Sequence preferences in the binding to DNA of triostin A and TANDEM as reported by DNase I footprinting

C.M. Loretta Low, Richard K. Olsen† and Michael J. Waring*

University of Cambridge Department of Pharmacology, Medical School, Hills Road, Cambridge CB2 2QD, England, and †*Utah State University, Department of Chemistry and Biochemistry, Logan, UT 84322, USA*

Received 17 August 1984

Six or seven triostin-binding sites have been identified in a 160-base-pair DNA restriction fragment containing the *tyr* T promoter sequence. Each is centred round a CpG step, and the minimum binding site-size appears to be six base pairs. The sites are practically the same as those reported for echinomycin by DNase I digestion. Only two sites are protected by binding of TANDEM, the des-*N*-tetramethyl analogue of triostin A; they are centred around the sequences ATA or TAT.

Sequence specificity *Triostin A* *TANDEM* *Quinoxaline antibiotic* Bis-*intercalation* *Footprinting*

1. INTRODUCTION

Quinoxaline antitumour antibiotics are produced by a variety of streptomycetes. They are characterised by a cross-bridged octadepsipeptide ring to which are attached two moieties of quinoxaline-2-carboxylic acid [1,2]. Depending upon the exact constitution of that cross-bridge, the antibiotics fall into two families called quinomycins and triostins, the former having a thioacetal cross-bridge and the latter a disulphide. The representative member of the quinomycins is echinomycin and that of the triostins is triostin A (fig.1a). All these antibiotics have been shown to bind strongly to double-stranded DNA by a mechanism of bifunctional intercalation [3,4].

Equilibrium binding studies have established that echinomycin and triostin A bind better to natural DNAs which are rich in G+C residues, although in general triostin A appears to display somewhat less pronounced sequence selectivity [5–7]. Moreover, triostin A binds more tightly to poly(dA-dT) than to poly(dG-dC) whereas the reverse is true for echinomycin. These results suggest that there are subtle differences in the precise base sequence preferences of the two antibiotics. We have previously reported sequence selectivity of echinomycin as revealed by DNase I footprinting of a 160-base-pair DNA fragment from *Escherichia coli* containing the *tyr T* promoter sequence [8]. A similar study using MPE·Fe (II) footprinting has also been published [9]. The binding sites were all found to contain the dinucleotide sequence CpG. Here, we describe the results of similar experiments to investigate the preferred binding sequence(s) of triostin A.

Unlike the naturally occurring quinoxaline antibiotics, the synthetic des-*N*-tetramethyl analogue of triostin A nicknamed TANDEM (fig.1b) displays a strong preference for binding to AT-rich DNAs, the extreme manifestation of which is seen with poly(dA-dT) to which it binds at least 5000 times more tightly than to poly(dG-dC) [3,10]. We demonstrate here that the preferred binding sites for TANDEM in our DNA fragment, revealed via blockages of DNase I digestion, contain alternating A and T residues. Thus in this case the sequence selectivity deduced from observations with synthetic polynucleotides applies to natural DNA species as well.

