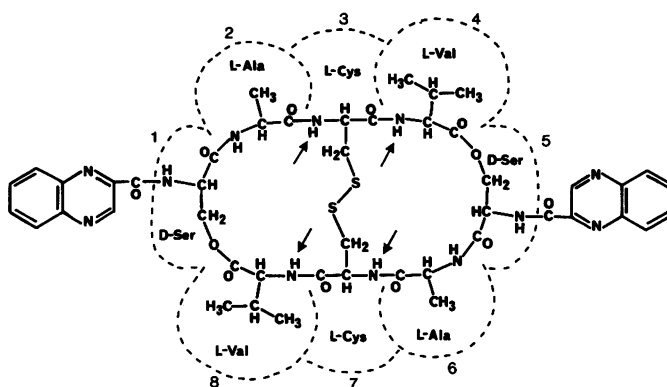
Figure 1 Structural formula of TANDEM (des-N-tetramethyl triostin A) showing the numbering scheme for the amino acid residues. Arrows indicate the NH moieties of peptide bonds which are methylated in the natural antibiotics.

been unambiguously identified, but is believed to be either TpA or ApT [4]. It is clear from these studies that the N-methyl groups play an important role in determining the AT or GC specificity of the quinoxaline antibiotics.

Crystal structures of TANDEM [8,9], triostin A [10], and of complexes between triostin A or echinomycin and a short DNA fragment [11,12] have recently been determined and have contributed a great deal to our understanding of the molecular basis of sequence recognition. The selectivity of triostin A and echinomycin for the dinucleotide step CpG appears to derive from interaction between the carbonyl groups of L-alanine residues and the 2-amino groups of guanine nucleotides [10-12]. In the crystal structure of TANDEM, which lacks the four N-methyl substituents, internal hydrogen bonds are formed between the NH groups of L-valine and the carbonyl groups of L-alanine [8,9]. These preclude the possibility of specific interactions between the alanine carbonyl groups and guanine residues, and change the structure so as to expose the amino groups of L-alanine. Viswamitra et al [8] suggested that these could now be available for interaction with the 2-keto groups of thymine in the DNA minor groove, and thus be responsible for the AT specificity.

In this paper, we examine further the role of the N-methyl groups on the depsipeptide ring by using two newly synthesised analogues of TANDEM; [N-MeVal[4]]TANDEM and [N-MeCys[3], N-MeCys[7]]TANDEM. The first analogue is modified at one of the L-valine residues which have been suggested as

important for determining the sequence-specificity of these ligands. The other analogue [N-MeCys[3], N-MeCys[7]]TANDEM permits the role of the functional groups on the amino acid residues forming the cross-bridge to be investigated. It has previously been shown that the cross-bridge itself is not instrumental in determining the sequence-specificity of these antibiotics since [Ala[3], Ala[7]]TANDEM, a compound without a cross-bridge, retains the AT specificity of TANDEM [13]. However, while the NH groups of the cysteine residues have not previously been implicated as determinants of sequence-selectivity, there is no information concerning their role, if any, in DNA binding.

To investigate the sequence binding preferences (if any) of these ligands we have used the technique of DNAase I "footprinting" with DNA fragments of known sequences. As a means to probe further the selectivity of [N-MeCys[3], N-MeCys[7]]TANDEM, we have also studied the differential inhibition of certain restriction enzymes.

MATERIALS AND METHODS
Quinoxaline depsipeptides
[N-MeVal[4]]TANDEM and [N-MeCys[3], N-MeCys[7]]TANDEM were prepared by chemical synthesis using the procedures developed previously for TANDEM and triostin A [14,15]. Details of the syntheses and characterization of the compounds, which were judged better than 99% pure, will be reported elsewhere. Stock solutions of each ligand were prepared in a methanol-buffer mixture (40/60,v/v), because of their low aqueous solubilities. The buffer used contained 10mM Tris-HCl, pH 7.5, and 10mM NaCl. The final concentration of methanol present in the digestion mixture did not exceed 20%. Controls were performed as previously described [3,4] to confirm that the presence of methanol did not significantly interfere with enzyme action.

Enzymes
Deoxyribonuclease I (DNAase I) was obtained from Sigma and prepared as a 7200 units/ml stock solution in 0.15M NaCl containing 1mM $MgCl_2$. It was stored at $-20^{\circ}C$ and diluted to working concentration immediately before use. The digestion buffer used for dilution contained 20mM NaCl, 2mM $MgCl_2$ and 2mM $MnCl_2$. All restriction enzymes were purchased from New England Biolabs.

Nucleic Acids and 3'-end-labelling with Reverse Transcriptase
The 160 base-pair duplex tyrT DNA fragment was isolated and labelled as previously described [3,16]. Incubation with reverse transcriptase, dGTP and $\alpha-[^{32}P]dCTP$ led to selective radiolabelling of the 3' end of the "Watson" strand (upper sequence in Figure 2b). Incubation with reverse transcriptase,

dTTP and α-[^{32}P] dATP was employed for selective labelling of the 3'-end of the "Crick" strand (lower sequence in Figure 2b).

The 160-mer pTyr2 DNA fragment was isolated from plasmid pMLB 1048, a gift from Dr A.A. Travers. The plasmid was first digested with BstEII and EcoRI to leave a mixture of the 160 base-pair fragment and linearized plasmid DNA [17]. From this mixture, the 160-mer pTyr2 duplex was isolated by electrophoresis on a 4% non-denaturing polyacrylamide gel followed by elution from a gel slice into 0.5M ammonium acetate containing 1mM EDTA. It was concentrated by ethanol precipitation and redissolved in 10mM Tris, pH 7.5, 0.1mM EDTA at a concentration of 2μg in 50μl. The upper sequence (Figure 4) was labelled at its 3'-end (BstEII site) with α-[^{32}P] dTTP, and the lower sequence (EcoRI site) with α-[^{32}P]dATP.

pUC13 DNA was purchased from P.L. Biochemicals Inc and was stored in buffer containing 10mM Tris, pH 7.5 and 1mM EDTA. The DNA was used as supplied without further purification. The top strand of the DNA (Figure 5b) was labelled at its 3'-end (EcoRI site) with dATP and α-[^{32}P] dTTP.

