Research Article

Journal of Physical Organic Chemistry

Received: 28 August 2007, Revised: 12 December 2007, Accepted: 14 December 2007, Published online in Wiley InterScience: 7 May 2008

(www.interscience.wiley.com) DOI 10.1002/poc.1323

Molecular recognition and physicochemical properties in the discovery of selective antibacterial minor groove binders

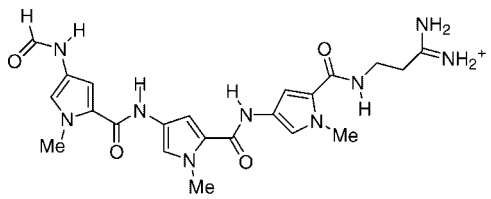
Colin J. Suckling*

The polyamide minor groove binders (MGBs), distamycin and netropsin, have been known for many years to have significant biological activities but high toxicity. Strategies are described for the development of more selective MGBs taking advantage of hydrophobic interactions with the minor groove of DNA. The introduction of branched alkyl side chain substituents, planar aromatic head groups and alkene isosteres of the amides have all been investigated. MGBs designed using these strategies and built from heterocyclic and aromatic amino acids with the ability to recognise short sequences of DNA have been found to be potent and selective antibacterial agents. Detailed structural and strength of binding investigations (NMR, capillary electrophoresis (CE), DNA footprinting, melting temperature measurement, ITC) show that their activity depends primarily upon molecular recognition in terms of both molecular shape and specific hydrogen bonding. However the lack of toxicity depends upon their basic tail group structure, the pK_a of which has a major influence on access to bacterial and mammalian cells. Lead compounds are active *in vivo* at doses competitive with recently introduced antibacterial drugs. Copyright © 2008 John Wiley & Sons, Ltd.

Keywords: DNA binding; minor groove binder; molecular recognition; antibacterial compounds; pK_a and log D

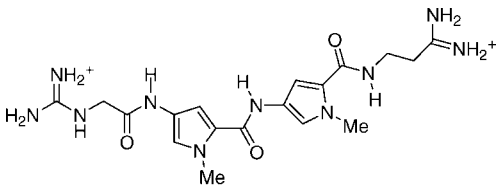
INTRODUCTION

Minor groove binders (MGBs) are compounds that possess an inherent curvature that matches approximately the helical curve of the minor groove of B-DNA. Their discovery was firstly in natural products, notably the oligoamides distamycin **1** and netropsin **2**, which are composed of short chains of *N*-methylpyrrole aminoacids with one or two cationic side chains.[1,2] Other natural products that bind to the minor groove include the spirocyclopropylquinone, CC1065, **3**[3] and the tricyclic alkaloid, anthramycin **4**.[4] Interest in these compounds comes not only from questions of molecular recognition for DNA but also from their potential as templates principally for anticancer drugs. Several of these templates (**1**, **3**, **4**) have been developed into compounds that have been taken into clinical trials for cancer treatment[5–9] and in all cases, a major part of their mechanism of action is forming covalent bonds with nucleophilic atoms of the DNA bases. In the case of the most active anthramycin derivatives, cross linking the strands of the DNA duplex occurs. In addition, there is significant selectivity for specific short sequences of DNA, which is a further important aspect of molecular recognition. In order to be effective anticancer drugs, these MGBs must cross the cell membrane and nuclear membrane to encounter human genomic DNA in the cell nucleus. Thus for success in discovering a selective drug, not only must recognition of the target be considered, in this case DNA, but also the ability of the compound to reach the target must be recognised. Lipid bilayers are major components of cell membranes and for compounds to cross such membranes, their physicochemical properties become very important; unless an active transport mechanism is available, highly polar molecules would not be expected to gain access to cells. Therefore in addition to molecular recognition, which depends upon shape and specific bonding interactions between DNA and the MGB ligand, it is important to take account of the lipophilicity and ionic state of potential drugs. This paper discusses how consideration of both molecular recognition and physicochemical properties has led to the discovery of selective antibacterial MGBs. Whilst several clear concepts have featured in this journey, the choices of the direction of research have had a substantial empirical component.

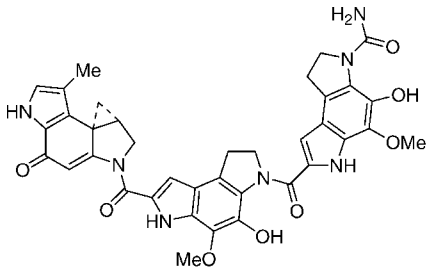


distamycin **1**

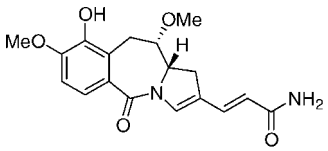
* *C. J. Suckling, WestCHEM Research School, Department of Pure & Applied Chemistry, University of Strathclyde, 295 Cathedral Street, Glasgow, G1 1XL, Scotland, UK.*
E-mail: c.j.suckling@strath.ac.uk



netropsin **2**



CC1065 **3**



anthramycin **4**

Oligoamide MGBs

The oligoamides, distamycin **1** and netropsin **2**, are highly polar compounds that nevertheless show significant cytoxic properties. However their ability to bind to DNA has made possible extensive studies on the sequence selectivity of binding which in turn has led to further developments in therapeutic and diagnostic applications.[8,9] It has been shown that distamycin and netropsin bind preferentially to AT-rich regions of DNA; binding adjacent to a GC base pair is hindered by a steric clash between H3 of the pyrrole rings and the 2-amino group of guanosine (Fig. 1).