* To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

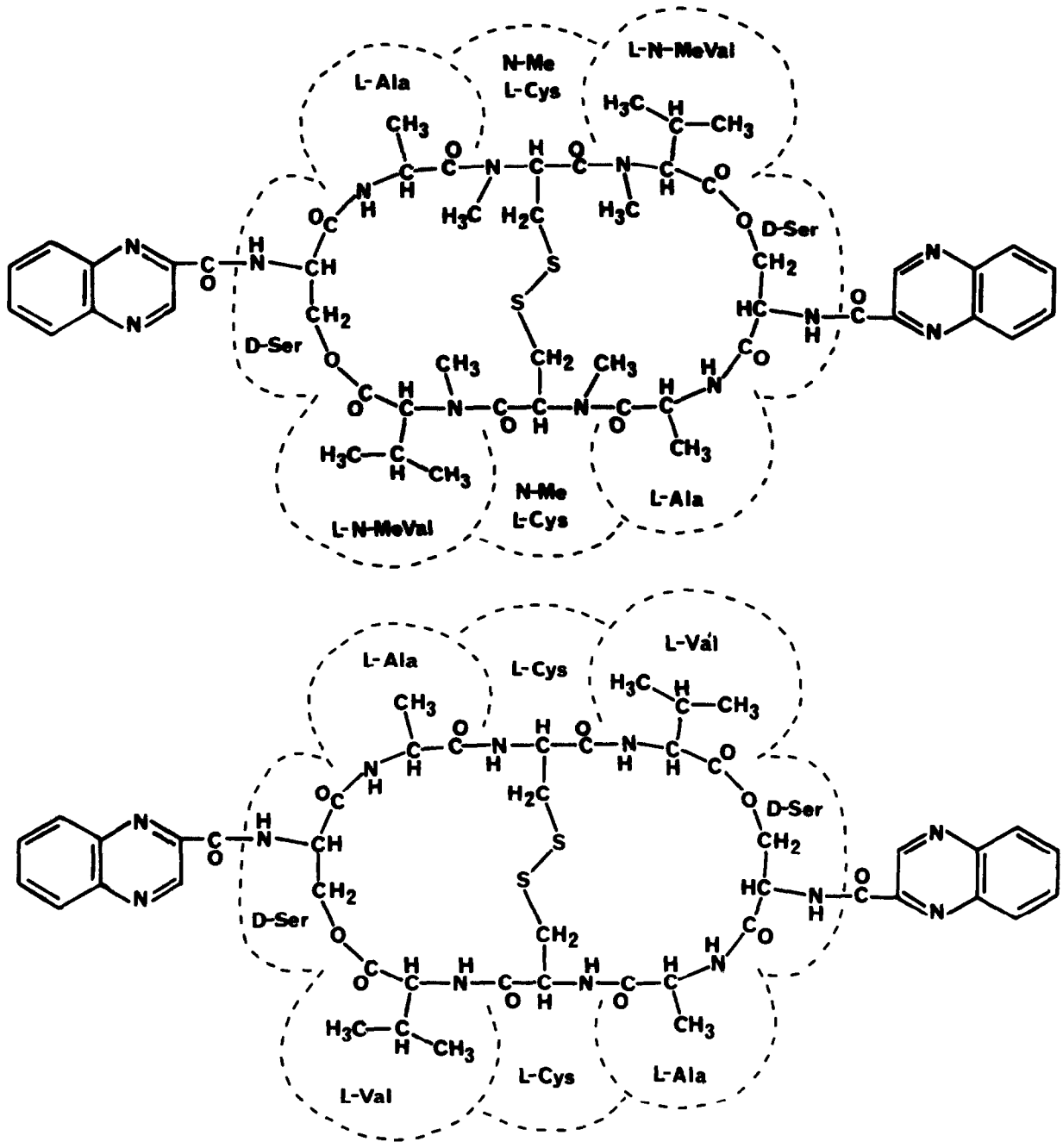

Fig.1. (a) Structure of triostin A. (b) Structure of TANDEM.

2. MATERIALS AND METHODS

Triostin A was a gift from Drs H. Otsuka and T. Yoshida of Shionogi and Co., Osaka, Japan. TANDEM was synthesised, characterised and purified as in [11]. Concentrations were determined spectrophotometrically from the absorbance at 325 nm using extinction coefficients of

10 900 $M^{-1} \cdot cm^{-1}$ and 12130 $M^{-1} \cdot cm^{-1}$ for triostin A and TANDEM, respectively [3,5]. Stock solutions of each ligand were prepared in a methanol-buffer mixture (40:60, v/v) because of their low aqueous solubility. The buffer used was 10 mM Tris–HCl, pH 7.5, containing 10 mM NaCl. The final concentration of methanol present in the digestion mixture did not exceed 20%. Controls were performed as before [8] to verify that the presence of methanol did not interfere with the enzyme action.

