

## A Role for Water Molecules in DNA–Ligand Minor Groove Recognition

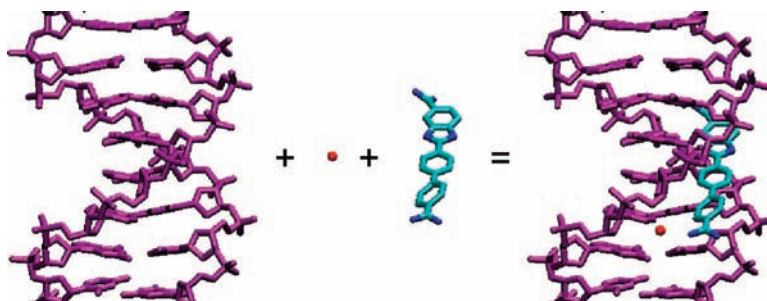
BINH NGUYEN,<sup>†</sup> STEPHEN NEIDLE,<sup>\*,‡</sup> AND  
W. DAVID WILSON<sup>\*,†</sup>

<sup>†</sup>Department of Chemistry, Georgia State University, Atlanta, Georgia 30302,

<sup>‡</sup>CRUK Biomolecular Structure Group, The School of Pharmacy, University of London, London WC1N 1AX, U.K.

RECEIVED ON JANUARY 16, 2008

### CON SPECTUS



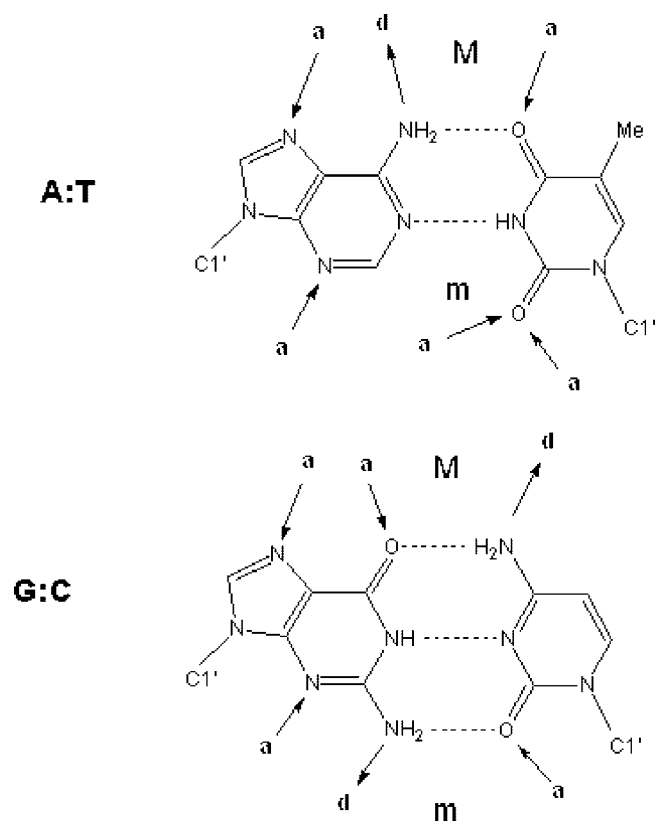
**T**argeting the minor groove of DNA through binding to a small molecule has long been considered an important molecular-recognition strategy in biology. A wide range of synthetic heterocyclic molecules bind noncovalently in the minor groove of the double helix and are also effective against a number of human and animal diseases. A classic structural concept, the isohelicity principle, has guided much of this work: such heterocyclic molecules require a shape that complements the convex surface of the minor groove. Researchers have used this principle to design molecules that can read DNA sequences. This principle also predicts that molecules that lack the complementary shape requirement would only bind weakly to DNA. Recently, however, researchers have unexpectedly found that some essentially linear compounds, which do not have this feature, can have high DNA affinity.

In this Account, we discuss an alternative recognition concept based on these new findings. We demonstrate that highly structured water molecules can play a key role in mediating between the ligand and DNA minor groove without loss of binding affinity. Combined structural and thermodynamic approaches to understanding the behavior of these molecules have shown that there are different categories of bound water in their DNA complexes. For example, application of this water-bridging concept to the phenylamidine platform has resulted in the discovery of molecules with high levels of biological activity and low nonspecific toxicity. Some of these molecules are now in advanced clinical trials.

### Introduction

Designed synthetic dicationic compounds that specifically recognize the DNA minor groove are providing clinically useful therapeutic agents against a number of diseases and serving as well-defined model systems giving insight into the structural and thermodynamic features that characterize DNA molecular recognition.<sup>1–4</sup> The interactions of dicationic organic molecules with the minor groove are characterized by both general

and specific factors. The former include electrostatics, van der Waals contacts, hydrophobicity, and groove hydration. Specific factors in particular sequences include H-bonding groups, steric differences, and water/ion arrangements. The minor groove contains primarily hydrogen-bond acceptor groups, the purine N3 and pyrimidine O2 at the floor of the groove as well as the deoxyribose O4' at the groove walls (Figure 1).<sup>3,5–7</sup> The 2-NH<sub>2</sub> of G makes the third H-bond in G·C base pairs



**FIGURE 1.** Schematic of the Watson–Crick base pairs showing hydrogen-bond donors and acceptors on the bases in the minor (m) and major (M) grooves.

and can prevent deep penetration into the minor groove. Structured solvent and increased electrostatic potential in A/T sequences also distinguish A/T from G/C minor groove binding sites.<sup>3,5–10</sup> Thus, for compounds that bind in the minor groove, shape that complements the groove curvature (“isohelicity”) and functional group placement, charge, and H-bonding groups have been identified as critical components for binding affinity and recognition.<sup>2,3,9–16</sup>

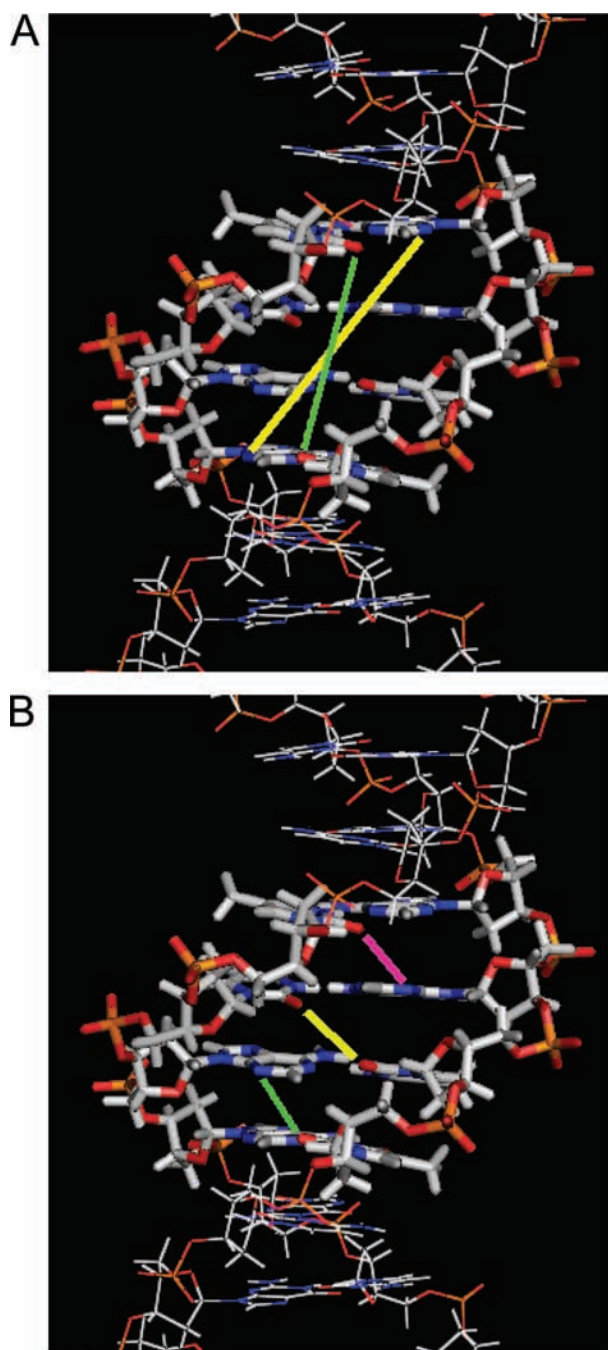
Recent studies on a range of heterocyclic dication, particularly diamidines and analogs, have significantly extended and modified our understanding of how such compounds can interact with the minor groove, and how ligand curvature, dihedral twist, charge, and in particular interactions with water molecules in the DNA complex can lead to highly specific and very strong interactions with the groove. Crystallographic and thermodynamic studies have provided complementary insights into the details of minor-groove binding. Diamidines are currently the primary class of reversible DNA minor groove binding compounds with therapeutic utility, and we survey here their DNA interactions, comparing and contrasting structural features for classical recognition of the minor groove with new models for nonclassical interactions. In particular, the fundamental molecular recognition features among compound

shape, water, and base edges will be described and correlated with DNA–ligand binding thermodynamics.