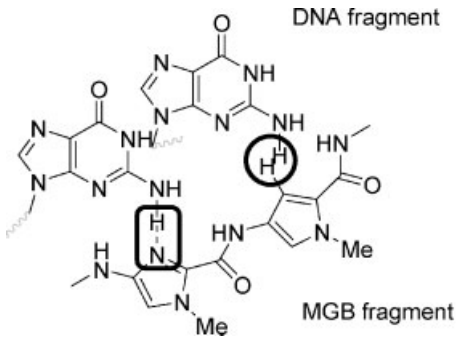
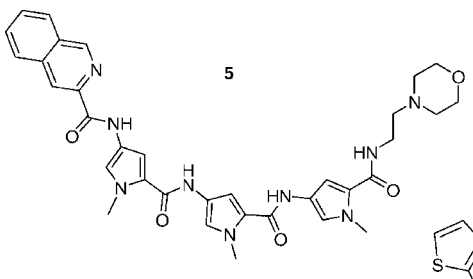


Figure 1. Favourable and unfavourable interactions in the binding of oligoamide MGBs composed of pyrroles and imidazole to DNA

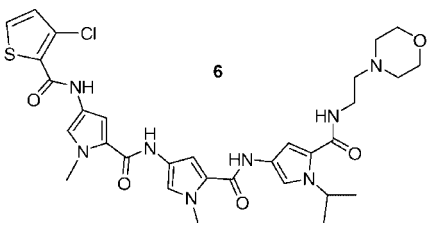
Therefore, in order to obtain distamycin and netropsin analogues that bind to GC segments of DNA, the steric clash was removed by replacing the *N*-methylpyrrole by *N*-methylimidazole[10–13] (Fig. 1). This seminal discovery has led to the synthesis of many hundreds of MGBs able to bind selectively to sequences of 20 and more base pairs and to studies of the uses of such compounds in controlling the function of DNA in cells.[14–18] Because of their large size, however, such MGBs are not normally considered appropriate for development as drugs.

In view of the sequence selectivity obtainable and the intrinsic activity of distamycin and netropsin, several groups have sought analogues that might form the basis of selective antibacterial or antifungal drugs.[19–25] Typical examples include compounds from Genesoft, **5** and **6** which notably feature extended and more lipophilic *N*-terminal groups, conventionally known as the head groups.[21–24] However the mechanism by which such compounds achieve antibacterial selectivity is not known and a quotation from one of Genesoft's papers is eloquent: 'Even though compound **6** appears to kill bacteria by interacting with A/T-rich DNA target sites ($K_d = 7.5$ nM) as demonstrated with earlier prototype molecules, its unique properties cannot be explained in a rational way'.

One factor that does not appear to have been strongly considered is drugs' crossing the cell wall of bacteria. Bacterial cell walls are much more complex than mammalian cell membranes and consist of several coats including lipid bilayers, as in



5



6

mammalian cell membranes, but importantly cross linked peptide-glycans which provide the physical strength for maintaining the integrity of the bacterial cell. The biosynthesis of these peptidoglycans is the target for the action of β-lactam antibiotics including the penicillins. The cell walls of the two main classes of bacteria, Gram positive and Gram negative, also differ from each other in terms of detailed molecular structure and overall architecture; these differences could form the basis for selectivity of action of a drug between Gram-positive and Gram-negative bacteria.

STRATHCLYDE MGBs

The Strathclyde concept: first and second strategies

At Strathclyde, we took the view that in order both to maximise hydrophobic interactions with the minor groove of DNA and to give the best chance of penetrating a bacterial cell wall, it was necessary to design MGBs with significantly increased lipophilicity than those described so far. The base of the minor groove in unliganded DNA contains a spine of hydrogen-bonded water molecules which are replaced by the MGB when it binds. The walls of the minor groove have substantial hydrophobic patches suitable for interaction with *N*-alkyl substituents (Fig. 2). Therefore to achieve improved binding, whilst leaving hydrogen bonding opportunities through the amide NHs intact, we firstly introduced *N*-alkyl groups larger than methyl, of which isopropyl proved especially important and secondly, we used *C*-alkylthiazoles in place of *N*-methylimidazoles to accommodate binding at GC sites. Thiazoles had been introduced before[26] but the use of

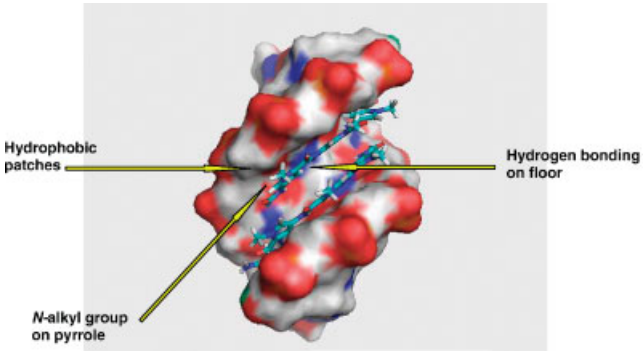
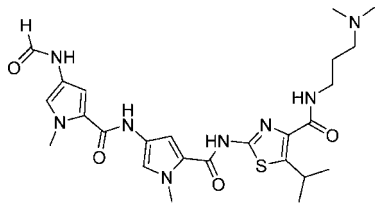
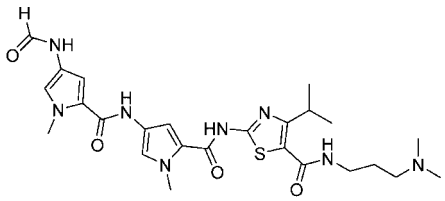


Figure 2. The basic design concept used in the development of the structures of Strathclyde MGBs. The representation of the DNA minor groove colours hydrophobic patches grey, negatively charged regions red and positively charged regions blue and is based upon NMR structures.

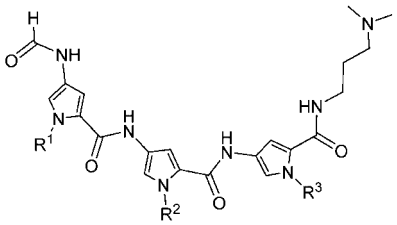
C-alkylthiazoles as substitutes for *N*-alkylimidazoles was novel. Examples of the compounds made include **7**–**12** and we have investigated the binding of these compounds by capillary electrophoresis (CE).[27] MGB **7**, which we call thiazotropsin A, incorporates both the thiazole and the isopropyl group and, as discussed below, binds strongly and selectively to DNA. The lipophilicity of thiazotropsin A as assessed by its log P value of 0.61 is over three log P units more positive than that of distamycin (**1**, log P = −2.94).