Deoxyribonuclease I (DNase I) was obtained from Sigma and prepared as a 7200 units/ml stock solution in 0.15 M NaCl containing 1 mM $MgCl_2$. It was stored at −20°C and diluted to working concentration immediately before use. The digestion buffer used for dilution contained 20 mM NaCl, 2 mM $MgCl_2$ and 2 mM $MnCl_2$.

A 160-base-pair duplex DNA fragment from *E. coli*, containing the tyrosine tRNA promoter together with its adjacent sequences, was a gift from Drs A.A. Travers and H.R. Drew. Its nucleotide sequence is represented in fig.3. The upper sequence ('Watson strand', reading 5′–3′ from left to right) can be labelled at the *AvaI* site at its 3′-end with [α-^{32}P]dCTP and the lower sequence ('Crick strand', reading 5′–3′ from right to left) can be labelled at the *EcoRI* site at its 3′-end with [α-^{32}P]dATP. 3′-End labelling was performed as in [8,12]. An aliquot (3 μl) of the labelled DNA (9 pmol in base pairs) was incubated with 5 μl triostin A (5–20 μM) or TANDEM (15–40 μM) at 37°C for 30 min, then digested with 2 μl DNase I (final concentration 0.05 units/ml). Samples (3 μl) were removed from the mixture after 1, 5 and 30 min digestion and the reaction stopped by adding 2.5 μl of 80% formamide solution containing 0.1% bromophenol blue and 10 mM EDTA. They were heated at 100°C for at least 2 min prior to electrophoresis [8].

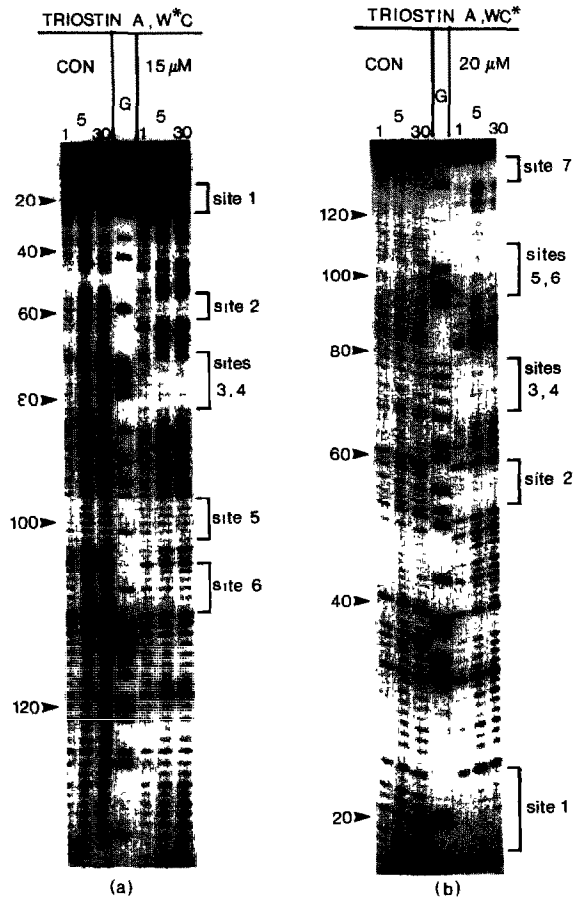

Fig.2. DNase I footprinting of triostin A on a 160-base-pair DNA fragment from *E. coli*. Symbols W*C and WC* indicate which of the two strands (Watson or Crick) bears a radioactive 3′-end label. Time in minutes (1, 5, 30) after the addition of enzyme is shown at the top of each gel lane. The extent of digestion was limited to 20–40% of the starting material so as to minimise the incidence of multiple cuts in any strand. Tracks labelled 'G' are dimethyl sulphate-piperidine marker lanes specific for guanine. Numbers on the left refer to the numbering scheme shown in fig.3, while sites of protection from DNase I digestion are identified on the right.