## General features of DNA–minor groove complexes at the -GAATTC- Site

The first B-form DNA crystal structure to be solved was of the DNA duplex,  $d(\text{CGCGAATTCGCG})_2$ , and this sequence has subsequently been studied in detail by both solution and crystallographic methods.<sup>3,6,7,17–22</sup> It contains a core of four A–T base pairs and has been frequently used as a platform for studies of specific DNA minor groove interactions. Figure 2 shows a solution structure of this duplex (“A2T2”) and some critical distances between H-bond groups on base pair edges at the floor of the groove. The right-hand twist of B-DNA creates unequal interstrand acceptor–acceptor distances for a given base pair separation (Figure 2A). For the AATT site, for example, the O2 of T8 to O2 of T20 is 9.1 Å, while the distance between N3 of the complementary bases A17 and A5 is 12.6 Å (the base numbering scheme on the duplex strands is given in Figure 3A). These are linear distances so that in order to form direct H-bonds with the two DNA groups, a bound ligand must have a concave curvature to complement the convex shape subtended by successive base edges along the floor of the groove. The distance difference for H-bond acceptors is a critically important feature for minor groove binding. Shorter ligands that just cover the four base pairs in the AATT sequence will opt for H-bonding with the O2 of T8 and O2 of T20 while longer, crescent-shaped, ligands prefer the N3s of the complementary A17 and A5. These distances change with sequence, and this can lead to specificity in recognition of different AT sequences with appropriately designed compounds.

Figure 2B illustrates the three shortest (3.7 Å) interstrand H-acceptor distances on adjacent bases (O2 of T8 to N3 of A18, O2 of T7 to O2 of T19, and N3 of A6 to O2 of T20) at the AATT site. These atoms frequently interact with ligand indole or benzimidazole -NH- groups to form bifurcated H-bonds with the central region of the AATT sequence (Figure 3A). The netropsin–AATT structure, (Figure 3B), provided the basis for the classical minor groove recognition model that emphasizes appropriate ligand curvature to match the groove, ligand, and DNA complementary H-bonding groups and displacement of water/ions from the -AATT- groove on complex formation.<sup>23,24</sup> Based on the shape-complementarity concept, the classical model has been used to predict that ligands with incorrect curvature would bind poorly to DNA. The netropsin amide -NH- groups arranged in a crescent fashion, take advantage of the shortest interstrand H-bond acceptor distances in



**FIGURE 2.** Interstrand distances of H-bond acceptor groups at the AATT site in  $d(\text{CGCGAATTCGCG})_2$ : (A) Two sets of H-bond acceptors are within a four-base pair length. The N3 of A5 to N3 of A17 distance (yellow) is longer than the O2 of T8 to O2 of T20 distance (green), and it is a through-base distance so that a bound compound must have some curvature in order to form H-bonds with the two Ns. (B) The shortest interstrand acceptor–acceptor distances (3.7 Å) within the minor groove of -AATT- are illustrated. The three lines indicate O2 of T8 to N3 of A18 (magenta), O2 of T7 to O2 of T19 (yellow), and N3 of A6 to O2 of T20 (green). Compound -NH<sup>+</sup> groups (from amidine, amide, benzimidazole, or indole) can form bifurcated H-bonds with these three pairs of H-bond acceptors.

the minor groove by forming an array of bifurcated H-bonds with N3 of A6 and O2 of T20, O2 of T7 and O2 of T19, and O2 of T8 and N3 of A18 (Figure 3B). In contrast to many diamidine compounds (discussed below), isothermal titration calorimetry (ITC) studies indicated that the interaction between netropsin and AATT is enthalpically driven with a free energy of  $-12$  kcal/mol, presumably due in part to the large number of H-bonds formed in this complex.<sup>25</sup>