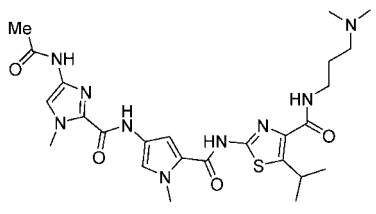
Thiazotropsin A **7**



8



	R^1	R^2	R^3
9	Me	*i*-Pr	Me
10	*i*-Pr	Me	*i*-Pr
11	*i*-Pr	*i*-Pr	*i*-Pr



12

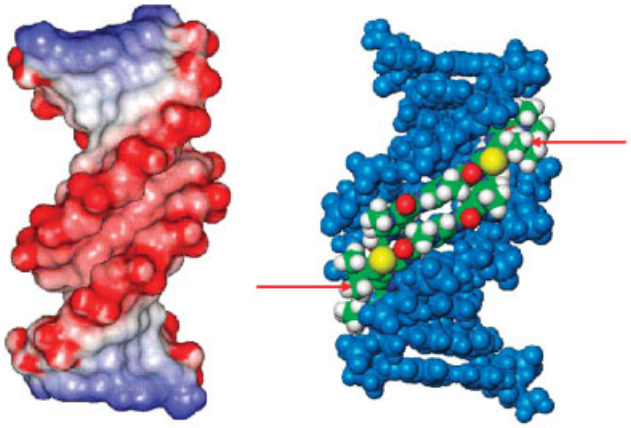


Figure 3. Left: canonical B-DNA with a standard width minor groove (pink and red shaded). Right: oligonucleotide $(CGACTAGTCG)_2$ (blue) with two molecules of thiazotropsin A bound in an expanded minor groove. The isopropyl groups pointing out of the minor groove and contributing strongly to dimer formation and binding are indicated by the red arrows. This structure was obtained from NMR studies[31].

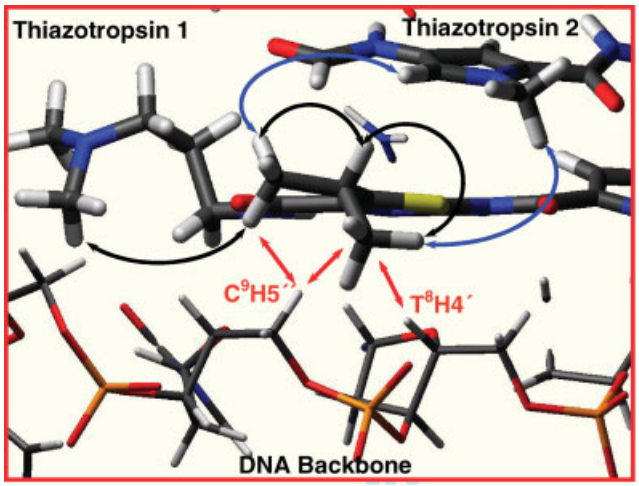


Figure 4. The intimate relationship of the formyl head group of the top molecule of thiazotropsin A (thiazotropsin 2) with the lower molecule (thiazotropsin 1) and the DNA backbone as shown by NMR.[31] The NOEs that give rise to this structure are shown by the red, blue and black coloured arrows that respectively show thiazotropsin 1 to DNA interactions, thiazotropsin 1 to thiazotropsin 2 interactions and intra thiazotropsin 1 interactions.

Amongst the notable features of thiazotropsin A is the fact that the thiazole isomer **8** does not bind to DNA,[28,29] which is attributable to the most unfavourable direct clash between the isopropyl group and the base of the minor groove of DNA. This result has not been studied in detail but it is possible that the exclusion of **8** relates to the side-by-side binding of MGBs to DNA, a possibility that becomes clearer when further compounds are considered. MGBs **9**, **10** and **11** contain respectively one, two and three *N*-isopropyl groups. Whereas both **9** and **10** were found to bind to AT-rich DNA (3′-AAATTATATTAT-5′/3′ATAATATAATTT-5′) by CE, **11** did not bind. At first sight, this result might appear paradoxical bearing in mind the above arguments about hydrophobic interactions, but when the stoichiometry and geometry of binding of such MGBs to DNA is considered, the reason is clear. CE showed that these MGBs bind to DNA in 2:1 ratio of MGB to duplex, consistent with X-ray crystallographic data for a DNA/distamycin complex.[30] In forming the complex, the minor groove broadens to accommodate two ligand molecules which are arranged closely side by side and antiparallel (Fig. 3). Bringing MGB **11** into this arrangement imposes unfavourable interactions between the *N*-isopropyl groups of the two molecules of MGB sufficient to prevent strong binding to DNA in this mode. These observations also emphasise the nature of DNA as a dynamic host in molecular recognition, capable of major conformational changes on ligand binding. It is important to note that the biological activity of many DNA-binding drugs can be attributed to the inhibition of binding of transcription factors and other proteins consequent upon conformational change.

Thiazotropsins A and B (*7* and 12)

In seeking an eventual biological application that might depend upon a DNA sequence, it was important to establish the sequence selectivity of our MGBs. In collaboration with Professors Keith Fox and Tom Brown (University of Southampton) we have studied the binding of our MGBs to specially designed sequences of DNA that contain all possible combinations of four base pairs probed by DNA footprinting.[27,31] Thiazotropsin A was found to bind especially strongly to the sequence (A)CTAG(T) and this result was confirmed in our laboratories by CE using the dodecamer CGACTAGTCG as target.[28,29] NMR studies using the same DNA oligomer allowed us to determine the details of binding of thiazotropsin A to DNA (Figs 3 and 4).[32] Notable features included a specific hydrogen bond between G and the thiazole nitrogen and that the antiparallel arrangement avoids steric hindrance between the *C*-isopropyl thiazole substituents. The side-by-side interactions of the two bound MGBs can be seen to be important from the intimate relationship between the *iso*-propyl group of one molecule of thiazotropsin A and the second molecule's *N*-methyl group. As noted above, it can also be seen that the minor groove has widened to accommodate the two ligands.