3. RESULTS

3.1. *Triostin A*

Patterns of DNase I digestion for the Watson and Crick strands of the 160-base-pair DNA fragment in the presence of triostin A are shown in fig.2a,b, respectively. Six major binding sites are apparent, located around positions 20, 58, 73, 76, 100 and 107 on both strands. A seventh site may be discerned from the gap in the Crick strand near position 140 at the top of fig.2b. The minimum binding site-size appears to be 6 base pairs, as can be seen from the blockage at positions 54–60 in fig.2b. The broad protected zone around position 105 seems in fact to contain two closely spaced sites of blockage, as is evident from fig.2a. No in-

termediate states of antibiotic-induced protection from cleavage were apparent at varying triostin concentrations. Little or no protection was observed at antibiotic concentrations below 5 μM while no further protection was afforded above 10 μM.

All 7 binding sites are associated with one or more steps of the dinucleotide sequence CpG (fig.3). The only such step that is not strongly protected is the one at position 35. In general, these footprints produced by triostin A are broadly similar to those observed with echinomycin [8], although the blockages are less pronounced.

Enhanced cleavage relative to that in the control occurs at flanking sequences adjacent to certain binding sites, notably at positions 30, 50, 65 and 130. These areas contain long runs of A and T, for example, ATTTTTCT around position 50 and AAAAAG at position 130. By contrast, there is no effect on the rate of cleavage at the GC-rich regions adjacent to binding sites 5 and 6.

The protected zones measured on the two strands appear to be staggered two bonds towards the 3'-end, as noted in previous studies with other ligands [8,12–14]. At 'site 2', for example, protection extends from positions 56 to 62 on the Watson strand, and from positions 54 to 60 on the Crick strand. The cross-strand stagger in the regions subject to enhanced cleavage is less obvious.

3.2. *TANDEM*

The DNase I digestion pattern for bonds 75–120 of the Watson strand in the presence of this ligand is shown in fig.4. Two binding sites can be discerned from the footprints, located near positions 88 and 110. They are each about 8 base pairs in length. No other protected regions could be seen. Parallel experiments on the Crick strand (gel not shown) revealed that the binding sites were again staggered by two base pairs towards the 3'-end (fig.3). No protection was observed below 15 μM TANDEM and no further protection beyond 30 μM.

The DNase I digestion patterns observed in the presence of TANDEM are clearly very different from those seen in the presence of triostin A and

