

Hydrogen-Bonding Patterns of Minor Groove-Binder–DNA Complexes Reveal Criteria for Discovery of New Scaffolds

Gudrun M. Spitzer,^{*†} Bernd Wellenzohn,[‡] Patrick Markt,[§] Johannes Kirchmair,[§] Thierry Langer,^{||} and Klaus R. Liedl[†]

Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, Innrain 52a, A-6020 Innsbruck, Austria, Boehringer Ingelheim Pharma GmbH & Co. KG, Birkendorfer Strasse 65, D-88397 Biberach/Riss, Germany, CAMD Group, Institute of Pharmacy, University of Innsbruck, Innrain 52c, A-6020 Innsbruck, Austria, and Prestwick Chemicals, Inc., Boulevard Gonthier d'Andernach, Parc d'Innovation, 67400 Illkirch, France

Received December 15, 2008

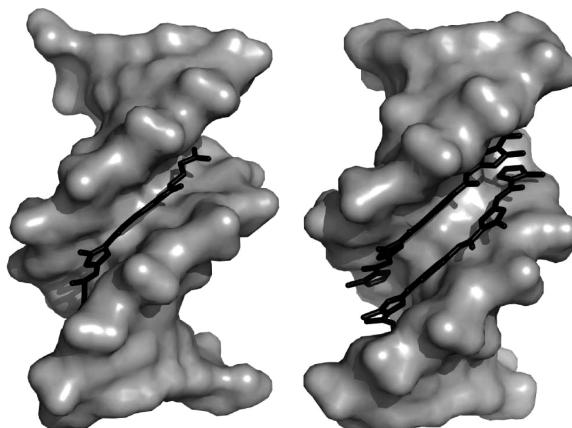
Minor groove-binding ligands are able to control gene expression and are of great interest for therapeutic applications. We extracted hydrogen-bonding geometries from all available structures of minor groove-binder–DNA complexes of two noncovalent binding modes, namely 1:1 (including hairpin and cyclic ligands) and 2:1 ligand/DNA binding. Positions of the ligand atoms involved in hydrogen bonding deviate from idealized hydrogen bond geometries and do not exploit the possibilities indicated by water molecules. Therefore, we suggest the inclusion of shape-based descriptors rather than hydrogen-bond patterns in virtual screening protocols for the identification of innovative minor groove-binding scaffolds.

INTRODUCTION

Small DNA minor groove-binding ligands provide the possibility to control the expression of genes,^{1–7} as has been shown by many examples.^{8,9} Their mechanism of action relies on noncovalent interactions with the floor and the walls of the minor groove, with sequence preference¹⁰ or, in other cases, even sequence selectivity.^{11–19} The development of new compounds is an active research topic^{20–23} because of their potential as therapeutic tools for various medicinal tasks. Examples reach from cancer therapy,^{24–28} antibacterial,^{29–33} antiprotozoal,^{21,34,35} or antifungal action,^{35–37} to the treatment of viral infections and other diseases. In contrast to conventional drugs, the target molecule is not a protein but the DNA itself, in particular, the sequences responsible for gene expression and consequentially for synthesis of the target protein.

A minor groove-binding ligand has to meet several requirements to be of therapeutic relevance: Oral bioavailability is desirable but—according to Lipinski's “rule of five”^{38,39}—demands an upper limit for the molecular weight and a balanced ratio of hydrophilic and hydrophobic parts in the molecule. Especially if aiming at full sequence readout, chemists tend to end up with molecules of considerable molecular weight^{40–45} and large numbers of hydrogen-bond donors. Moreover, the set of scaffolds used is rather limited, in particular, for the sequence-specific side-by-side binding molecules^{6,46,47} composed of heteroaromatic rings (mostly pyrrole) concatenated by amide bonds (see right side of Scheme 1).

Scheme 1. Left: A typical single-ribbon minor groove binder (PDB 1D86). Right: a double-ribbon ligand (PDB 1CVX)



With the dimension of chemical space in mind, it seems likely that alternative scaffolds could be identified. Other scientists searched for new scaffolds among natural products in sea creatures.⁴⁸ This strategy led to the discovery of trabectedin, a minor groove binding agent that is now approved for use in patients with advanced soft tissue sarcoma.⁴⁹ Rather than products from the sea, we intend to exploit collections of already known, commercially available or even virtual compounds. A detailed understanding of the interactions involved helps to reasonably choose a small subset of molecules for closer examination.