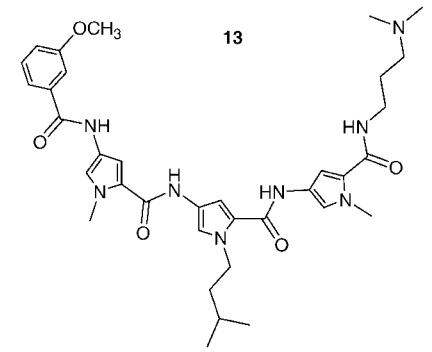
Thiazotropsin B, **12**, contains an imidazole in place of the *N*-terminal pyrrole of thiazotropsin A; this would be expected to promote binding to the sequence ACGCGT and strong footprints were found at this site.[31] Confirmation of the selectivity of binding came from melting temperature measurements which showed substantial increases in the range 12–19°C but only when thiazotropsin A bound to its preferred target, ACTAGT, and thiazotropsin B bound to its preferred target, ACGCGT. We thus have significant sequence preferences for short MGBs binding to short oligonucleotides. The increases in melting temperature are quite large for short oligonucleotides and typical of what we find for MGBs of this class. In order to determine the strength of binding corresponding to such melting temperatures, we have examined the affinity of thiazotropsin A to DNA by isothermal titration calorimetry. Binding was found to be exothermic with a K_d approximately 50 nM, indicative of strong binding to its preferred target sequence and a very acceptable value for a potential drug. It is worth emphasising, however, that such selectivity is not absolute in the sense that one might consider the affinity of a potent inhibitor for an enzyme. MGBs interact with several closely related sequences in a strand of DNA duplex

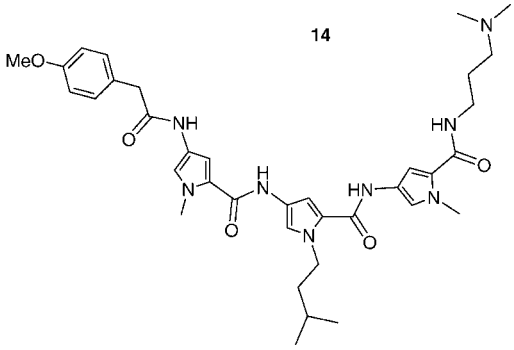
Journal of Physical Organic Chemistry

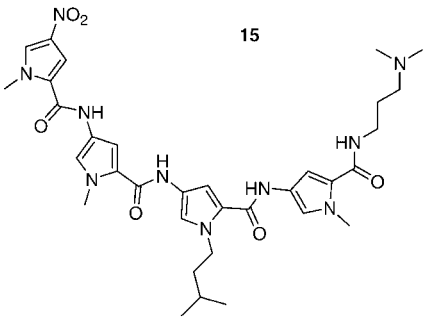
but in general show a preference for one sequence, typically of at least 100-fold in our experience.

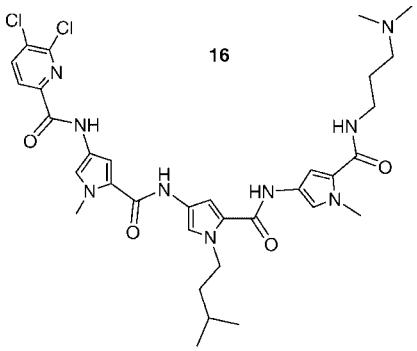
The third and fourth strategies

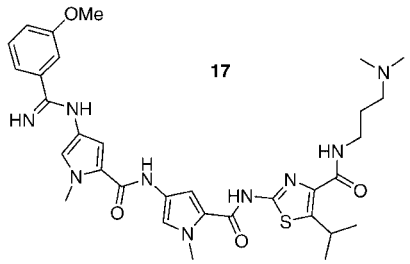
Having established the significance of *N*-alkyl groups and thiazoles in binding MGBs to DNA through the pattern of heterocyclic groups, attention turned to the head and tail groups. Conventionally, the head groups of many MGBs are formyl groups, which cannot contribute to hydrophobic interactions. Similarly, the tail groups in most studied compounds were simple tertiary aliphatic amines, which are largely protonated at physiological pH. The Genesoft scientists[21–24] had recognised this as shown by their inclusion of hydrophobic head groups and the substantially less basic alkyl morpholine tail group (**5**, **6**). Genelabs had also used *N*-alkyl groups in their compounds. Table 1 shows a small selection of Strathclyde MGBs to illustrate the effect of changing the head and chain groups. In general, hydrophobic head groups with a hydrogen bonding atom or substituent are required for optimal antibacterial activity (although not simply for binding to DNA). Conversely, the most hydrophilic compounds such as thiazotropsin B are inactive. Bearing in mind that we are dealing with living cells, it would be expected that biological activity would respond to more than one property of a potential drug; metabolism may, for example, be important.

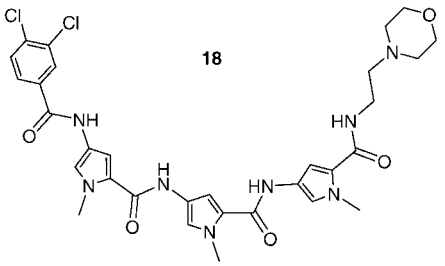


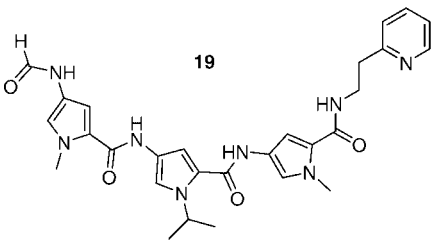












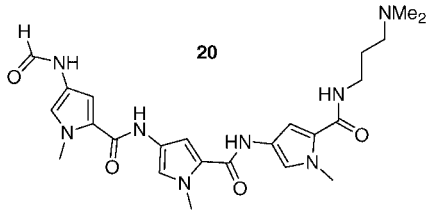


Table 1. Selected biological activities in relation to structure of Strathclyde oligoamide MGBs