The DNA substrate used in the restriction enzyme studies was a gift from Dr H.R. Drew. It was derived from an EcoRI and Tth 111 I restriction digest of plasmid pKM Δ-98 [18,19]. The DNA was 3'-end-labelled at the Tth 111 I site with α-[^{32}P] dCTP.

DNAase I footprinting

Aliquots (3μl) of the labelled DNA (9 pmoles in base pairs) were incubated with 5μl of the ligand (10-40μM) at 37°C for 30 min, then digested with 2μl DNAase I (final concentration 0.05 units/ml). Samples (3μl) were removed from the mixture after 1, 5 and 30 minutes digestion and the reaction stopped by adding 2.5μl of 80% formamide solution containing 0.1% bromophenol blue and 10mM EDTA. These were heated at 100°C for at least 2 minutes prior to electrophoresis as described below.

Restriction enzyme digestion

Experiments were performed at 37°C in buffer containing 10mM Tris, pH 7.8, 50mM NaCl, 6mM MgCl$_2$ and 6mM mercaptoethanol. A sample of the labelled DNA was first incubated with the test ligand in the above buffer for approximately 15 minutes, before subjecting to digestion by the enzyme at an appropriate dilution. An equal volume of 15% sucrose containing 0.1% bromophenol blue and 10mM EDTA was added at the end of the incubation and the products of digestion were separated by electrophoresis on 4.5% non-denaturing polyacrylamide gels.

Gel electrophoresis

Products of digestion by DNAase I were fractionated on polyacrylamide gels (0.3mm thick) prepared in Tris-borate-EDTA buffer containing 7M urea, 12% w/v for pUC 13 DNA, and 8% w/v for tyrT and pTyr2 DNA fragments. Products of restriction enzyme cleavage were fractionated in 4.5% polyacrylamide gels containing Tris-borate-EDTA but without urea. Gels were fixed in 10% acetic acid, transferred to Whatman 3MM paper, dried under vacuum and subjected to autoradiography with an intensifying screen.

Densitometry

Autoradiographs were scanned using a Joyce-Loebl microdensitometer to produce profiles from which the relative intensity of each band was measured. The data are expressed in terms of fractional cleavage (f)= A_i/A_t as previously described [3,20], where A_i is the area under band i and A_t is the sum of the intensity under all bands in any gel lane. They are presented in the form of $\ln(f_{antibiotic}) - \ln(f_{control})$, representing the differential cleavage at each bond relative to that in the control. Positive values indicate enhancement, negative values blockage.

RESULTS

DNAase I footprinting on tyrT DNA

Patterns of DNAase I digestion for the lower (Crick) strand of this 160-base-pair DNA fragment in the absence and presence of [N-MeVal[4]]TANDEM and [N-MeCys[3], N-MeCys[7]]TANDEM are shown in Figure 2a along with those for lacTANDEM, an analogue which does not bind to DNA. The binding sites are revealed as gaps in the otherwise continuous ladder of bands, and it is immediately apparent that only [N-MeCys[3], N-MeCys[7]]TANDEM produces a digestion pattern which is different from that of the control. The other two analogues do not significantly affect the cleavage pattern, suggesting that under these conditions neither [N-MeVal[4]]TANDEM nor lacTANDEM interact with the DNA. The digestion was repeated several times with varying concentrations of each analogue. In every case the changes in the digestion pattern produced by [N-MeCys[3], N-MeCys[7]]-TANDEM were "all-or-none", with little or no protection observed below 10μM ligand, and no further protection observed beyond 15μM. No variation in the pattern of digestion was observed with [N-MeVal[4]]TANDEM at concentrations up to 40μM.