In our present work, we chose a statistical approach to inspect positions of interaction centers in order to revise the assumption of the central role of hydrogen-bonding pattern complementarity between the ligand and DNA. From the set of all public DNA–minor groove-binder complex structures, we extracted hydrogen-bond coordinates, analyzed distances between interaction centers, and compared geometries by aligning bases and base pairs.

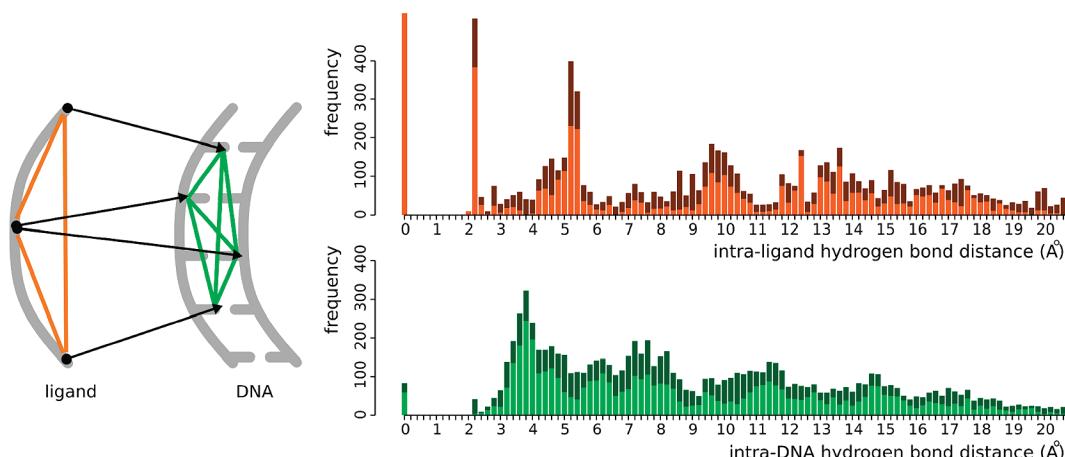
* Corresponding author e-mail: gudrun.spitzer@uibk.ac.at.

† Department of Theoretical Chemistry—University of Innsbruck.

‡ Boehringer–Ingelheim Biberach.

§ CAMD—Group University of Innsbruck.

|| Prestwick Chemicals.

Scheme 2. Plotted Distances between Hydrogen-Donating or -Accepting Atoms^a

^a Orange: distances between hydrogen-bonding atoms within the ligand. Green: distances between hydrogen-bonding atoms in the DNA minor groove. Light colors: data from single-ribbon complexes. Dark colors: data from double-ribbon complexes.

METHODS

We searched the Protein Data Bank (PDB)⁵⁰ as well as the Nucleic Acid Database (NDB)⁵¹ for structures of minor groove binders in complex with B-DNA. In all, 107 entries were selected from a total of 226. For our analysis, we did not consider structures with covalent bonds between the ligand and DNA, intercalators or intercalators fused with minor groove binding molecules, ligands binding via metal complexation, and protein-DNA complexes, as most of them induce severe distortions in the DNA structures. We did not filter by resolution, as we assume structural artifacts to be diminished by the number of structures. Resolution values did not exceed 2.65 Å, with a mean value of 2.1 Å. Six complex structures resulted from NMR analysis; the remaining 101 were resolved by X-ray diffraction. In cases with several minor groove binders within one structure file, the coordinates were split into the respective number of files and analyzed separately. This procedure yielded 124 complexes.

The complexes were grouped by interaction type: (i) 95 monomer minor groove binders, for example, natural products like distamycin or netropsin, the dyes from Hoechst, or the drug pentamidine and its analogs. We will further refer to them as “single-ribbon” ligands to distinguish them from (ii) 29 double-ribbon minor groove binders. The latter occupy the minor groove whether in a side-by-side dimeric binding mode or with both parts of the dimer connected to form a hairpin or a circle. See Scheme 1 for a graphical representation of the two ligand binding types.