Compound	Head group	Hydrogen bonding	Lipophilic feature	*S. aureus* MIC (μg/ml)	*A. niger* MIC (μg/ml)
7	Formyl	Thiazole	*C-i*-propyl	4.7	76.1
12	Acetyl	Imidazole and thiazole	*C-i*-propyl	Inactive	Inactive
13	3-MeO benzoyl	3-MeO	*N*-pentyl and head group	2.0	Inactive
14	4-MeO phenylacetyl	4-MeO	*N*-pentyl and head group	15.9	15.9
15	3-NO_2-pyrrolyl	3-NO_2	*N*-pentyl and head group	31.6	15.8
16	2,3-DiCl pyridyl	Pyridine N	*N*-pentyl and head group	8.0	128
17	3-MeO benzamidine	3-MeO and thiazole	*C-i*-propyl and head group	3.1	Inactive
18	2,3-Dichloro benzoyl	Morpholine tail group	Head group	16.0	Inactive

For comparison, the minimum inhibitory concentration (MIC) of the standard antibacterial drug, ampicillin, against *Staphylococcus aureus* in the same tests was 16.1 μg/ml. The standard antifungal drug had an MIC of >300 μg/ml against *Aspergillus niger* in the same tests.

A significant component of the binding energy of MGBs to DNA can be expected from ionic interactions between the protonated tail group and the phosphate anions of DNA. It was therefore of interest to investigate the extent to which this interaction is important. We therefore prepared the MGB **19**, which bears a 2-pyridyl side chain of p$K_a \sim 6$, for comparison with **20**, which bears the standard dimethylaminopropyl tail of p$K_a \sim 9.5$ (Suckling, Yule, unpublished results). The affinities of these compounds were evaluated by T_m measurements using the oligonucleotide duplex 3′-AAATTATATTAT-5′/3′ATAATATAATTT-5′ as a function of pH. Table 2 shows that the affinity as measured by the difference between the T_m of the oligonucleotide itself compared with that of the MGB complexes decreases substantially for the pyridyl compound **19** at pHs above its pK_a as the degree of protonation decreases. On the other hand, for the more basic trialkylamino tail group MGB, **20**, the difference in T_m remains more or less constant. The protonation state of the tail group in alkyl amines is evidently important in binding to DNA.

Table 2. Binding to DNA as a function of pH by MGBs differing in the pK_a of their tail group. The target sequence was AAATTATATTAT/ATAATATAATTT

pH	ΔT_m (°C) of **20**	ΔT_m (°C) of **19**
5.0	19.6	20.1
7.0	20.1	14.1
9.0	16.1	10.1

Oligonucleotide concentration was 2 μM and MGB concentration 4 μM in PBS (sodium phosphate, 10 mM; sodium chloride, 0.3 M).

Tail group pK_a is also important in at least one more context, namely transport across biological cell membranes and cell walls. As explained in the introduction, the lipid bilayer component of such structures presents a serious barrier to the passive transport of polar compounds. Clearly, the ionic state of acids and bases is very significant in this context and it might therefore be expected that the pK_a of the tail group would be an important consideration. With our biological data, we can compare three different types of tail group namely amidines, such as distamycin (**1**) itself, simple trialkyl amines, such as the common dimethylaminopropyl, and *N*-morpholinoethyl, as introduced by Genesoft.[21–24] Table 3 summarises the structures, relevant physicochemical properties and biological characteristics.

The usual measure of the lipophilicity of a compound is the octanol-water partition coefficient, expressed as log P. Compounds with log P in the range 2–4 are usually considered to be possible drugs. However with ionised compounds, the fraction ionised at physiological pH (7.4) must be used, taking the view that non-ionised molecules will more effectively cross the lipid bilayer by passive diffusion. This is expressed by the

Table 3. Antibacterial activity of MGBs as a function of the pK_a of the tail group and its implications for distribution and partition

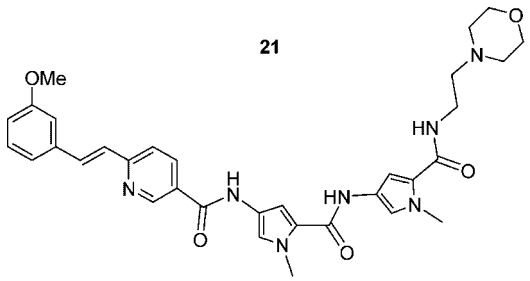
Compound	Tail	pK_a	log P	log $D^{7.4}$	Structural feature	Biological activity
1	**Amidine**	**12.4**	**−2.94**	**−7.94**	**Amidine**	**Toxic**
20	*t*-Amine	9.99	−1.98	−4.57	No lipophilic groups	Weak antibacterial
12	*t*-Amine	9.99	0.85	−1.74	Imidazole and *C*-alkylthiazole	Inactive antibacterial
7	***t*-Amine**	**9.99**	**0.61**	**−1.98**	***C*-alkylthiazole**	**Antibacterial and antifungal**
13	*t*-Amine	9.99	1.25	−1.34	H-bond head and *N*-pentyl	Selective antibacterial
21	Morpholine	7.41	1.10	0.79	H-bond head and alkene	Selective antibacterial
22	**Morpholine**	**7.41**	**2.22**	**1.91**	**H-bond head and alkene**	**Selective antibacterial**

Log P values are calculated. The bold entries highlight examples of major significance to the discussion. Numerical values for biological activity of new MGBs in this study are given in Table 1.