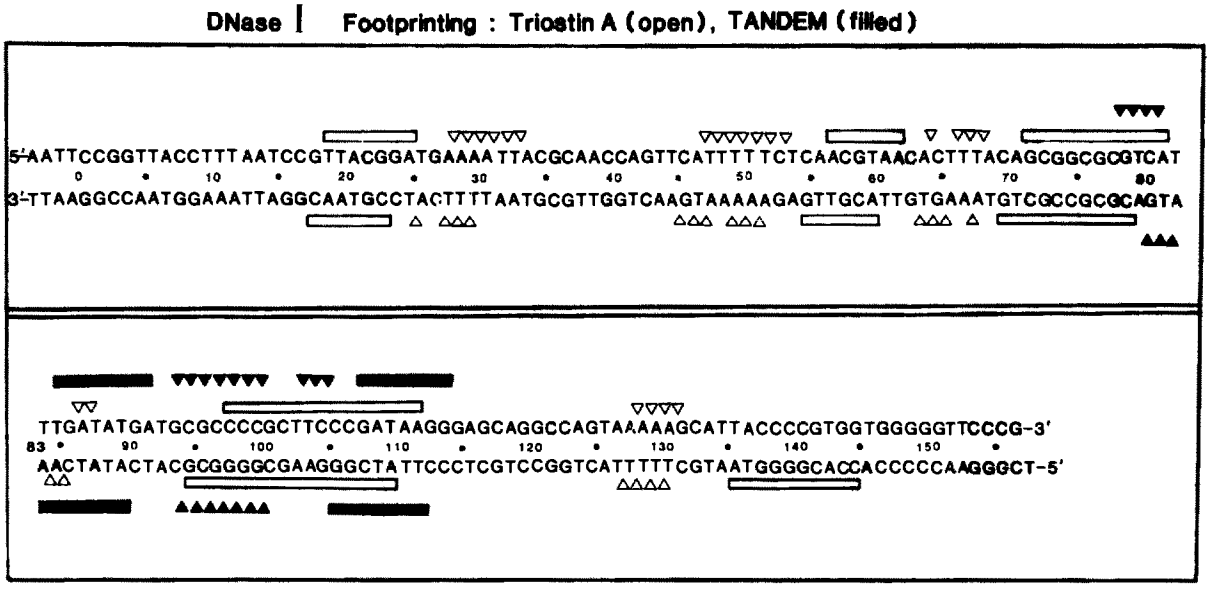

Fig.3. Summary of footprinting results for triostin A (unfilled bars, open triangles) and TANDEM (filled bars, filled triangles). Bars indicate protected regions while triangles show positions of enhanced cutting. The upper sequence represents the 'Watson' (antisense) strand and the lower sequence the 'Crick' (sense, coding) strand. These maps were compiled from visual inspection of numerous gels as well as from densitometric tracings, and may be considered as a set of averaged values. Note that the increased sensitivity to DNase I in the presence of TANDEM is much less pronounced than that seen with triostin A.

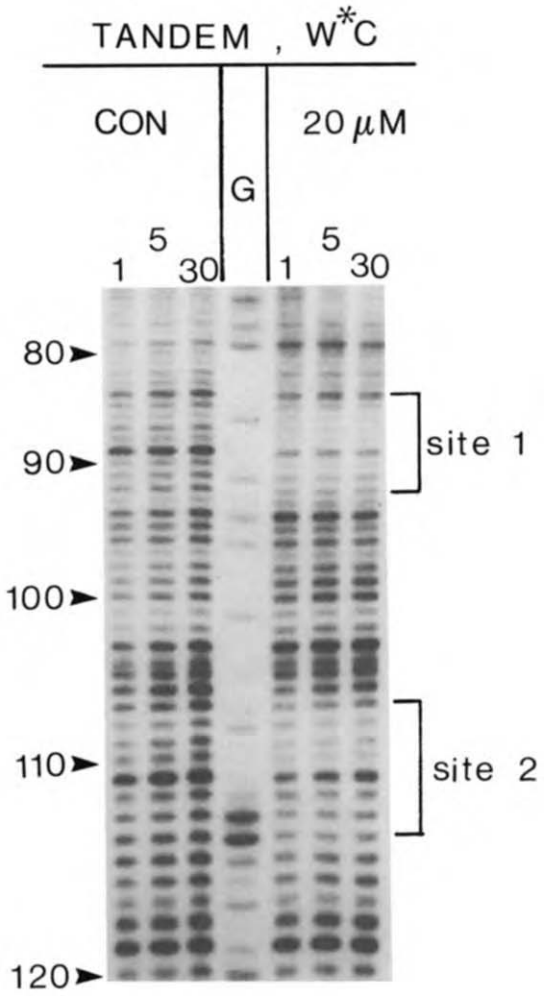

Fig.4. DNase I footprinting of TANDEM on the 160-base-pair DNA fragment with Watson strand bearing a radioactive 3'-end label. Labelling as in fig.2.

echinomycin, and the binding sites appear to be centred around the sequences ATAT and ATAA. The protection afforded by TANDEM is very much weaker, as expected since its binding constant for natural DNAs is known to be about 20-fold lower than that of echinomycin or of triostin A [2,3]. The binding of TANDEM appears to be rather specific, since not all AT-rich sequences are protected.

Enhancement of enzymatic cleavage at sequences flanking the binding sites can again be observed, especially on the Watson strand, although the phenomenon is very much less pronounced. The most prominent effect occurs between positions 78–81.