Two major sites of blockage appear in the presence of [N-MeCys[3], N-MeCys[7]]TANDEM, located around position 85 (site 1) and 110 (site 2), each about 7 base pairs in length (Figure 2a). Similar regions of protection
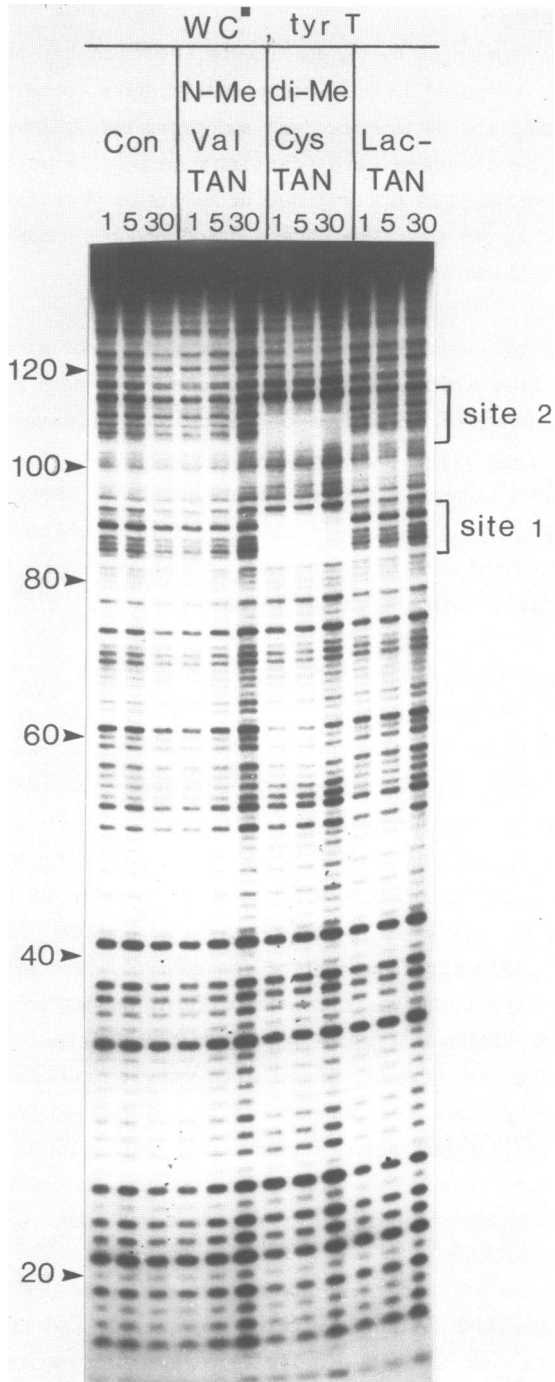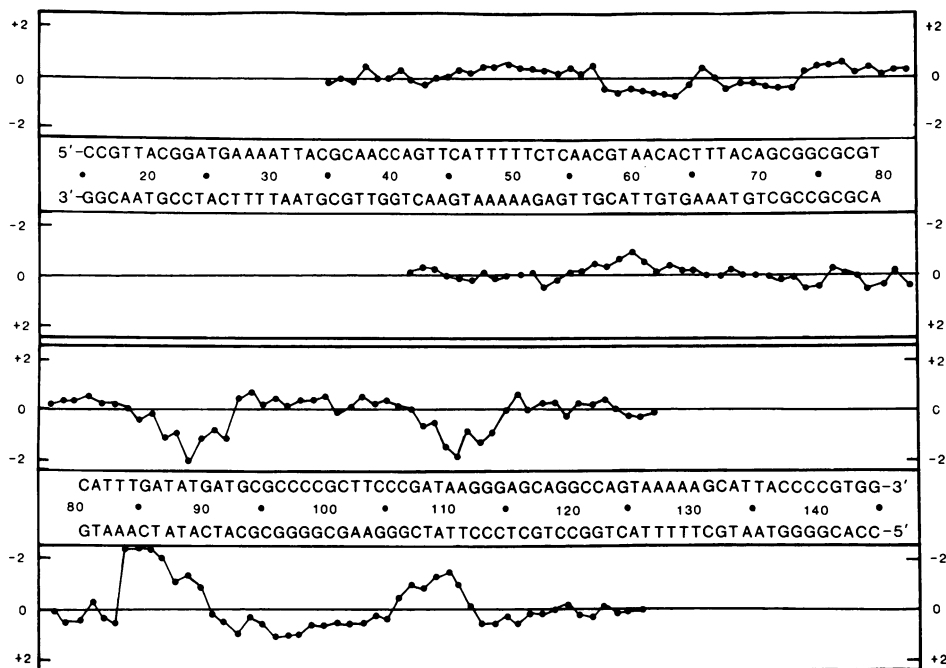
Figure 2 (a) DNAase I digestion patterns in the absence (CON) and presence of [N-MeVal⁴]TANDEM (N-MeValTAN) or [N-MeCys³, N-MeCys⁷]TANDEM (diMeCys TAN) for the tyrT DNA fragment. Symbol WC indicates that the Crick (lower) strand bears the radioactive 3'-end label. Time in minutes (1,5,30) after the addition of enzyme is shown at the top of each gel lane. The extent of digestion was limited to 20-40% of the starting material so as to minimise the incidence of multiple cuts in any one strand. Numbers on the left refer to the numbering scheme shown in Figure 2(b), while sites of protection from DNAase I digestion by [N-MeCys³, N-MeCys⁷]TANDEM are identified on the right. Also shown is the digestion pattern in the presence of a lactyl analogue of TANDEM (LacTAN), the synthesis and results for which will be presented in detail elsewhere.
(b) Differential cleavage plot for [N-MeCys³, N-MeCys⁷]TANDEM-induced differences in susceptibility to DNAase I digestion. Vertical scales are in units of $\ln f(a) - \ln f(c)$, where f(a) is the fractional cleavage at any bond in the presence of the antibiotic and f(c) is the fractional cleavage of the same bond in the control, for closely similar extents of overall digestion. Positive values indicate enhancement, negative values blockage.

were also observed for the upper (Watson) strand (gel not shown). A possible third site may be discerned around position 60 but the protection is marginal and must be considered of dubious significance. Differential cleavage plots, determined by densitometric analysis of each gel lane, are shown in Figure 2b. The diagram depicts the difference between enzymatic cleavage in the presence

and absence of the ligand, with each point plotted on a logarithmic scale so that positive values indicate enhanced cleavage by DNAase I, and negative values protection from enzymatic cleavage.

The gel and the differential cleavage plot together reveal that both major regions of protection are centred around AT clusters: ATAT at site 1 and ATAA at site 2. At each site the block is staggered across the two strands by about three bonds towards the 3'end, as previously observed with DNAase I footprinting [3,4,20], presumably because DNAase I cuts bonds which lie in close proximity across the minor groove. The stagger is particularly marked at site 1 between positions 84 and 90, possibly because the blockage at this region appears somewhat stronger than at site 2 (positions 107 to 113). These regions protected from cleavage by [N-MeCys[3], N-MeCys[7]]TANDEM are identical to those previously observed with TANDEM using the same DNA fragment [4].However, the blockage seen with this analogue is much more pronounced. Since the new compound is structurally so similar to TANDEM, it is reasonable to suppose that it binds to double-helical DNA by the same mechanism, with two base pairs sandwiched between the quinoxaline chromophores in the bis-intercalated complex. The sandwiched base sequence cannot be unambiguously determined from the present data alone and could be either ApT or TpA, since both binding sites contain the sequence ATA. Moreover, the true binding preference must involve more than mere specificity for AT residues since not all ApT or TpA steps are susceptible. The ApT steps at positions 14, 25, 31, 47, 82, 92 and 134 are unaffected by the presence of the ligand, as are the TpA steps at positions 12, 20, 33, 68 and 136. There is only a marginal effect around the TpA step at position 60. It is noteworthy that those dinucleotide steps that are insensitive to [N-MeCys[3], N-MeCys[7]]TANDEM are often preceded or followed by runs of A or T which, it has been suggested, can adopt a peculiar helical conformation [16,21,22]. For example, the ApT at position 31 is preceded by the sequence GAAA while that at position 47 is followed by TTTCT: neither site is protected at all from enzymatic cleavage. It is possible then that local variations in DNA structure may play an important role in determining which TpA or ApT steps can or cannot bind the ligand.