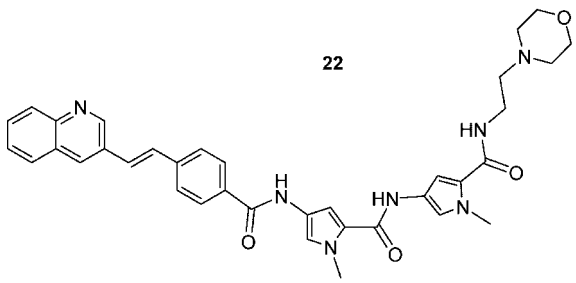
Journal of Physical Organic Chemistry

distribution coefficient expressed as log $D^{7.4}$.[33] The tendency from the data in Tables 1 and 3 is that antibacterial activity is best found in those compounds that possess the least basic tail group, the morpholinoethyl group; these compounds also have the highest log D. Such an effect has also been found by others.[21–24] On the other hand, the toxicity of distamycin, which has the highest p*K* tail group, an amidine, stands out. Whilst the data are clear, this leaves a problem to provide a molecular mechanism for the toxicity of distamycin. A possible solution is that amidines are transported actively into mammalian cells. There is some evidence that distamycin binds to AT-rich segments of control regions in mammalian DNA,[34] the sequence selectivity being consistent with what we would expect from its structure. Moreover, many of the anticancer MGBs also have amidine tail groups and clearly they must also penetrate mammalian cells and nuclei in order to exert the anticancer activity.[35] The properties of tail group amidines may therefore be principally due more to an accident of biology in providing an active transport mechanism than to the physical chemical properties of the MGBs themselves. This reinforces the previous comment that the ultimate biological properties of an MGB depend upon a balance between many structural factors.

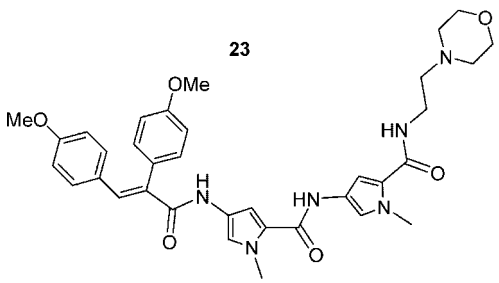
21



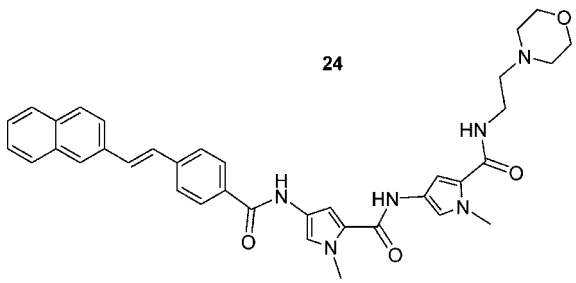
22

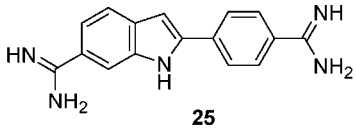


23



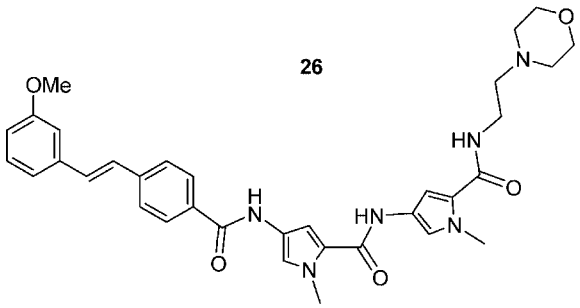
24





25

26



A fifth strategy: peptide isosteres

The antibacterial activity of most of the compounds in Table 3 was not sufficient to warrant investigation *in vivo* and we wished to improve the activity. In order to achieve this, a further approach to increasing the lipophilicity of our compounds was introduced, namely a peptide isostere. By replacing an amide with the isosteric alkene, a compound with essentially the same geometry but without the hydrogen bond donor and acceptor of an amide is obtained. This structural change led to an order of magnitude improvement in the antibacterial activity giving MGBs **21** (1.10log P) and **22** (2.22log P) which are effective at less than 1 μg/kg in *in vitro* tests and at 20–40 mg/kg in *in vivo* tests in mice, both against *Staphylococcus aureus*, including methicillin resistant strains (MRSA). The alkene-containing MGBs were also shown to be bactericidal.[36,37] Moreover, an approximately 1000-fold higher concentration was required to produce toxicity to mammalian cell lines than the concentration required to kill bacterial cells. These most encouraging results made it important to understand the physicochemical properties of the new MGBs.

There appear to be several parallels with the all amide MGBs, for example the importance of a hydrogen bonding head group; MGB **22** is one of our most active compounds and contains a quinolyl head group but its naphthyl analogue **24** is completely inactive. Other similar comparisons reinforce the conclusion that a hydrogen bonding head group is required for antibacterial activity. But despite this activity and the obvious geometric similarity between amide and alkene, the actual affinity of alkene-containing MGBs for DNA cannot be presumed. CE studies have shown that **22** binds to the DNA oligomers GCGATATATGCG/CGCTATATACGC in a clean 2:1 ratio (Waigh, Araya, unpublished results). We have therefore compared alkene-containing MGBs and their amide partners with hydrogen bonding head groups known to produce high antibacterial activity and studied their binding to DNA oligomers by melting temperature measurements (Donoghue, Suckling, unpublished results). For example, for the alkene-containing MGB **22** binding also to GCGATATATGCG/CGCTATATACGC, an increase in T_m on complexation of 16°C was observed; the corresponding figure for the amide analogue was 18°C. From such measurements at least, the behaviour of alkene-containing MGBs seems similar to that of their traditional amide analogues. However it is important to recognise that the selected target oligonucleotide may not be an ideal binding site for the different MGBs. In anti-infective applications, for example, other sites in DNA may be hit and these points are now under investigation.