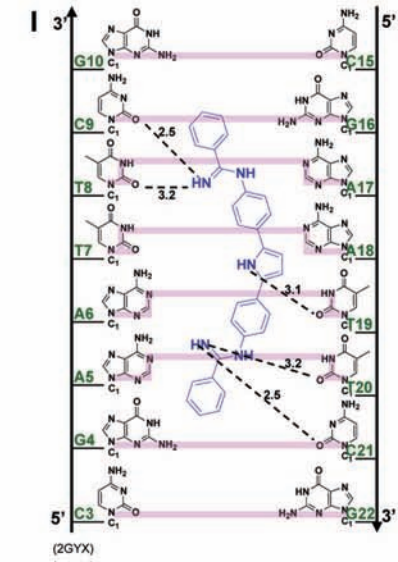
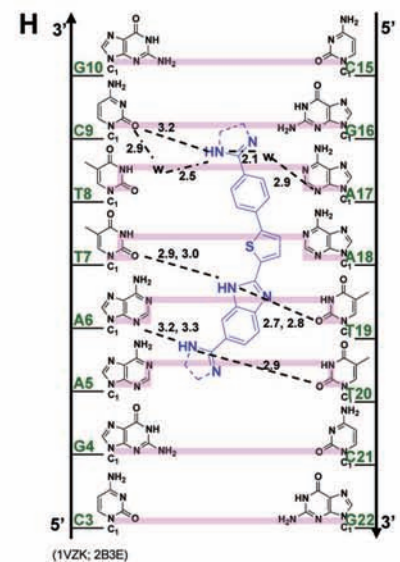
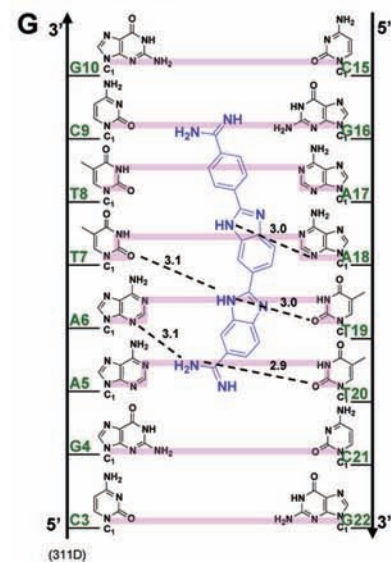
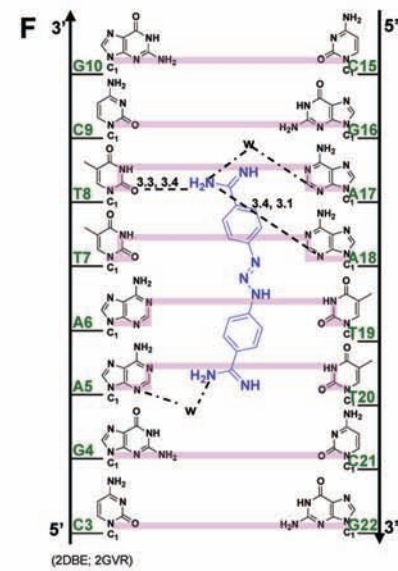
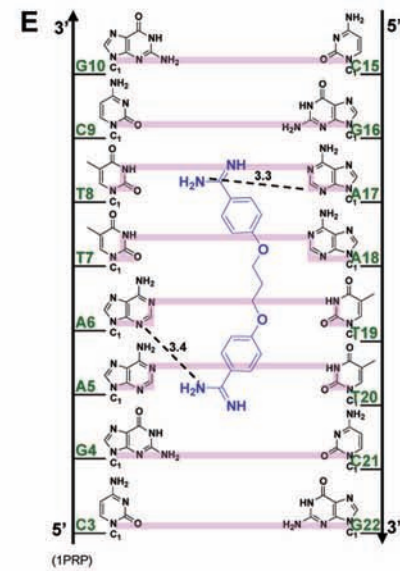
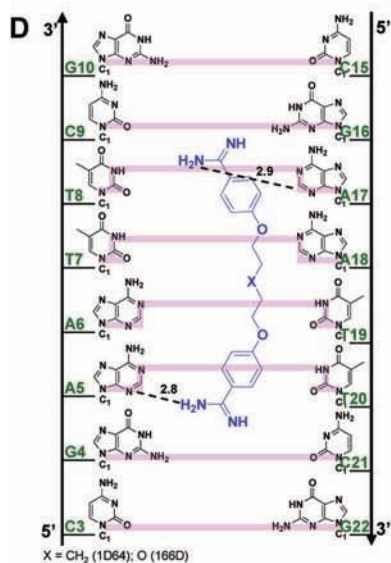
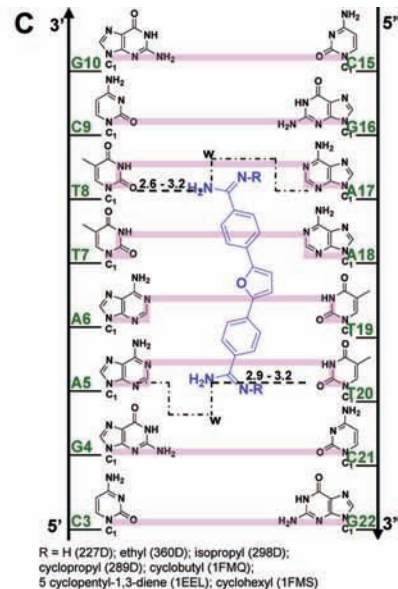
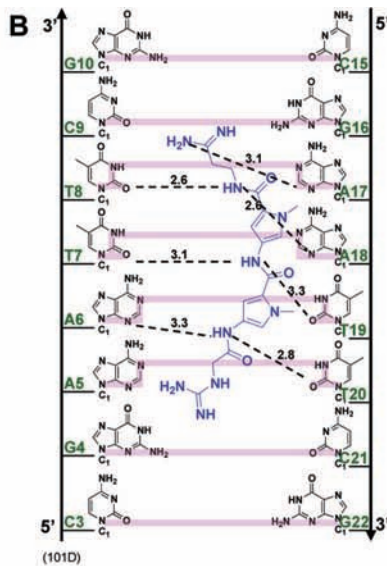
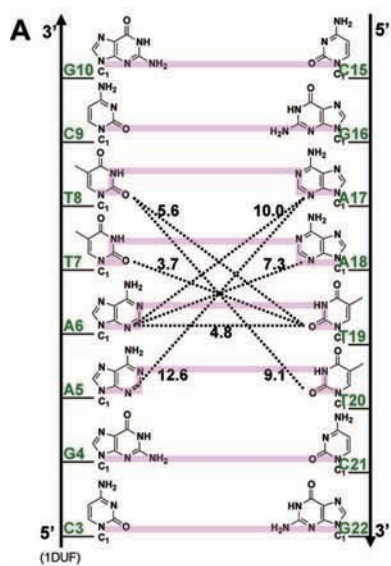
Figure 3 illustrates specific interactions of all synthetic dicationic amidine analogs that have been crystallized with AATT. For assistance in correlating the compound names, structures, and PDB identifications, a reference list is shown in Table 1. The diamidines and analogs in Figures 3C–I closely match the curvature and H-bonding groups of AT sequences in the DNA minor groove, an isohelical interaction, as well as complement the negative charge of the DNA phosphates. This is the “classical type” of minor groove complex that provides AT base pair sequence recognition. A different and unexpected type of complex, based on the classical model, however, is observed for the compounds bound to AATT in Figures J–L. Comparison of these two distinct complex classes shows that the requirements for minor groove interactions to occur can result in the shape of the compounds deviating significantly from an isohelical match to the groove.

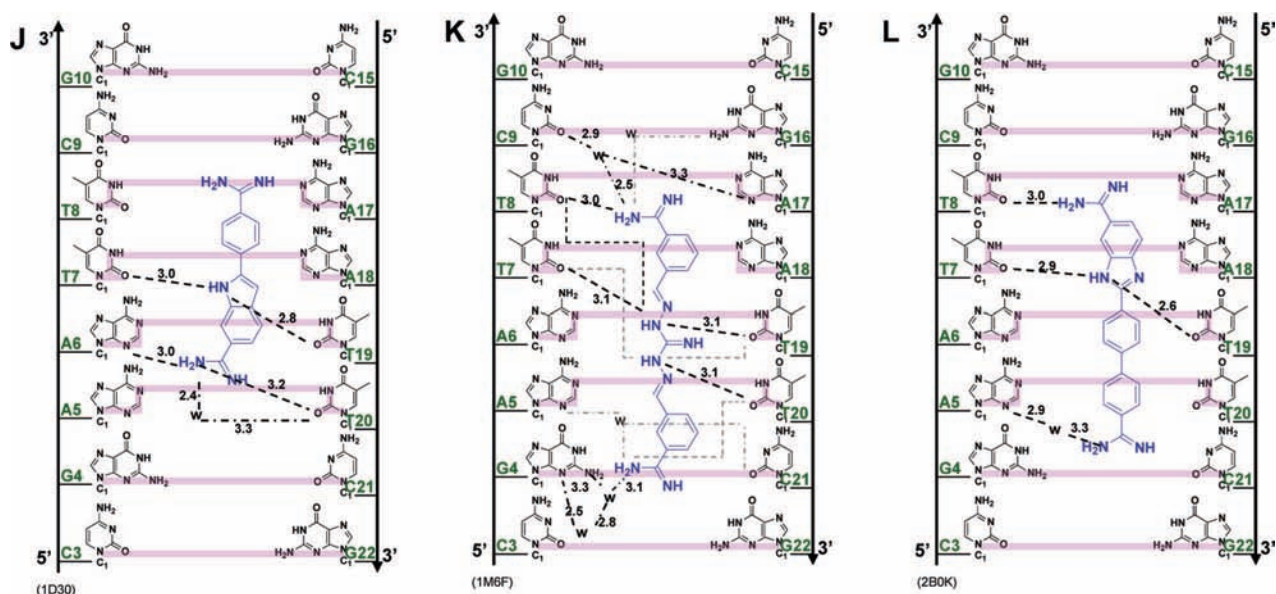
### Structures and Thermodynamics of Diamidine Complexes with the A/T Minor Groove

In many of the DNA complexes in Figure , the amidine or modified-amidine cationic group is linked to the groove floor-base edge, groove wall-O4', or both by bridging water molecules that flank the ends of the bound molecule. While interactions with the groove walls and phosphate groups can enhance overall complex stability, they generally have more limited effects on sequence specificity. We will therefore emphasize the interactions among the ligand, the base edges, and bridging water molecules at the floor of the minor groove where sequence specificity is primarily defined.