We note in passing that one of the sites most favoured for binding of [N-MeCys[3], N-MeCys[7]]TANDEM, and of TANDEM itself, is the Pribnow box of the tyrT promoter sequence TATGATG at position 87 – 94 (the start site is numbered 100). This region becomes moderately sensitive to SI nuclease attack under superhelical stress [23] suggesting that TANDEM, and ligands related to it, would serve as ideal probes for examining the biological function of promoters

– a notion which lay behind the original selection of the tyrT fragment as a substrate for footprinting analysis [3].

Subsequent experiments were designed to identify the precise recognition sequence of [N–MeCys[3], N–MeCys[7]]TANDEM; in particular to determine whether the preferred binding sites contain the sequence ApT or TpA. The results should shed light on the recognition sequence of TANDEM itself since both ligands appear to share the same sequence-specificity. The first two experiments involved the use of different DNA fragments for further footprinting studies, while in the third set of experiments the inhibition of restriction enzyme digestion by [N–MeCys[3], N–MeCys[7]]TANDEM was examined.

DNAase I footprinting with pTyr2 DNA

Patterns of cleavage inhibition produced by [N–MeCys[3], N–MeCys[7]]TANDEM on the lower strand of pTyr2 DNA are shown in Figure 3a. DNAase I cuts this piece of DNA more evenly than tyrT DNA (cf Figure 2a). There are three major regions in this substrate which are protected by the ligand, situated around positions 35, 85 and 134 on both strands. Another possible binding site may be discerned around position 20 on the lower strand (Figure 3a). A differential cleavage plot derived from densitometric tracing of this and other gels is shown in Figure 3b for the 100 base pairs that are sufficiently well-resolved for quantitative determination.

The three binding sites revealed by the differential cleavage plot contain both ApT and TpA steps; most clearly the binding sites at position 35 (ATAT) and position 85 (TATA). It is worth noting that the isolated ApT or TpA steps which are not protected by the ligand are often associated with runs of A or T; for example, the TpA steps at positions 73, 94 and 117 are preceded by runs of A at around positions 65, 90 and 112 respectively. The only isolated TpA step to which the ligand appears to bind is located around position 21 and is situated within a region of mixed DNA sequence. By way of contrast, echinomycin, included in the experiment for purposes of comparison, produces clear blockages in pTyr2 DNA which are centred around the dinucleotide step CpG, in accordance with previous observations [3,5].

Once again it is not possible to identify conclusively the sequence recognised by [N–MeCys[3], N–MeCys[7]]TANDEM, because the protected regions on pTyr2 DNA contain consecutive AT and TA pairs such as ATAT and TATA. A third DNA fragment containing only two TpA or ApT sequences (one of each) was therefore sought, hopefully to provide a definitive conclusion.