4. DISCUSSION

4.1. *Triostin A*

The DNase I protection patterns reveal that all the triostin A binding sites contain the dinucleotide step CpG. However, the sequence ACGC at position 35 is only weakly protected, in contrast with the same sequence at position 78. Perhaps the juxtaposition, in this case, of the CpG dinucleotide beside the long AT-rich region from positions 27–34 somehow hinders its interaction with the antibiotic. It is also difficult to determine whether CpG at position 95 is truly protected or not since it is only included in the blockage site on the Crick strand but not the Watson strand of the DNA.

The footprinting patterns obtained for triostin A do not differ significantly from those obtained with echinomycin [8]. As with echinomycin, the preferred binding sequence can be deduced to be XCGY, where X,Y are either purine or pyrimidine. The similarity of the footprints for the two antibiotics is surprising since their relative binding affinities for poly(dA-dT) and poly(dG-dC) are quite different, implying definite differences in the precise base sequence(s) of their preferred binding site(s) [5]. However, these polynucleotides may adopt different helical structures as compared to natural DNAs [15,16], so that binding data obtained with them may tend to exaggerate small differences in the intrinsic sequence preferences of the antibiotics. It may therefore be advisable to assess the footprinting patterns in the light of equilibrium binding data for natural DNAs. The equilibrium binding experiments reported several years ago showed that echinomycin and triostin A bind to various natural DNAs in much the same ranking order, although triostin A displays a smaller overall variation in binding constants than does echinomycin [7].

4.2. *TANDEM*

The preferred binding sites of TANDEM are clearly different from those of triostin A or echinomycin. The regions blocked do not contain the step CpG but rather are centred around the sequences ATA or TAT. This is perfectly consistent with equilibrium binding studies which revealed

that TANDEM binds best to those DNAs which are rich in AT residues [3]. The structure of TANDEM [17] is such as to predict that when it binds to double-helical DNA two base pairs are sandwiched between the quinoxaline chromophores in the *bis*-intercalated complex. Unfortunately, the present data are not sufficient to determine unambiguously whether the sandwiched base sequence is ApT or TpA, nor whether particular sequences are required in the flanking regions. It appears that structural differences must play a significant part in determining those ApT (or TpA) steps that will or will not bind the ligand. Most probably these dinucleotides will not provide TANDEM binding sites if they are either preceded or followed by runs of A or T. For example the ApT at position 31 is preceded by the sequence GAAA while that at position 47 is followed by TTTTCT: neither site is protected from DNase digestion by TANDEM. It has been suggested that runs of A and T possess a peculiar helical structure [12,18–20], and this may set up long-range changes affecting the ability of nearby ApT (or TpA) sequences to interact with the ligand.

Nevertheless, our experimental findings as summarised in fig.3 permit two further observations. If we assume that ApT is the sandwiched dinucleotide sequence, as suggested from the crystal structure of TANDEM [17], then it is noticeable that in both binding sites one of the flanking bases is A or T, continuing the alternating purine-pyrimidine sequence, while the other base can be G or C. On the other hand, it has been reported that poly d(TAC)·poly d(GTA) binds TANDEM while poly d(ATC)·poly d(GAT) does not, suggesting that the ligand recognises TpA rather than ApT sequences [21]. If this interpretation is correct, the preferred sequence appears to be XTAY where X and Y can be either A or T.

4.3. *Enhancement of susceptibility to nuclease attack*

The increases in cleavage rate at certain sequences observed in the presence of triostin A are similar to those seen with echinomycin [8], and occur at AT-rich regions adjacent to the binding sites. With TANDEM there is a very slight enhancement of cleavage at flanking GC-rich regions, observed mainly on the Watson strand.