Furamidine, DB75 ( $R = \text{H}$  in Figure 3C), is a therapeutically useful heterocyclic dication that has been crystallized as an AATT complex (PDB 227D). A neutral, orally available prodrug of the compound has reached phase III clinical trials against human African trypanosomiasis.<sup>26</sup> The compound selectively targets repeated A/T sequences in the trypanosome mitochondrial kinetoplast DNA and has exhibited no significant toxicity to patients at therapeutic levels.<sup>1,2</sup> The primary H-bond contacts from DB75 and derivatives to DNA are consistently between the amidines and O2s of T8 and T20.<sup>27–31</sup>







**FIGURE 3.** (A) Crystal structures (with PDB IDs) of the dodecamer duplex  $d(CGCGAATTCGCG)_2$  with some important interstrand H-bond acceptor distances and with minor groove binding diamidines and analogs (B - L). The light magenta lines depict the base edges that form the floor of the groove at which specific interactions occur and the dashed lines indicate compound-DNA H-bonds. Water molecules that connect the bound compound to the minor groove floor are indicated with “w”. In (K), lighter gray lines indicate the interaction of another equivalent structure.

**TABLE 1**

ligand	figure in which structure is shown	PDB IDs
netropsin	3B	101D
furamide and closely related ligand	3C	227D, 360D, 298D, 289D, 1FMQ, 1EEL, 1FMS
pentamidine and derivative	3D	1D64, 166D
propamidine	3E	1PRP
berenil	3F	2DBE, 2GVR
DB185	3G	311D
DB818 and DB819	3H	1VZK, 2B3E
DB884	3I	2GYX
DAPI	3J	1D30
CGP 40215A	3K	1M6F
DB921	3L	2B0K

In addition, water molecules are observed to bridge the amidine  $-NH_2$  and N3 of A5 (PDB 1EEL and 289D) or both N3 of A5 and N3 of A17 (298D, 360D, and 1FMQ). Under physiological conditions, DB75 and close analogs have a binding constant,  $K_a$ , with the AATT sequence in the range of  $(1-5) \times 10^7 M^{-1}$  with a small  $\Delta H$  and large  $T \cdot \Delta S$  for binding.<sup>31,32</sup> Shape recognition and van der Waals interactions thus appear to play a larger role in specific recognition of -AATT- by these compounds than with netropsin. This set of diamidines (Figure 3C) thus provides a well-defined structural and thermodynamic benchmark for comparison with other minor groove binders. It should be noted that the thermodynamic values reported in this paper are at 25 °C, and since these complexes have negative heat capacity change values for binding, the  $\Delta H$

and  $T \cdot \Delta S$  will become more negative or less positive, depending on the initial values.

Unlike the relatively rigid diphenylfuran core of DB75 and analogs (Figure 3C), pentamidine<sup>33</sup> and derivatives<sup>34</sup> (Figure 3D) have flexible linkers connecting two benzamidines. Pentamidine also binds to the four base pairs of AATT (PDB 1D64), but it utilizes the N3s of A5 and A17 as amidine H-bond acceptors rather than O2 of the complementary T20 and T8 as with furamide. Since the separation of the two inner-facing amidine  $-NH_2$  groups in bound pentamidine (15.3 Å) is significantly longer than in DB75 (12.8 Å), pentamidine interacts with the longer interstrand H-bond acceptor pair within AATT (H-bond distances shown in Figure 3D). The shorter diphenylfuran derivatives select the closer O2 of T8 to O2 of T20 for H-bonding (Figure 3C).

Propamidine (Figure 3E) is a shorter version of pentamidine, but the amidine N–N distance (14.6 Å) in its AATT complex (PDB 1PRP) is longer than that of furamide. While one of the amidines of propamidine forms a direct H-bond with N3 of A17 (as with pentamidine), the other amidine cannot reach the N3 of A5 of the opposite strand, and it is closer to N3 of A6 (3.4 Å). The propamidine molecule is longer than furamide but interacts with N3 of A17 and N3 of A6 because it is more acutely concave than furamide. This curvature helps it wrap around the base edges for the N3 distances noted above. Despite its relatively long H-bonds (to N3 of A6), propamidine binds more strongly than pentamidine,<sup>35</sup> probably



because of a smaller conformational entropy loss when the trimethylene linker fits into the minor groove relative to the pentamethylene of pentamidine. Both pentamidine and propamidine bind more weakly than furamidine, and ITC studies indicate that their interactions are entropically driven.

Berenil (diminazene) (Figure 3F)<sup>36</sup> resembles furamidine and is an important antitrypanosomal drug in animals. In its AATT complex, berenil (PDB 2DBE, 2GVR) has an amidine N–N distance (13.3 Å) that is slightly longer than that of furamidine. However, it only binds across the -ATT- site with a weak (3.4 Å) H-bond from one amidine to O2 of T8. The other amidine interacts with N3 of A5 through a bridging water molecule in the groove. The longer propamidine also covers three base pairs (-ATT-), but no water molecule is required to link it to the floor of the groove. A crystal structure with a longer AAATTT site<sup>37</sup> shows that berenil binds to the central -AATT- by forming H-bonds with interstrand O2s of T in much the same way that DB75 binds to DNA. Berenil binds with a slightly lower affinity than furamidine ( $1 \times 10^7$  and  $2 \times 10^7$  M<sup>-1</sup>, respectively)<sup>38</sup> and with a small  $\Delta H$  and large  $T \cdot \Delta S$  of binding. The furamidine, pentamidine, and berenil group of compounds thus illustrate the flexibility of amidine–DNA H-bonding interactions that are possible for ligands of similar size within an expanded classical model of minor groove complex formation. The compounds, and to a lesser extent DNA, have sufficient conformational flexibility to allow these differences in AATT recognition to optimize the Gibbs energy of binding in each case.

### Compounds with Additional Types of H-Bond Interactions

DB185 (Figure 3G) is distinct from furamidine but similar in structure to Hoechst 33258, another well-established minor groove binder. DB185 binds to the minor groove of AATT much more strongly than furamidine or Hoechst 33258.<sup>39</sup> It is one of a small number of DNA minor groove binders with a high affinity ( $K_a > 10^{11}$  M<sup>-1</sup>). The strong binding results in part from the H-bond network between the ligand and DNA, van der Waals interactions, and extensive water release from the minor groove and ligand.<sup>39–41</sup> A bifurcated H-bond is formed between the terminal benzimidazole -NH- and the interstrand O2 of T7 and T19 as well as between the amidine attached to this benzimidazole and N3 of A6 and O2 of T20 (Figure 3G). Extensive water bridges are also observed at the amidine attached to benzimidazole with water molecules linked to O2 of C21 and O4' of G22 on the backbone.

DB818 (amidine) and DB819 (imidazoline) have the furan of DB75 converted to a thiophene and a phenyl replaced by

a benzimidazole (Figure 3H).<sup>42,43</sup> The larger size of the sulfur atom (compared with oxygen in furan) and resulting bond angle changes become magnified at the molecular extremities in terms of overall ligand shape. This leads to a lower curvature and a superior fit within the minor groove with a binding affinity that is over 30 times greater for DB818 than for furamidine.<sup>43</sup> This change is due to the thiophene substitution since the benzimidazole derivative of DB75 has similar binding affinity to DB75.<sup>44,45</sup> With DB818 (Figure 3H, PDB 1VZK), there are direct amidine H-bonds to O2 of C9 and bifurcated H-bonds to O2 of T20 and N3 of A6. There are also two water molecules mediating the interactions between one of the imidazoline ends and O2 of C9 and N3 of A17.