pUC 13 DNA

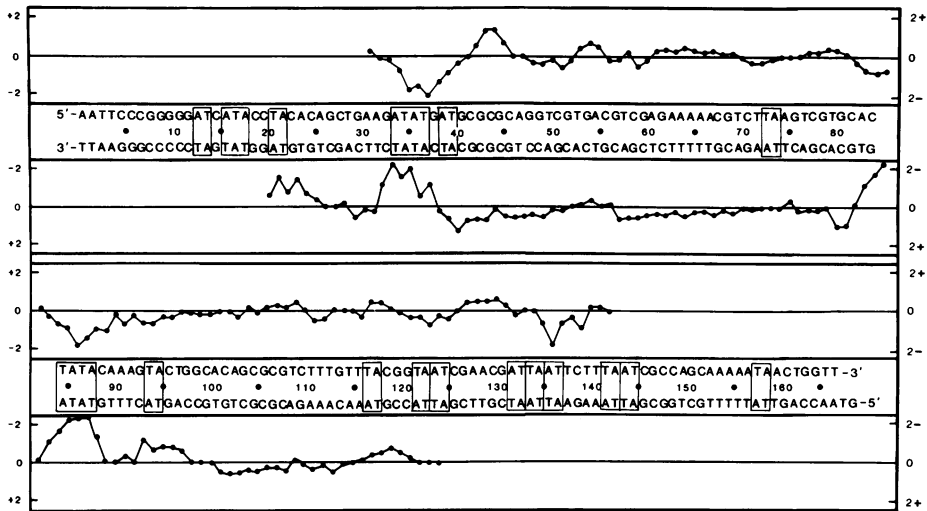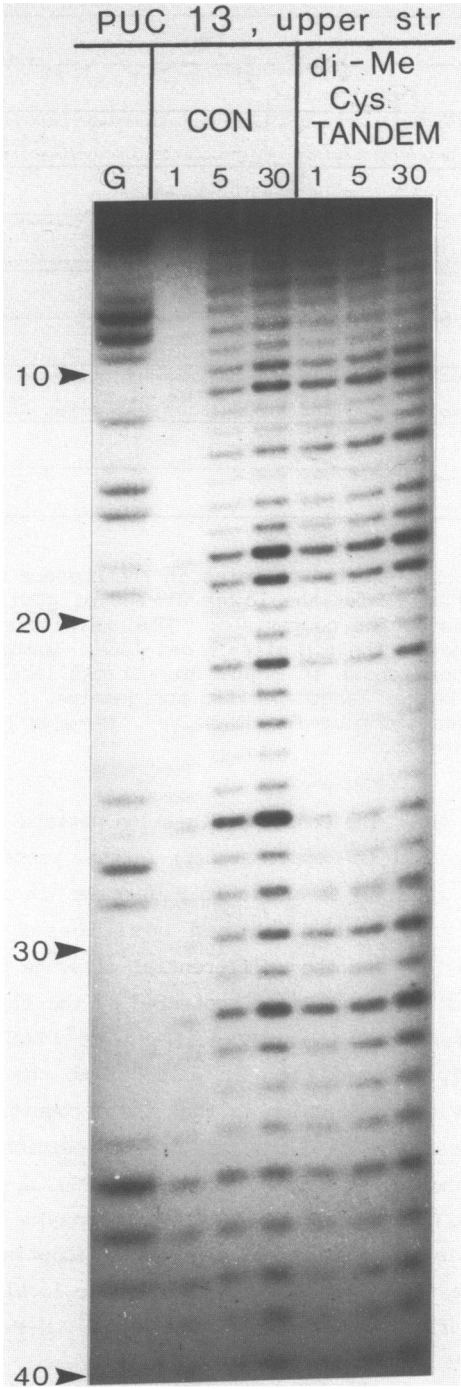In this DNA isolated TpA and ApT steps occur at positions 24 and 30
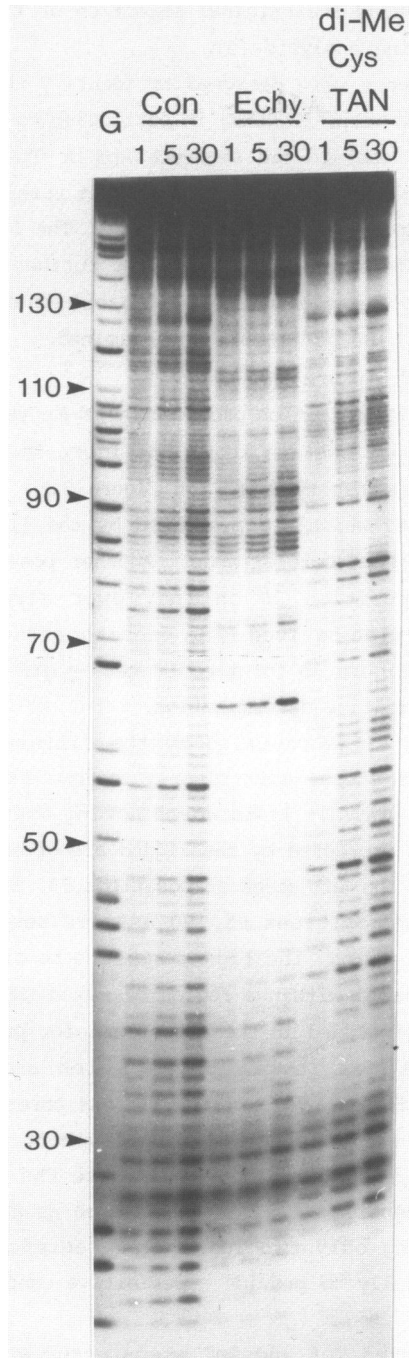
Figure 3 (a) DNAase I digestion patterns in the absence (CON) and presence of [N-MeCys3, N-MeCys7]TANDEM for the lower strand of pTyr2 DNA. Echinomycin (Echy) was also included for comparison. Time in minutes after the addition of the enzyme is shown at the top of each gel lane. Numbers on the left refer to the numbering scheme shown in Figure 3b. Tracks labelled 'G' are dimethyl sulphate-piperidine marker lanes specific for guanine.
(b) Differential cleavage plot for [N-MeCys3, N-MeCys7]TANDEM on pTyr2 DNA. Lettering as in Figure 2b.

respectively (Figure 4b). The DNAase I digestion pattern of this DNA fragment (labelled at the 3'-end of the upper strand) in the presence and absence of [N-MeCys3, N-MeCys7]TANDEM is shown in Figure 4a. It can be seen that the relative intensity of bands between positions 23 and 28 is reduced corresponding to the "dip" in the differential cleavage plot shown in Figure 4b. No other regions in this DNA are protected by the ligand from cleavage.

It appears then that [N-MeCys3, N-MeCys7]TANDEM binds to the dinucleotide step TpA at position 24, but not to the ApT at position 30 (Figure 4b). The obvious conclusion is that the recognition sequence must be TpA. However, it should be noted that the ApT at position 30 is preceded by a GG sequence and followed by a long GC-rich run CCCCGGGCG. The central part of this long run is not cleaved efficiently by the enzyme in the control, and sequences such as this are known to be capable of adopting unusual structures [24]. It is possible therefore that an unfavourable local DNA conformation is responsible for hindering the binding of [N-MeCys3, N-MeCys7]TANDEM to the adjacent dinucleotide step ApT.
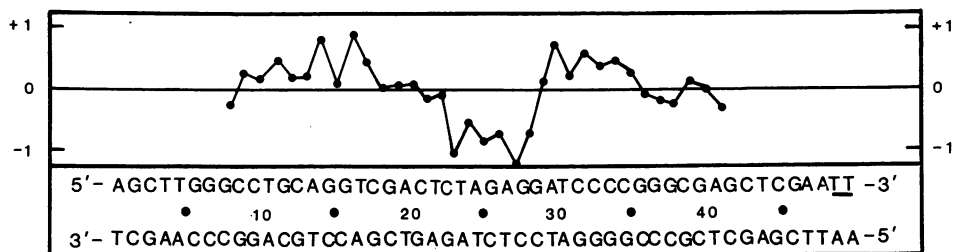
Differential cleavage


```
+1                                                                    +1
 0                                                                     0
-1                                                                    -1

5'- AGCTTGGGCCTGCAGGTCGACTCTAGAGGATCCCCGGGCGAGCTCGAATT -3'
        •   10    •    20    •    30    •    40    •
3'- TCGAACCCGGACGTCCAGCTGAGATCTCCTAGGGGCCCGCTCGAGCTTAA -5'
```

Figure 4 (a) DNAase I footprinting of [N-MeCys^3, N-MeCys^7]TANDEM on the upper strand of pUC 13 DNA. Lettering as in Figure 3a. The numbering scheme refers to the sequence shown in Figure 4b.
(b) Differential cleavage plot for [N-MeCys^3, N-MeCys^7]TANDEM on pUC 13 DNA. Symbols are as described in the legend to Figure 2b.