There remains the question of biological selectivity. By choosing to synthesise alkene-containing MGBs, another significant property was consciously built in, namely fluorescence. Fluorescent labels have been attached to large amide MGBs in order to follow their distribution into cells but these labels are additions to the MGB structure and cannot be presumed to have no effect on uptake into cells and internal distribution therein.[44,45] In contrast, our fluorescent MGBs are intrinsically fluorescent and should therefore report accurately their cellular uptake and location. All of the alkene head groups lead to fluorescent properties to some extent but the MGB **23** was both active in antibacterial screens and sufficiently fluorescent for us to investigate its transport into cells of different types by microscopy (Ellis, Hunter, Khalaf, Suckling, unpublished results). These cells were treated with the MGB **23** and it was allowed to diffuse into the cells if it could. Fluorescence at the expected wavelength of emission for **23** was observed with *S. aureus* cells but no uptake was visible with *E. coli* cells. This is consistent with the observed biological activity because **23**, like most MGBs of this type, is inactive against *E. coli*. A similar experiment was undertaken with V79 cells, a mammalian cell line. As a control, the synthetic MGB, DAPI **25**, which is a fluorescent, amidine-containing compound known to penetrate mammalian cells and nuclei, was used. Whilst it was clear that DAPI does indeed reach the nuclei of V79 cells, as expected, the other hand, the uptake of **23** by V79 cells was at best weak and penetration into the nucleus was not observed. Notably, compounds in the polyamide MGB class with amidine tail groups have proved to be very successful as anti-tumour agents and are in clinical trials,[8,9] a situation which emphasises the importance of the tail group.

The great steps forward in potential therapeutic applications of the alkene-containing MGBs, however, come at a cost. The solubility of these compounds in physiological buffers such as phosphate buffered saline (typically phosphate buffer, 10 mM pH 7.4, with sodium chloride 50 mM) is low. At concentrations above 50 μM, turbidity is observed and in some cases, for example **21**, precipitation as an oil then occurs. That intermolecular association should occur with alkene-containing MGBs is no surprise taking into account the lipophilic head and ionic tail. With a view to designing out the solubility problem whilst maintaining DNA binding and antibacterial activity, we have begun studies of self-association by NMR (Suckling, Parkinson, Breen, unpublished results). The first experiments in DMSO solution with **26** showed clearly from concentration and temperature variation that intermolecular association takes place with a dissociation constant of approximately 2.5 mM. However the more significant point was that the data did not support stacking of the hydrophobic sections of the MGBs, as might be expected from its structure. A more appropriate description was an infinite planar array. It is not immediately obvious what the detailed geometry of such an array might be and tiling arrays of various types are possibilities. Further investigations of these phenomena in water are in hand.

Taken together, the above results define the structural and physicochemical properties of compounds required for high antibacterial activity and suggest ways in which the profiles of apparently similar MGBs can be modified in order to provide a useful biological or therapeutic outcome. In summary, for antibacterial activity, we can identify the following:

- A hydrophobic head group bearing a hydrogen bonding atom or substituent.
- A sequence of aryl or heteroaryl groups that permit antiparallel binding in the minor groove with a 2:1 ratio.
- A tail group of pK_a greater than 5 but less than 9.
- The ability to bind to DNA at least at a single GC site within predominantly AT-rich sequences.
- Studies of the therapeutic potential and the associated physicochemical properties of these compounds continue.

Acknowledgements

Financial support from the Synergy Fund, from the Scottish Executive Proof of Concept Fund and from the University of Strathclyde is acknowledged with thanks. This paper is an edited version of a plenary lecture presented at the 2007 European Symposium on Organic Reactivity in Faro, Portugal.

REFERENCES

[1] L. Grehn, U. Ragnarsson, B. Eriksson, B. Oberg, *J. Med. Chem.* **1983**, *26*, 1042–1049.
[2] A. C. Finlay, F. A. Hochstein, B. A. Sobin, F. X. Murphy, *J. Am. Chem. Soc.* **1951**, *73*, 341–343.
[3] L. J. Hanka, A. Dietz, S. A. Gerphiede, S. L. Kuentzel, D. G. Martin, *J. Antibiot.* **1978**, *31*, 1211–1217.
[4] W. Leimgruber, A. D. Batcho, F. Schenker, *J. Am. Chem. Soc.* **1965**, *87*, 5793–5795.
[5] S. R. Alberts, C. Erlichman, J. M. Reid, J. A. Sloan, M. M. Ames, R. L. Richardson, R. M. Goldberg, *Clin. Cancer Res.* **1998**, *4*, 2111–2117.
[6] D. L. Boger, D. S. Johnson, W. Yun, C. M. Jarby, *Bioorg. Med. Chem.* **1994**, *2*, 115–135.
[7] S. J. Gregson, P. W. Howard, D. R. Gullick, A. Hamaguchi, K. E. Corcoran, N. A. Brooks, J. A. Hartley, T. C. Jenkins, S. Patel, M. J. Guille, D. E. Thurston, *J. Med. Chem.* **2004**, *47*, 1161–1174.
[8] A. J. Ten Tije, J. Verweij, A. Sparreboom, A. van der Gaast, C. Fowst, F. Fiorentini, J. Tursi, A. Antonellini, M. Mantel, C. M. Hartman, G. Stoter, A. S. T. Planting, M. J. A. de Jonge, *Clin. Cancer Res.* **2003**, *9*, 2957–2964.
[9] A. Fedier, C. Fowst, J. Tursi, C. Geroni, U. Haller, S. Marchini, D. Fink, *Br. J. Cancer* **2003**, *89*, 1559–1565.
[10] B. P. Reddy, S. M. Sondhi, J. W. Lown, *Pharmacol. Ther.* **1999**, *84*, 1–111.
[11] B. H. Geierstanger, M. Mrksich, P. B. Dervan, D. E. Wemmer, *Science* **1994**, *266*, 646–650.
[12] P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.
[13] J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1996**, *382*, 559–561.
[14] S. K. Sharma, J.-N. Billaud, M. Tandon, O. Billet, S. Choi, M. L. Kopka, T. R. Phillips, J. W. Lown, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2007–2010.
[15] L. A. Dickinson, R. J. Gulizia, J. W. Trauger, E. E. Baird, D. E. Mosier, J. M. Gottesfeld, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12890–12895.
[16] L. A. Dickinson, R. Burnett, C. Melander, B. S. Edelson, P. S. Arora, P. B. Dervan, J. M. Gottesfeld, *Chem. Biol.* **2004**, *11*, 1583–1594.
[17] C. Bailly, J. B. Chaires, *Bioconjug. Chem.* **1998**, *9*, 513–538.
[18] K. Maeshima, S. Janssen, U. K. Laemmli, *EMBO J.* **2001**, *20*, 3218–3228.
[19] A. I. Khalaf, A. R. Pitt, M. Scobie, C. J. Suckling, J. Urwin, R. D. Waigh, S. C. Young, R. V. Fishleigh, W. A. Wylie, *Tetrahedron* **2000**, *56*, 5225–5239.
[20] A. I. Khalaf, A. R. Pitt, M. Scobie, C. J. Suckling, J. Urwin, R. D. Waigh, S. C. Young, R. V. Fishleigh, *J. Chem. Res.* **2000**, 264–265.
[21] R. W. Bürli, Y. Ge, S. White, S. M. Touami, M. Taylor, J. A. Kiazerman, H. E. Moser, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2591–2594.
[22] R. W. Bürli, D. McMinn, J. A. Kaizerman, W. Hu, Y. Ge, Q. Pack, V. Jiang, M. Gross, M. Garcia, R. Tanaka, H. E. Moser, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1253–1257.
[23] R. W. Bürli, P. Jones, D. McMinn, Q. Le, J.-X. Duan, J. A. Kaizerman, S. Difuntorum, H. E. Moser, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1259–1263.
[24] R. W. Bürli, J. A. Kaizerman, J.-X. Duan, P. Jones, K. W. Johnson, M. Iwamoto, K. Truong, W. Hu, T. Stanton, A. Chen, S. Touami, M. Gross, V. Jiang, Y. Ge, H. E. Moser, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2067–2072.