DB884 was the first reversed-amidine (rev-Am) to be crystallized with DNA (Figure 3I).<sup>46</sup> In its AATT complex, one rev-Am -NH- forms a bifurcated H-bond with intrastrand O2 of T8 and O2 of C9, while the other reversed-amidine interacts with the opposite strand through a bifurcated H-bond with intrastrand O2 of T20 and O2 of C21 (Figure 3I). An important feature in this complex that differs from furamidine is an H-bond between the pyrrole -NH- and O2 of T19. Although reversed-amidines generally bind more weakly than the corresponding amidines, DB884 is an exception with a  $K_a$  similar to that of furamidine.<sup>46</sup> Replacement of the pyrrole NH with a NCH<sub>3</sub> or with O (furan) leads to a destabilization of the complex indicating the importance of the pyrrole–DNA H-bond.<sup>46</sup> The extra H-bond contributes to a larger  $\Delta H$  on binding of DB884 relative to DB75.<sup>46</sup>

The phenyl-indole diamidine DAPI has been used for many years as a DNA stain. Although DAPI is small, it binds to AATT with a higher affinity than DB75.<sup>25,31</sup> The indole -NH- forms a bifurcated H bond with the interstrand O2s of T7 and T19 (Figure 3J).<sup>47</sup> The amidine attached to this indole forms a bifurcated H-bond with N3 of A6 and O2 of T20 while a flanking water molecule also bridges the amidine with O2 of T20. The other amidine is above the floor of the minor groove, but no water molecule is observed in the relatively low-resolution DAPI structure. As with many diamidine compounds, ITC shows that the binding enthalpy contributes less than half of the binding free energy.<sup>25</sup>

### “Nonclassical” Minor Groove Complexes

The minor groove ligands described above obey the requirement for isohelicity to complement the DNA minor groove. Recent results, however, have shown that the “classical model”, which is described above and which originated when the database of structures was small, is only one way to obtain high-affinity and specific DNA minor groove recognition. The

AATT complex of CGP 40215A (Figure 3K) differs significantly from the classical model (Figures 3C–I)<sup>48</sup> because of the linear shape of the compound. In contrast to pentamidine (of similar size), only one of the amidines of CGP 40215A can form direct H-bond contacts with AATT. While this would be expected to significantly reduce the binding affinity, CGP 40215A actually binds more strongly than the curved pentamidine (40–50 times more strongly) or furamidine (5 times more strongly).

In the DNA complex (PDB 1M6F), one amidine of CGP 40215A forms H-bonds with O2 of T8 along with indirect H-bonds, through a bridging water molecule, with N3 of the complementary A17 and O2 of C9. An -NH- on the linker forms a bifurcated H-bond with the interstrand O2 of T7 and T19, the shortest interstrand H-bond acceptor distance in the minor groove (Figure 3A). The amidine group at the other end of CGP 40215A cannot form a direct H-bond with the base edges because it protrudes away from the groove floor, but instead forms an H-bond bridge with two water molecules to N2 and N3 of G4. A molecular dynamics simulation suggests that due to the symmetry of the ligand and the AATT DNA duplex, an equivalent structure exists with a similar pattern of H-bonds. This compound thus shows a distinct mode of groove binding in which water molecules can compensate for the lack of curvature of the ligand and can complete the crescent shape requirement for minor groove recognition and strong binding.<sup>48</sup> Thermodynamic studies indicate that the binding enthalpy contributes slightly more than entropy to the binding free energy.<sup>49</sup>

To provide additional insight into and tests of the curvature requirements for minor groove binding, DB921 and other relatively linear diamidines have been designed and synthesized. DB921 is an adaptation of DAPI that lengthens the molecule without introducing additional curvature. In the DB921 complex (PDB 2B0K), the benzimidazole group is in the groove, while the biphenyl-amidine projects to the top edge of the minor groove. The benzimidazole -NH- forms a bifurcated H-bond with the interstrand O2s of T7 and T19, and the amidine on the benzimidazole forms an H-bond with the O2 of T8. This structure creates an optimum pocket for a direct bridging water molecule between the phenyl-amidine and N3 of A5 (Figure 3L),<sup>50</sup> which forms a link of flexible orientation and H-bonding ability between the amidine and the N3 H-bond acceptor on the groove floor and is sandwiched between the base and the amidine. This water interaction differs from that of berenil in which a terminal water molecule is linked to an acceptor atom of the next base along the sequence. Surprisingly, the binding constant for DB921 is

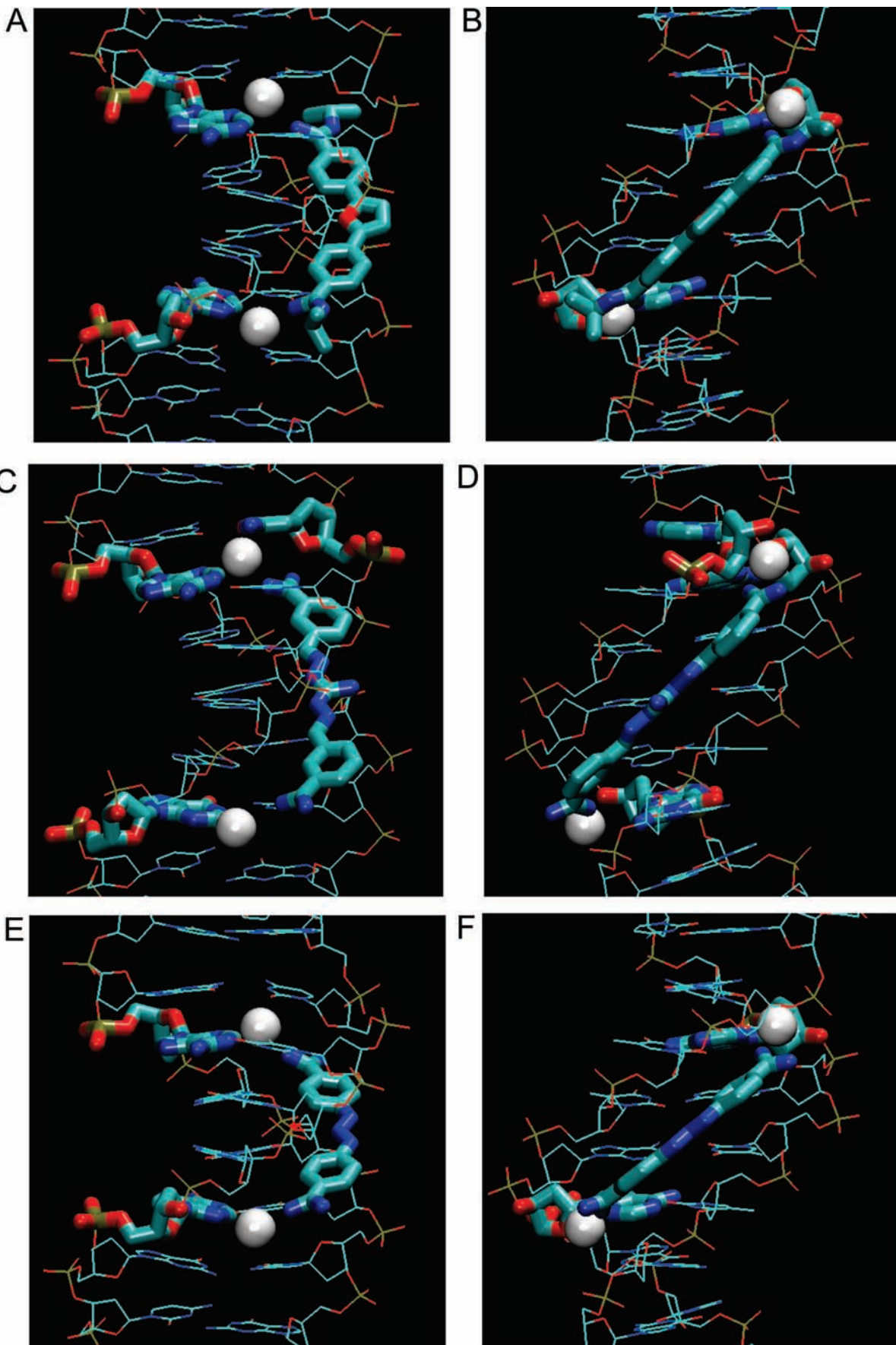
more than 25 times higher than that of furamidine, and ITC results reveal that the interaction with AATT is entropically driven.<sup>50</sup> This demonstrates that a concave shape is not a requirement for a ligand to have strong and specific interactions with the minor groove.