The possibility that local structural factors materially affect the interaction of [N-MeCys^3, N-MeCys^7]TANDEM with its preferred binding sequence cannot be ignored. Consequently the interpretation of DNAase I footprinting patterns observed with various DNAs may not be straightforward. To resolve this dilemma, we decided to adopt a different approach and have probed the recognition sequence of the ligand by studying its effects on the activity of several restriction enzymes.

Inhibition of restriction nuclease digestion

The two restriction enzymes studied were RsaI (GTAC) and Sau3AI (GATC), for which the central dinucleotides in the recognition sequences are TpA and ApT respectively. The effect of echinomycin on these enzymes was included in the study to provide a point of comparison; since their recognition sites do not contain the dinucleotide step CpG to which echinomycin binds we anticipated that this antibiotic should not interfere with the activity of either enzyme.

The effects of [MeCys^3, MeCys^7]TANDEM and echinomycin on digestion by RsaI and Sau3AI are illustrated in Figure 5. At high concentrations echinomycin blocks the activities of both enzymes, while [MeCys^3, MeCys^7]-TANDEM only inhibits RsaI (GTAC) and has no effect at all on Sau3AI (GATC). At first sight, the results seem to confirm that the recognition sequence of the TANDEM analogue is indeed TpA. However, the interpretation is complicated by the unexpected finding that echinomycin, which recognises CpG, also inhibits both enzymes. A possible explanation for this observation is that
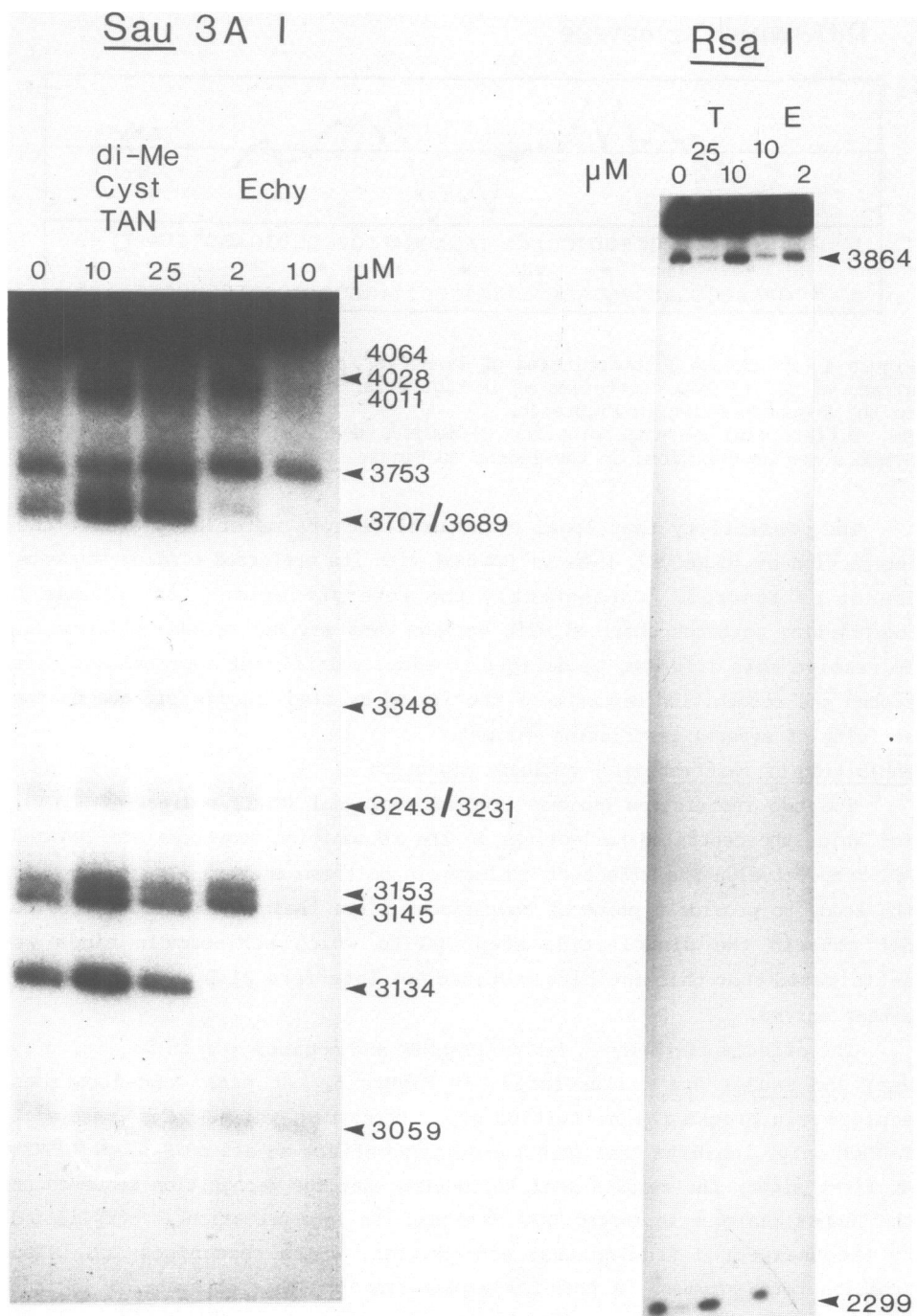
the enzyme activity can be affected by local structural distortions of DNA induced by antibiotic binding to regions adjacent to the enzyme cutting site, as well as direct steric blockage by the ligand itself. One example of the former effect is evident at the Sau3AI restriction site at position 3689, which is effectively blocked by echinomycin. This probably results from the ligand binding to the adjacent CpG step at position 3688 (Table 1). However, some of the blockages produced by echinomycin are not so easily explained, such as the inhibition of cutting at RsaI sites at positions 2299 and 3864. With regard to the former, the nearest CpG is found at position 2282. As regards the latter, there are CpG steps at 3837 and 3888. Each of these CpG steps is more than one helical turn displaced from the cutting site, which seems too far away to exert either local structural or steric effects.