Automated software tools turned out to be not sufficiently reliable in assigning protonation and tautomerization states, hydrogen placement, interaction, and atom types. Therefore, these duties were carefully completed manually. For consistent interaction pattern retrieval, it was necessary to accept hydrogen-bond lengths of up to 4 Å. We are aware that this exceeds default values of numerous software tools by about 0.5 Å. Nevertheless, patterns found in the distance plot presented in Scheme 2 hardly change when lowering the distance cutoff during analysis of the data collection. It is known that hydrogen-bond angles in minor groove-binder complexes depart from linearity; therefore, no angle restrictions were introduced.⁵² No examples of unexpected hydrogen-bond angles were found anyway. We treated modified bases

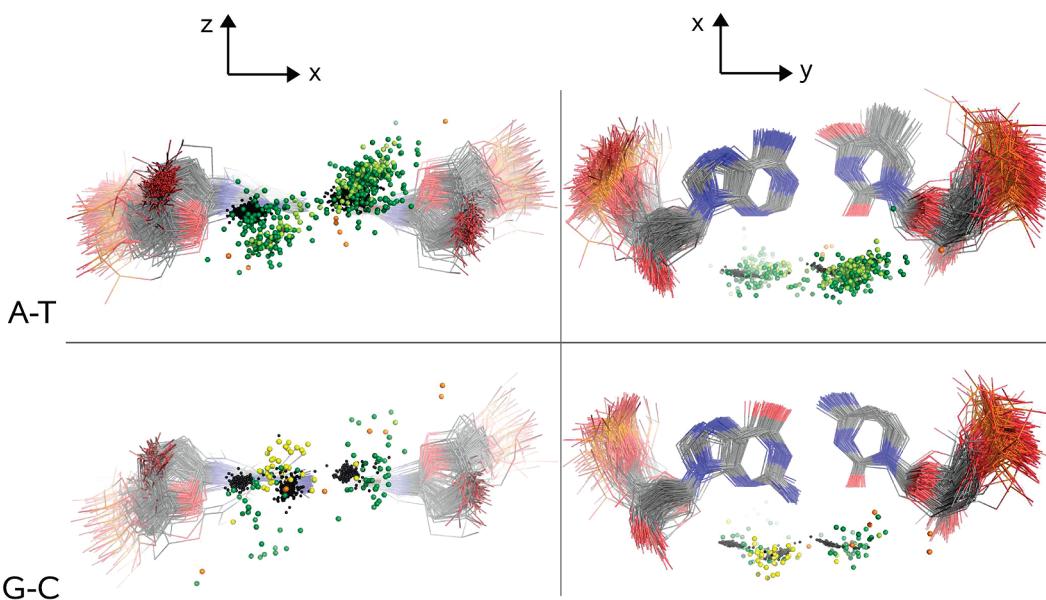
like their unmodified counterparts as long as the modification did not affect the minor groove.

We searched the NDB for solvated DNA duplexes for analysis of water–DNA interactions. Selection criteria were B-DNA structure and absence of modified bases, mismatch base pairs, ligands, or metal ions. We found 43 structures. For comparability reasons, the cutoff for hydrogen bonds was again chosen to be 4 Å. Comparison with the results automatically created by an online service⁵³ which uses 3.5 Å as distance cutoff shows that our limit does not introduce artificial confusion to the solvent positions.

For the base pair superposition, two vectors were defined. The first vector connected two points in the middle of the base pair: (i) the center between the hydrogen-bond acceptor atoms and (ii) the center between two atoms at the major groove side of the bases (C4 for thymine and cytosine, C6 for adenine and guanine). The second vector joined the hydrogen-bond acceptor centers of each nucleobase. During the superposition process, we translated the whole base pair to bring the first vector to the center of origin and performed three rotations. In the end, the first vector coincided with the DNA *x* axis, and the second vector was coplanar with the *x*–*y* plane. Axis nomenclature follows the recommendations given by Dickerson.⁵⁴ In contrast to examples found in the literature,^{53,55,56} we did not idealize base pair structures but regarded them as rigid units. Therefore, DNA flexibility and the diversity of deviations from idealized structures remain visible in our pictures.

We obtained the coordinates of the idealized hydrogen-bond interaction partners by simple vector calculation on the basis of nucleobase coordinates. The length of idealized hydrogen bonds was chosen to be 3 Å. This is consistent with default definitions in several software packages, with values reported in the literature⁵⁷ and with the average value we found for the analyzed structures of minor groove binders. In all images, these optimal positions appear as black spheres. They are intended as references for the visualization of ligand atom distribution and give an impression of the diversity of base pair geometries.