## Water-Mediated H-Bonds in Diamidine Complexes

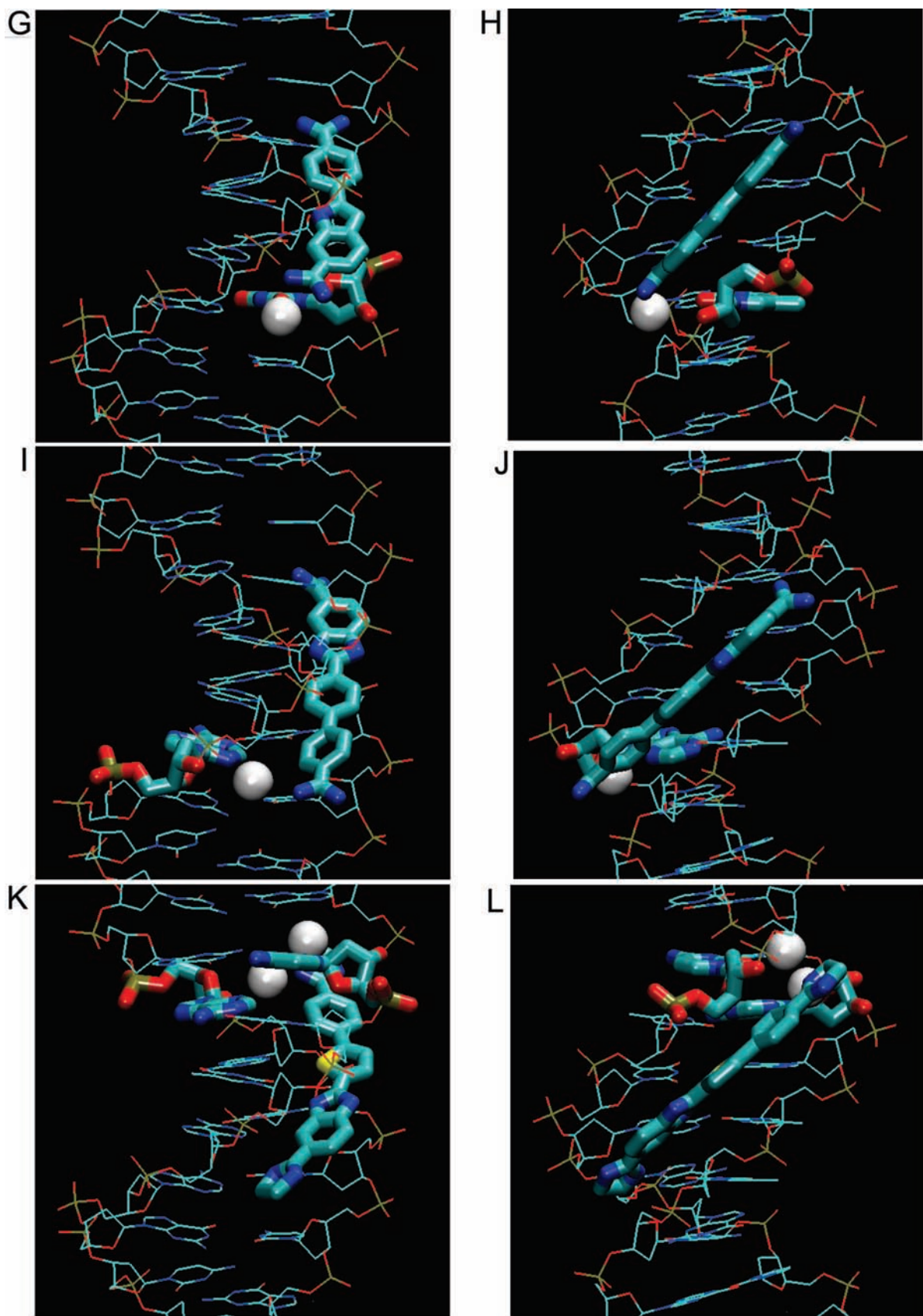
Figure 4 illustrates examples of hydrated amidine groups (from Figure 3) within the minor groove, showing different types of ligand–DNA–water bridges. Only water molecules that are within 3.3 Å between ligand and base edge(s) are shown. Figures 4A, 4B show different views of a furamidine derivative (PDB 298D) that forms two water bridges with N3s of A17 and A5. Figures 4C, 4D illustrate water bridges of the linear CGP 40215A with -NH<sub>2</sub> of G4 (lower) or with O2 of C9 and N3 of A17. Figures 4E, 4F show berenil with water bridges to N3 of A5 or A17. DAPI forms a water bridge at its indole end with O2 of T20 (Figures 4G, 4H). In a water-dependent interaction that is very different from the classical model, the phenylamidine of DB921 and the first A base of AATT sandwich a buried water molecule that is an essential and intrinsic part of the complex (Figures 4I, 4J). The imidazoline–thiophene derivative DB819 (Figures 4K, 4L) has two types of water interactions: one water molecule is between the imidazoline and N3 of A17, while the other reaches to the O2 of C9 on the opposite strand. Clearly, water is an important component of most diamidine minor groove complexes, and in cases such as DB921, it is an essential part of the complex structure.

## New and Extended Design Features for Biologically Active Diamidines

Binding to sites, such as the A/T-rich kinetoplast of trypanosomes and leishmania or the A/T-rich nuclear DNA of the malaria parasite, provide excellent targets for biological activity with little host toxicity.<sup>2</sup> This mirrors the “AT island” concept proposed for minor groove binding anticancer drugs, and it may be possible that appropriately designed diamidines can also be developed as anticancer agents.<sup>51</sup> For effective cell uptake, the compound size should be kept small, and synergistic biological effects that can take place at multiple A/T sequences, such as those in kinetoplast DNA, offer attractive biological targets. The inclusion of water interactions and ligand flexibility into the design of new ligands is a challenging problem, and additional compounds and DNA interactions studies will be required to establish the limits of this new approach to achieve satisfactory binding affinity, sequence







**FIGURE 4.** Types of amidinium–water–base interactions: bridging waters in furamidine (A, B), CGP40215A (C, D), berenil (E, F), and DAPI (G, H), sandwiching water of DB921 (I, J), and a mix of both types with DB819 (K, L). Water molecules are shown as white spheres, ligands and interacting nucleotides as tubes, and other nucleotides as stick representations.

selectivity, site recognition, and biological activity. Results with DB921 (Figures 4I, 4J) illustrate the utility of water as a flexible adapter for DNA complexes and provide optimism for this design approach.

More drastic differences in diamidine–DNA interactions are observed with DB293 and DB1242, which form cooperative stacked dimers to recognize more complex sequences with G–C base pairs than is possible for the compounds in Figure 3.<sup>44,52</sup> A single molecule of DB1242, for example, can extend over 3–4 base pairs but by stacking of two molecules, a five base-pair sequence including four G–C base pairs can be targeted.<sup>52</sup> Positively charged amidine, reversed-amidine, or guanidine groups (Figure 3) are favorable for hydration, solubility, stacking with the minor groove, H-bonding and electrostatic interactions with the anionic field of the minor groove. Benzamidine motifs are not only important for DNA binding because of charge complementarity to the DNA and as H-bond donors but also for cellular uptake.<sup>53–55</sup> Amidine derivatives, thus, continue to offer exciting opportunities for development of drugs that target DNA. Polyamides that target the minor groove have excellent sequence recognition capability<sup>9,10,12,14,56</sup> and many useful applications in fundamental molecular recognition. Unlike the amidines, however, the polyamides have poor cell uptake and limited stability with no potential for oral drug administration. They have, therefore, not been used in any clinical application. Ongoing research with new amidine derivatives, such as DB293 and DB1242, which recognize mixed DNA sequences, may provide a bridge between the therapeutic and sequence design areas.

*Long-term collaborative efforts on diamidine–DNA complexes with David W. Boykin and co-workers (Georgia State University, Atlanta, GA), Christian Bailly (Institut de Recherche Pierre Fabre, Toulouse, France), J. Ed Hall, and Richard R. Tidwell (University of North Carolina, Chapel Hill, NC) have been both scientifically productive and personally enjoyable. Almost all of the diamidines in this Account were synthesized by Professor Boykin and co-workers and this project would not have been possible without their efforts, suggestions, and collaboration. Support from the National Institutes of Health, the Bill and Melinda Gates Foundation, the Georgia Research Alliance (to W.D.W.), and CRUK (to S.N.) is gratefully acknowledged.*

#### BIOGRAPHICAL INFORMATION

**Binh Nguyen** received his Ph.D. in biophysical chemistry (2002) from Georgia State University, Atlanta, GA. He conducted post-doctoral research at Georgia State University and is currently at Washington University in St. Louis.