A crystalline complex between triostin A and the hexanucleotide CGTACG has been described in which the TA step is a Hoogsteen base pair [22]. However, we do not consider it likely that the enhancement of cleavage at AT sequences observed here results from drug-induced Hoogsteen base-pairing for two reasons. Firstly, chemical reactions using dimethyl sulphate and bromoacetaldehyde have failed to detect any Hoogsteen pairings (H.R. Drew, personal communication). Secondly, similar enhancements are also produced by elevated temperature and dimethyl sulphoxide [12]. Hence the enhanced cleavages seen here are more likely to reflect some local helical distortion, caused by unwinding of adjacent regions of DNA by intercalation of the antibiotic [7]. It may well prove to be a general property of sequence-specific intercalators that they cause enhancement of enzymic cleavage at certain regions flanking their binding sites.

ACKNOWLEDGEMENTS

This work was supported by grants from the Medical Research Council, the Royal Society and the Cancer Research Campaign. R.K.O. acknowledges support from the National Institute of Allergy and Infectious Diseases (AI 15759). We thank Drs H.R. Drew, A.A. Travers and K.R. Fox for their help and Mrs T. Douglas for able technical assistance.

REFERENCES

[1] Katagiri, K., Yoshida, T. and K. (1974) in: Antibiotics III: Mechanism of Action of Antimicrobial and Antitumour Agents (Corcoran, J.W. and Hahn, F.E. eds) pp.234–251, Springer, Berlin.

[2] Waring, M.J. (1979) in: Antibiotics V/Part 2: Mechanism of Action of Antieukaryotic and Antiviral Compounds (Hahn, F.E. ed.) pp.173–194, Springer, Berlin.

[3] Lee, J.S. and Waring, M.J. (1978) Biochem. J. 173, 129–144.

[4] Waring, M.J. (1977) in: Drug Action at the Molecular Level (Roberts, G.C.K. ed.) pp.167–189, Macmillan, London.

[5] Lee, J.S. and Waring, M.J. (1978) Biochem. J. 173, 115–128.

[6] Fox, K.R. and Waring, M.J. (1981) Biochim. Biophys. Acta 654, 279–286.

[7] Waring, M.J. and Fox, K.R. (1983) in: Molecular Aspects of Anti-cancer Drug Action (Waring, M.J. and Neidle, S. eds) pp.137–141, Macmillan, London.
[8] Low, C.M.L., Drew, H.R. and Waring, M.J. (1984) Nucleic Acids Res. 12, 4865–4879.
[9] Van Dyke, M.W. and Dervan, P.B. (1984) Science, in press.
[10] Fox, K.R., Olsen, R.K. and Waring, M.J. (1982) Biochim. Biophys. Acta 696, 315–322.
[11] Ciardelli, T.L. and Olsen, R.K. (1977) J. Am. Chem. Soc. 99, 2806–2807.
[12] Drew, H.R. and Travers, A.A. (1984) Cell 37, 491–502.
[13] Van Dyke, M.W. and Dervan, P.B. (1983) Cold Spring Harbor Symp. Quant. Biol. 47, 347–353.
[14] Scamrov, A.V. and Beabealashvilli, R.S. (1983) FEBS Lett. 164, 97–101.
[15] Klug, A., Jack, A., Viswamitra, M.A., Kennard, O., Shakked, Z. and Steitz, T.A. (1979) J. Mol. Biol. 131, 669–680.
[16] Cohen, J.S., Wooten, J.B. and Chatterjee, C.L. (1981) Biochemistry 20, 3049–3055.
[17] 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.
[18] Fratini, A.V., Kopka, M.L., Drew, H.R. and Dickerson, R.E. (1982) J. Biol. Chem. 257, 14686–14707.
[19] Rhodes, D. (1982) in: Topics in Nucleic Acid Structure, Part II (Neidle, S. ed.) pp.287–304, Macmillan, London.
[20] Arnott, S., Chandrasekaran, R., Hall, I.H. and Puigjaner, L.C. (1983) Nucleic Acids Res. 11, 4141–4155.
[21] Evans, D.H., Lee, J.S., Morgan, A.R. and Olsen, R.K. (1982) Can. J. Biochem. 60, 131–136.
[22] 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, in press.