[25] A. I. Khalaf, R. D. Waigh, A. J. Drummond, B. Pringle, I. McGroarty, G. G.. Skellern, C. J. Suckling, *J. Med. Chem.* **2004**, *17*, 2133–2156.
[26] S. K. Sharma, M. Tandon, J. W. Lown, *Tetrahedron* **2002**, *58*, 3417–3421.
[27] I. I. Hamdan, G. G. Skellern, R. D. Waigh, *Nucleic Acids Res.* **1998**, *26*, 3053–3058.
[28] N. G. Anthony, K. R. Fox, B. F. Johnston, A. I. Khalaf, S. P. Mackay, I. S. Mc Graorty, J. A. Parkinson, G. G. Skellern, C. J. Suckling, R. D. Waigh, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1353–1356.
[29] P. L. James, E. E. Merkina, A. I. Khalaf, C. J. Suckling, R. D. Waigh, T. Brown, K. R. Fox, *Nucleic Acids Res.* **2004**, *32*, 3410–3417.
[30] S. N. Mitra, M. C. Wahl, M. Sundaralingam, *Acta Cryst. D* **1999**, *55*, 602–609.
[31] A. J. Hampshire, H. Khairallah, A. I. Khalaf, A. H. Ebraimabadi, R. D. Waigh, C. J. Suckling, T. Brown, K. R. Fox, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3469–3474.
[32] N. G. Anthony, B. F. Johnston, A. I. Khalaf, S. P. MacKay, J. A. Parkinson, C. J. Suckling, R. D. Waigh, *J. Am. Chem. Soc.* **2004**, *126*, 11338–11349.
[33] A. I. Khalaf, A. R. Pitt, M. Scobie, C. J. Suckling, J. Urwin, R. D. Waigh, S. C. Young, R. V. Fishleigh, K. R. Fox, *J. Med. Chem.* **2000**, *43*, 3257–3266.
[34] A. Taylor, K. A. Webster, T. A. Gustafson, L. Kedes, *Mol. Cell. Biochem.* **1997**, *169*, 61–72.
[35] A. Lansiaux, F. Tanious, Z. Mishal, L. Dassoneville, A. Kumar, C. E. Stephens, Q. Hu, W. D. Wilson, D. W. Boykin, C. Bailly, *Cancer Res.* **2002**, *62*, 7219–7229.
[36] A. I. Khalaf, R. D. Waigh, C. J. Suckling, UK Patent Application 0619325.4 (30 September **2006**).
[37] C. J. Suckling, D. Breen, A. I. Khalaf, E. M. Ellis, I. S. Hunter, G. Ford, C. G. Gemmell, N. Anthony, J.-J. Helsebeux, S. P. Mackay, R. D. Waigh, *J. Med. Chem.* **2007**, *50*, 6116–6125.
[38] J. M. Gottesfeld, C. Melander, R. K. Suto, H. Raviol, K. Luger, P. B. Dervan, *J. Mol. Biol.* **2001**, *309*, 615–629.
[39] B. Dudouet, R. Burnett, L. A. Dickinson, M. R. Wood, C. Melander, J. M. Belitsky, B. Edelson, N. Wirts, C. Breihn, P. B. Dervan, J. M. Gottesfeld, *Chem. Biol.* **2003**, *10*, 859–867.
[40] K. S. Crowley, D. P. Phillion, S. S. Woodard, B. A. Schweitzer, M. Singh, H. Shabany, B. Burnette, P. Hippenmeyer, M. Heitmeier, J. K. Bashkin, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1565–1570.
[41] S. K. Sharma, A. T. Morrissey, G. G. Miller, W. H. Gmeiner, J. W. Lown, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 769–772.
[42] E. J. Fechter, B. Olenyuk, P. B. Dervan, *J. Am. Chem. Soc.* **2005**, *127*, 16685–16691.
[43] G. Nickols, C. S. Jacobs, M. E. Farkas, P. B. Dervan, *Nucleic Acids Res.* **2007**, *35*, 363–370.
[44] K. Kayser, K. Baumgartner, H. J. Gabius, *Anal. Quant. Cytol. Histol.* **1996**, *18*, 115–120.
[45] K. Jansen, P. N. Lincoln, *Biochemistry* **1993**, *32*, 6605–6612.