**Stephen Neidle** has a Ph.D. in chemical crystallography from Imperial College, London (1970), and a D.Sc. from the University of London. He is currently Professor of Chemical Biology at the School of Pharmacy, University of London.

**W. David Wilson** obtained his Ph.D. in biophysical chemistry in 1970 at Purdue University. He began work on peptide–DNA interactions as well as anticancer drug design with Dr. Edmond Gabbay at the University of Florida in Gainesville, Florida. He is currently Regents' Professor of Chemistry at Georgia State University. All three authors have a strong interest in DNA structure and interactions, in the design of compounds to specifically recognize DNA, and in the targeting of DNA sequences and structures for therapeutic uses.

#### FOOTNOTES

\*To whom correspondence should be addressed. E-mail addresses: wdw@gsu.edu; stephen.neidle@pharmacy.ac.uk.

#### REFERENCES

- Tidwell, R. R.; Boykin, D. W. In *DNA and RNA Binders: From Small Molecules to Drugs*; Demeunynck, M., Bailly, C., Wilson, W. D., Eds.; WILEY-VCH: Weinheim, Germany, 2003; Vol. 2; pp 414–460.
- Wilson, W. D.; Nguyen, B.; Tanius, F. A.; Mathis, A.; Hall, J. E.; Stephens, C. E.; Boykin, D. W. Dications that target the DNA minor groove: compound design and preparation, DNA interactions, cellular distribution and biological activity. *Curr. Med. Chem. Anticancer Agents* **2005**, *5*, 389–408.
- Neidle, S. DNA minor-groove recognition by small molecules. *Nat. Prod. Rep.* **2001**, *18*, 291–309.
- Mathis, A. M.; Bridges, A. S.; Ismail, M. A.; Kumar, A.; Francesconi, I.; Anbazhagan, M.; Hu, Q.; Tanius, F. A.; Wenzler, T.; Sautler, J.; Wilson, W. D.; Brun, R.; Boykin, D. W.; Tidwell, R. R.; Hall, J. E. Diphenyl furans and aza analogs: Effects of structural modification on in vitro activity, DNA binding, and accumulation and distribution in trypanosomes. *Antimicrob. Agents Chemother.* **2007**, *51*, 2801–2810.
- Seeman, N. C.; Rosenberg, J. M.; Rich, A. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 804–808.
- Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Crystal structure analysis of a complete turn of B-DNA. *Nature* **1980**, *287*, 755–758.
- Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry* **1998**, *37*, 8341–8355.
- Lavery, R.; Pullman, B.; Zakrzewska, K. Intrinsic electrostatic properties and base sequence effects in the structure of oligonucleotides. *Biophys. Chem.* **1982**, *15*, 343–351.
- Dervan, P. B.; Burli, R. W. Sequence-specific DNA recognition by polyamides. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688–693.
- Reddy, B. S.; Sondhi, S. M.; Lown, J. W. Synthetic DNA minor groove-binding drugs. *Pharmacol. Ther.* **1999**, *84*, 1–111.
- Goodsell, D.; Dickerson, R. E. Isohelical analysis of DNA groove-binding drugs. *J. Med. Chem.* **1986**, *29*, 727–733.
- Dervan, P. B.; Edelson, B. S. Recognition of the DNA minor groove by pyrrole–imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284–299.
- Zasedatelev, A. S. Geometrical correlations useful for design of sequence-specific DNA narrow groove binding ligands. *FEBS Lett.* **1991**, *281*, 209–211.
- Wemmer, D. E. Designed sequence-specific minor groove ligands. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 439–461.
- Nikitin, A. M.; Gursky, G. V. New structural motifs isohelical to DNA. *Dokl. Biochem. Biophys.* **2003**, *390*, 139–142.
- Nikitin, A. M.; Rodin, S. A.; Pis'menskii, V. F.; Surovaya, A. N.; Gursky, G. V. A new pseudopeptide motif for designing specific DNA-binding compounds capable of recognizing long DNA sequences. *Dokl. Biochem. Biophys.* **2002**, *384*, 167–171.
- Drew, H. R. Geometry of hydration. *J. Mol. Biol.* **1981**, *151*, 535–556.
- Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179–2183.
- Tereshko, V.; Minasov, G.; Egli, M. A “hydrat-ion” spine in a B-DNA minor groove. *J. Am. Chem. Soc.* **1999**, *121*, 3590–3595.