Turning again to [N-MeCys^3, N-MeCys^7]TANDEM, the RsaI site (GTAC) at position 2299 is closely preceded by an ApT dinucleotide at position 2296 (Table 1), whereas the other RsaI site at position 3864 is 15 bases away from the nearest ApT step at position 3879. Therefore inhibition of cleavage at 2299 might be explained by [N-MeCys^3, N-MeCys^7]-TANDEM binding to either ApT or TpA. However, the inhibition of the restriction enzyme cutting at position 3864 is much more likely to result from direct blockage by the ligand binding to the step TpA.

In a further series of experiments we investigated possible inhibitory effects of [N-MeCys^3, N-MeCys^7]-TANDEM on the restriction enzyme HaeIII which cuts the sequence GGCC. The only restriction site on the fragment blocked by the TANDEM analogue was located at position 2969 within the sequence GAAGTGGTGGCCTAACTACG. It seems reasonable to suppose that the observed inhibition reflects binding of the ligand to the adjacent TpA step at 2972. All the other HaeIII sites on this fragment were not affected by the analogue and upon examination were found to be remote from any TpA (or ApT) steps.

DISCUSSION

[N-MeVal^4]TANDEM

The apparent failure of [N-MeVal^4]TANDEM to bind to the tyrT DNA is rather surprising since it was anticipated that the loss of one of the

Figure 5 Differential effects on restriction enzyme digestion of echinomycin (abbreviated as Echy or E), and [N-MeCys^3, N-MeCys^7]TANDEM (abbreviated as di-Me Cyst TAN or T). RsaI recognises the sequence GTAC and Sau3AI GATC. The numbering scheme for nucleotides in the DNA substrate is as defined by Drew and Travers [18]. Table 1 lists the cutting sites together with their flanking sequences extending for one turn of the ten-fold helix in either direction.

TABLE 1

Enzyme	Cutting Site	5'-flanking		3'-flanking
RsaI	3864	ACTGGTGA	[GTAC]	TCAACCAA
	2299	AGCAGATT		TGAGAGTG
Sau3AI	4064	ACCCAACT	[GATC]	TTCAGCAT
	4028	GCTGTTGA		CAGTTCGA
	4011	CTCTCAAG		TTACCGCT
	3753	GGTCCTCC		GTTGTCAG
	3707	AGTTACAT		CCCCATGT
	3689	TTCCCAAC		AAGGCGAG
	3348	ATCTCAGC		TGTCTATT
	3243	TTCACCTA		CTTTTAAA
	3231	TCAAAAAG		TTCACCTA
	3153	GATCCTTT		TTTTCTAC
	3145	CTCAAGAA		CTTTGATC
	3134	AAAAAAAG		TCAAGAAG
	3059	TAGCTCTT		CGGCAAAC

internal hydrogen bonds might cause the compound to revert to a conformation akin to that of triostin A. In this analogue the carbonyl group of one of the alanines (Ala[7]) should be free to interact with the amino group of a guanine nucleotide, while the other alanine (Ala[3]) could still be involved in forming an internal hydrogen bond with the NH of Val[8]. It can even be imagined that one half of the molecule might recognise guanine residues while the other would recognise an AT-base-pair providing a compound with unusual intermediate sequence-selectivity. In the event neither of these notions proved to be correct. Most probably the conformation of the depsipeptide in solution is different from that of either triostin A or TANDEM; the asymmetric peptide ring may be twisted in such a way that the quinoxaline chromophores are no longer positioned roughly parallel to allow for bifunctional

intercalation into the DNA helix - like the situation previously postulated for L-serine-containing analogues of TANDEM [6,13].

[N-MeCys^3, N-MeCys^7]TANDEM

All the experimental results consistently point to the same sequence-selectivity for [N-MeCys^3, N-MeCys^7]TANDEM as for TANDEM itself. The cysteine -NH groups do not therefore appear to play an important role in either holding the peptide portion of the antibiotic in any particular conformation, or in determining the sequence-selectivity of the ligand. In fact [Ala^3, Ala^7]TANDEM, a compound lacking a cross-bridge altogether, has been shown to interact with DNA (albeit weakly) and again displays the same AT specificity as does TANDEM [13]. It seems likely that the function of the cysteine residues in providing the cross-bridge is to restrict the conformational flexibility of the peptide ring so as to reduce unfavourable entropy terms in the binding reaction, rather than to participate in direct interactions with functional groups on the DNA. Our observation that methylation of the cysteine residues produces more pronounced "footprints" probably means that the analogue binds more tightly to DNA than does TANDEM, presumably as a consequence of the increased hydrophobic character of the interaction.