- 20 Tjandra, N.; Tate, S.-i.; Ono, A.; Kainosho, M.; Bax, A. The NMR structure of a DNA dodecamer in an aqueous dilute liquid crystalline phase. *J. Am. Chem. Soc.* **2000**, *122*, 6190–6200.
- 21 Woods, K. K.; McFail-Ison, L.; Sines, C. C.; Howerton, S. B.; Stephens, R. K.; Williams, L. D. Monovalent cations sequester within the A-tract minor groove of  $[d(CGCGAATTCGCG)]_2$ . *J. Am. Chem. Soc.* **2000**, *122*, 1546–1547.
- 22 Johansson, E.; Parkinson, G.; Neidle, S. A new crystal form for the dodecamer C-G-C-G-A-A-T-T-C-G-C-G: Symmetry effects on sequence-dependent DNA structure. *J. Mol. Biol.* **2000**, *300*, 551–561.
- 23 Goodsell, D. S.; Kopka, M. L.; Dickerson, R. E. Refinement of netropsin bound to DNA: Bias and feedback in electron density map interpretation. *Biochemistry* **1995**, *34*, 4983–4993.
- 24 Patel, D. J. Antibiotic-DNA interactions: Intermolecular nuclear Overhauser effects in the netropsin-d(C-G-C-G-A-A-T-T-C-G-C-G) complex in solution. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 6424–6428.
- 25 Freyer, M. W.; Buscaglia, R.; Nguyen, B.; Wilson, W. D.; Lewis, E. A. Binding of netropsin and 4,6-diamidino-2-phenylindole to an A2T2 DNA hairpin: A comparison of biophysical techniques. *Anal. Biochem.* **2006**, *355*, 259–266.
- 26 Midgley, I.; Fitzpatrick, K.; Taylor, L. M.; Houchen, T. L.; Henderson, S. J.; Wright, S. J.; Cybulski, Z. R.; John, B. A.; McBurney, A.; Boykin, D. W.; Trendler, K. L. Pharmacokinetics and metabolism of the prodrug DB289 (2,5-bis[4-(N-methoxyamidino)phenyl]furan monomaleate) in rat and monkey and its conversion to the antiprotozoal/antifungal drug DB75 (2,5-bis[4-(N-guanylphenyl)furan dihydrochloride]). *Drug Metab. Dispos.* **2007**, *35*, 955–967.
- 27 Laughton, C. A.; Tanious, F.; Nunn, C. M.; Boykin, D. W.; Wilson, W. D.; Neidle, S. A crystallographic and spectroscopic study of the complex between  $d(CGCGAATTCGCG)_2$  and 2,5-bis[4-(N-guanylphenyl)furan], an analogue of berenil. Structural origins of enhanced DNA-binding affinity. *Biochemistry* **1996**, *35*, 5655–5661.
- 28 Guerri, A.; Simpson, I. J.; Neidle, S. Visualisation of extensive water ribbons and networks in a DNA minor-groove drug complex. *Nucleic Acids Res.* **1998**, *26*, 2873–2878.
- 29 Trent, J. O.; Clark, G. R.; Kumar, A.; Wilson, W. D.; Boykin, D. W.; Hall, J. E.; Tidwell, R. R.; Blagburn, B. L.; Neidle, S. Targeting the minor groove of DNA: Crystal structures of two complexes between furan derivatives of berenil and the DNA dodecamer  $d(CGCGAATTCGCG)_2$ . *J. Med. Chem.* **1996**, *39*, 4554–4562.
- 30 Simpson, I. J.; Lee, M.; Kumar, A.; Boykin, D. W.; Neidle, S. DNA minor groove interactions and the biological activity of 2,5-bis-[4-(N-alkylamidino)phenyl] furans. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2593–2597.
- 31 Mazur, S.; Tanious, F. A.; Ding, D.; Kumar, A.; Boykin, D. W.; Simpson, I. J.; Neidle, S.; Wilson, W. D. A thermodynamic and structural analysis of DNA minor-groove complex formation. *J. Mol. Biol.* **2000**, *300*, 321–337.
- 32 Liu, Y.; Kumar, A.; Boykin, D. W.; Wilson, W. D. Sequence and length dependent thermodynamic differences in heterocyclic diamidine interactions at AT base pairs in the DNA minor groove. *Biophys. Chem.* **2007**, *131*, 1–14.
- 33 Edwards, K. J.; Jenkins, T. C.; Neidle, S. Crystal structure of a pentamidine-oligonucleotide complex: implications for DNA-binding properties. *Biochemistry* **1992**, *31*, 7104–7109.
- 34 Nunn, C. M.; Jenkins, T. C.; Neidle, S. Crystal structure of gamma-oxapentamidine complexed with  $d(CGCGAATTCGCG)_2$ . The effects of drug structural change on DNA minor-groove recognition. *Eur. J. Biochem.* **1994**, *226*, 953–961.
- 35 Cory, M.; Tidwell, R. R.; Fairley, T. A. Structure and DNA binding activity of analogues of 1,5-bis(4- amidinophenoxy)pentane (pentamidine). *J. Med. Chem.* **1992**, *35*, 431–438.
- 36 Brown, D. G.; Sanderson, M. R.; Skelly, J. V.; Jenkins, T. C.; Brown, T.; Garman, E.; Stuart, D. I.; Neidle, S. Crystal structure of a berenil-dodecanucleotide complex: The role of water in sequence-specific ligand binding. *EMBO J.* **1990**, *9*, 1329–1334.
- 37 Brown, D. G.; Sanderson, M. R.; Garman, E.; Neidle, S. Crystal structure of a berenil-d(CGCAAATTTGCG) complex. An example of drug-DNA recognition based on sequence-dependent structural features. *J. Mol. Biol.* **1992**, *226*, 481–490.
- 38 Nguyen, B.; Lee, M. P.; Hamelberg, D.; Joubert, A.; Bailly, C.; Brun, R.; Neidle, S.; Wilson, W. D. Strong binding in the DNA minor groove by an aromatic diamidine with a shape that does not match the curvature of the groove. *J. Am. Chem. Soc.* **2002**, *124*, 13680–13681.
- 39 Tanious, F. A.; Hamelberg, D.; Bailly, C.; Czarny, A.; Boykin, D. W.; Wilson, W. D. DNA sequence dependent monomer-dimer binding modulation of asymmetric benzimidazole derivatives. *J. Am. Chem. Soc.* **2004**, *126*, 143–153.
- 40 Kopka, M. L.; Fratini, A. V.; Drew, H. R.; Dickerson, R. E. Ordered water structure around a B-DNA dodecamer. A quantitative study. *J. Mol. Biol.* **1983**, *163*, 129–146.
- 41 Clark, G. R.; Boykin, D. W.; Czarny, A.; Neidle, S. Structure of a bis-amidinium derivative of Hoechst 33258 complexed to dodecanucleotide  $d(CGCGAATTCGCG)_2$ : The role of hydrogen bonding in minor groove drug-DNA recognition. *Nucleic Acids Res.* **1997**, *25*, 1510–1515.
- 42 Campbell, N. H.; Evans, D. A.; Lee, M. P.; Parkinson, G. N.; Neidle, S. Targeting the DNA minor groove with fused ring dicationic compounds: Comparison of in silico screening and a high-resolution crystal structure. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 15–19.
- 43 Mallena, S.; Lee, M. P.; Bailly, C.; Neidle, S.; Kumar, A.; Boykin, D. W.; Wilson, W. D. Thiophene-based diamidine forms a “super” AT binding minor groove agent. *J. Am. Chem. Soc.* **2004**, *126*, 13659–13669.
- 44 Wang, L.; Bailly, C.; Kumar, A.; Ding, D.; Bajic, M.; Boykin, D. W.; Wilson, W. D. Specific molecular recognition of mixed nucleic acid sequences: An aromatic dication that binds in the DNA minor groove as a dimer. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12–16.
- 45 Bailly, C.; Tardy, C.; Wang, L.; Armitage, B.; Hopkins, K.; Kumar, A.; Schuster, G. B.; Boykin, D. W.; Wilson, W. D. Recognition of ATGA sequences by the unfused aromatic dication DB293 forming stacked dimers in the DNA minor groove. *Biochemistry* **2001**, *40*, 9770–9779.
- 46 Munde, M.; Lee, M.; Neidle, S.; Arafa, R.; Boykin, D. W.; Liu, Y.; Bailly, C.; Wilson, W. D. Induced fit conformational changes of a “reversed amidine” heterocycle: Optimized interactions in a DNA minor groove complex. *J. Am. Chem. Soc.* **2007**, *129*, 5688–5698.
- 47 Larsen, T. A.; Goodsell, D. S.; Cascio, D.; Grzeskowiak, K.; Dickerson, R. E. The structure of DAPI bound to DNA. *J. Biomol. Struct. Dyn.* **1989**, *7*, 477–491.
- 48 Nguyen, B.; Hamelberg, D.; Bailly, C.; Colson, P.; Stanek, J.; Brun, R.; Neidle, S.; Wilson, W. D. Characterization of a novel DNA minor-groove complex. *Biophys. J.* **2004**, *86*, 1028–1041.
- 49 Nguyen, B.; Stanek, J.; Wilson, W. D. Binding-linked protonation of a DNA minor-groove agent. *Biophys. J.* **2006**, *90*, 1319–1328.
- 50 Miao, Y.; Lee, M. P.; Parkinson, G. N.; Batista-Parra, A.; Ismail, M. A.; Neidle, S.; Boykin, D. W.; Wilson, W. D. Out-of-shape DNA minor groove binders: Induced fit interactions of heterocyclic dications with the DNA minor groove. *Biochemistry* **2005**, *44*, 14701–14708.
- 51 Woynarowski, J. M. Targeting critical regions in genomic DNA with AT-specific anticancer drugs. *Biochim. Biophys. Acta* **2002**, *1587*, 300–308.
- 52 Munde, M.; Ismail, M. A.; Arafa, R.; Peixoto, P.; Collar, C. J.; Liu, Y.; Hu, L.; David-Cordonnier, M. H.; Lansiaux, A.; Bailly, C.; Boykin, D. W.; Wilson, W. D. Design of DNA minor groove binding diamidines that recognize GC base pair sequences: A dimeric-hinge interaction motif. *J. Am. Chem. Soc.* **2007**, *129*, 13732–13743.
- 53 De Koning, H. P.; Jarvis, S. M. Adenosine transporters in bloodstream forms of *Trypanosoma brucei brucei*: Substrate recognition motifs and affinity for trypanocidal drugs. *Mol. Pharmacol.* **1999**, *56*, 1162–1170.
- 54 Barrett, M. P.; Fairlamb, A. H. The biochemical basis of arsenical-diamidine cross-resistance in African trypanosomes. *Parasitol. Today* **1999**, *15*, 136–140.
- 55 Maser, P.; Luscher, A.; Kaminsky, R. Drug transport and drug resistance in African trypanosomes. *Drug Resist. Updates* **2003**, *6*, 281–290.
- 56 Lacy, E. R.; Madsen, E. M.; Lee, M.; Wilson, W. D. In *DNA and RNA Binders: From Small Molecules to Drugs*; Demeunynck, M., Bailly, C., Wilson, W. D., Eds.; WILEY-VCH: Weinheim, Germany, 2003; Vol. 2; pp 384–413.