Viswamitra _et al_ [8] proposed that the sandwiched base pair in the intercalated TANDEM complex should be ApT (rather than TpA) since the vector joining the alanine NH protons in TANDEM is parallel to the vector joining the 2-keto groups of thymines for the sequence ApT, whereas it is perpendicular for the sequence TpA. This should favour interaction with ApT rather than TpA. By contrast, however, it has been reported from thermal denaturation experiments that TANDEM binds to poly d(TAC). poly d(GTA) but not to poly d(ATC). polyd(GAT) [25], suggesting that the ligand binds to the sequence TpA rather than ApT. The DNAase I footprinting results presented here and in ref. [4] confirm that both TANDEM and [N-MeCys^3, N-MeCys^7]TANDEM indeed bind to sequences containing AT residues, but we cannot unambiguously identify the sandwiched base sequence because the binding site often comprises consecutive AT and TA base pairs such as ATAT, TATA or TAAT. Moreover, local deviations from classical B-DNA caused by the nature of surrounding sequences appear to affect profoundly the binding of these ligands. In the last analysis we can confidently state that [N-MeCys^3, N-MeCys^7]TANDEM does bind to the dinucleotide step TpA, though we cannot exclude the possibility that it interacts with the sequence ApT as well.

It is important to note that the ability to recognise a particular

dinucleotide step as the preferred site for binding does not necessarily identify that step as the sequence sandwiched between the quinoxaline rings in a bis-intercalative complex. The TpA (pyrimidine-purine) sequence is known to unstack more easily than other sequences [26] which ought to render it more favourable as a site for penetration (intercalation) of one of the chromophores. But if so we would expect to see some definite preference for the sequence TpApTpA to allow for the bifunctional reaction: such a sequence is present at positions 84 - 88 in the pTyr2 fragment but is not noticeably more susceptible than other sites to protection from nuclease attack (Fig 3).

Although we have established that the NH groups of the cysteine residues play no significant role in determining the AT sequence-selectivity of TANDEM, it remains to be seen whether the same is true for the GC selectivity of triostin A. The present model for the interaction of triostin A with the dinucleotide step CpG [10,11] suggests that these entities are indeed unimportant. Perhaps the methyl groups are present in the naturally-occuring antibiotics to strengthen the interaction with DNA by increasing its hydrophobic character, and/or to protect the compounds from protease digestion. It will be of considerable interest to examine the sequence-selectivity of [N-MeVal^4, N-MeVal^8]TANDEM if the compound becomes available. We predict that it will display the same CpG selectivity as does the natural antibiotic triostin A.

REFERENCES
1   Gale, E.F.,Cundliffe,E.,Reynolds, P.E., Richmond, M.H. and Waring, M.J. (1981). In The Molecular Basis of Antibiotic Action, pp 173-277. John Wiley & Sons Ltd, London.
2   Waring,M.J. and Fox, K.R (1983) in Molecular Aspects of Anti-cancer Drug Action, eds Waring, M.J. and Neidle, S., pp 127-156. Macmillan, London.
3   Low, C.M.L., Drew, H.R. and Waring, M.J. (1984) Nucleic Acids Res., 12, 4865-1879.
4   Low, C.M.L., Olsen, R.K. and Waring, M.J. (1984) FEBS Letters 176, 414-419.
5   Van Dyke, M.W. and Dervan, P.B. (1984) Science, 225, 1112-1127.
6   Lee, J.S. and Waring, M.J. (1978) Biochem. J., 173, 129-144.
7   Fox, K.R., Olsen, R.K. and Waring, M.J. (1982) Biochim. Biophys. Acta., 696, 315-322.
8   Viswamitra,M.A.,Kennard,O.,Cruse,W.B.T., Egert, E., Sheldrick, G.M.,

Jones, P.G., Waring, M.J., Wakelin, L.P.G. and Olsen, R.K. (1981) Nature, 289, 817-819.

9   Hossain, M.B., van der Helm, D., Olsen, R.K., Jones, P.G., Sheldrick, G.M., Egert, E., Kennard, O., Waring, M.J. and Viswamitra, M.A. (1982) J. Amer. Chem. Soc. 104, 3401-3408.

10 Sheldrick, G.M., Guy, J.J., Kennard, O., Rivera, V. and Waring, M.J. (1984) J. Chem. Soc. Perkin Trans II, 1601-1605.

11 Wang, A.H-J., Ughetto, G., Quigley, G.J., Hakoshima, T., van der Marel, G.A, van Boom, J.H. and Rich, A. (1984) Science, 225, 1115-1121.

12 Ughetto, G., Wang, H-J., Quigley, G.J., van der Marel, G.A., van Boom, J.H., and Rich, A. (1985) Nucleic Acids Res. 13, 2305-2323.

13 Fox, K.R., Olsen, R.K. and Waring, M.J. (1980) Br. J. Pharmac. 70, 25-40.

14 Ciardelli, T.L. and Olsen, R.K. (1977) J. Amer. Chem. Soc. 99, 2806-2807.

15 Chakravarty, P.K. and Olsen, R.K. (1978) Tetrahedron Lett. 1613-1616.

16 Drew, H.R. and Travers, A.A. (1984) Cell 37, 491-502.

17 Lamond, A.I. and Travers, A.A. (1983) Nature, 305, 248-250.

18 Drew, H.R. and Travers, A.A. (1985) Nucleic Acids Res., 13, 4445-4467.

19 Travers,A.A., Lamond, A.I., Mace, H.A.F. and Berman, M.L. (1983) Cell 35, 265-273.

20 Fox, K.R. and Waring, M.J. (1984) Nucleic Acids Res., 12, 9271-9285.

21 Fratini, A.V., Kopka, M.L., Drew, H.R. and Dickerson, R.E. (1982) J. Biol. Chem., 257, 14686-14707.

22 Arnott, S., Chandrasekaran, R., Hall, I.H. and Puigjaner, L.C. (1983) Nucleic Acids Res., 11, 4141-4155

23 Drew, H.R., Weeks, J.R. and Travers, A.A. (1985) EMBO J., 4, 1025-1032.

24 McCall, M., Brown, T. and Kennard, O. (1985) J. Mol. Biol. 183, 385-396.

25 Evans, D.H., Lee, J.S., Morgan, A.R. and Olsen, R.K. (1982) Can. J. Biochem. 60, 131-136.

26 Ornstein, R.L., Rein, R., Breen, D.L. and MacElroy, R.D. (1978) Biopolymers 17, 2341-2360.