Pymol⁵⁸ was used for creation of all molecular representations.

Scheme 3. Base Pair Alignment with Ligand Atoms^a

^a Small spheres show the positions of ligand atoms relative to the aligned base pairs. Five types of hydrogen-bonding atoms are color-coded: (i) dark green, amide nitrogen atoms donating hydrogen; (ii) light green, nitrogen atoms in aromatic environment donating hydrogen; (iii) yellow, nitrogen atoms accepting hydrogen; (iv) orange, oxygen; (v) black and slightly smaller, idealized hydrogen-bond partners calculated on the basis of nucleobase coordinates.

RESULTS AND DISCUSSION

Distance Patterns. Identification of an appropriate filter method is a crucial task toward the identification of innovative minor groove-binder scaffolds in large chemical databases. In an initial step, we examined hydrogen-bond patterns for their usability as filter criteria. For this purpose, we calculated all intraligand distances between hydrogen-bond-accepting or -donating heavy atoms. The resulting histogram is shown in the upper part of Scheme 2; distances in double-ribbon structures are colored darker than those collected in single-ribbon structures. The bar at 0 Å is cut for y-axis comparability with the second diagram. Originally, it had a height of nearly 1000 and represented numerous bifurcated hydrogen bonds mainly occurring in single-ribbon ligands. The second outstanding bar represents the distance between two nitrogen atoms that belong to an amidinium or guanidinium group both interacting with the DNA.

The next noticeable bars arise at 5.2 and 5.4 Å. They represent hydrogen-bond-donor centers separated by five stretched covalent bonds, as can be found in all distamycin-like polyamides and other structures dominated by a sequence of aromatic rings. The smaller bars at shorter distances either belong to hydrogen bonds separated by five bonds in a less stretched conformation or to interaction centers separated by less than five bonds. However, smaller distances can rather be observed in terminal regions of ligands due to tail group variation. The next maximum can be found around 10 Å. By far less distinct, it is still twice the previously discussed distance of five stretched bonds between two interaction points taking into account the bend of the ligand. The addition of another unit of five bonds leads to a distance of 13–14 Å in the bent ligand. The bars are elevated at these distance values but cannot be clearly distinguished anymore.

After straightforward analysis of ligand distances, we expected to obtain patterns similar or even more evident for intra-DNA hydrogen-bond interaction centers, as there are

only four different bases and clear rules for the composition of B-DNA. Although the first maximum is interpretable as the distance between two bases on the same strand or two bases on different strands not belonging to the same base pair, the initial assumption did not turn out to be correct. The light-green bars (bottom half of Scheme 2) show a slight wavelike alteration between higher and lower bars, whereas the dark bars are more even in length and often rather balance the light-green pattern.

Since the ligands show a far more distinct hydrogen-bond pattern than the minor groove floor, this pattern is not valuable as a filter for database screening. Using it would end in an extraction of well-known minor groove binders. These molecules do not seem to exploit the range of hydrogen-bond geometries that can be derived from the diversity of intra-DNA distances.

Base Pair Superposition. After analyzing the distances between hydrogen bonds, we focused on their respective geometries. Alignment of all base pairs with the adjacent ligand atoms gives a statistical overview over preferred interaction positions. We restricted the data set to exclusively contain uncharged hydrogen-bond interactions, as they are more specific than charged ones. The resulting pictures are presented in Scheme 3.

For the AT base pair alignment, the statistic sampling is considerably better than for GC. On the one hand, this is determined by the preference of single-ribbon ligands for AT-rich sequences and the numerical dominance of these complexes (95 single-ribbon versus 29 double-ribbon structures). On the other hand, there are 18 structures containing inosine instead of guanine, leading to a further decrease of the number of aligned GC base pairs. In Scheme 3, the heavy atoms involved in hydrogen bonding are color-coded: (i) dark green, amide nitrogen atoms donating hydrogen; (ii) light green, nitrogen atoms in an aromatic environment donating hydrogen; (iii) yellow, nitrogen atoms accepting hydrogen;

(iv) orange, oxygen; and (v) black and slightly smaller, idealized hydrogen-bond partners calculated on the basis of nucleobase coordinates. Whereas pictures on the left-hand side of Scheme 3 are intended for geometry inspection, pictures on the right-hand side give a general idea of the alignment quality.

The black spheres of idealized hydrogen-bond centers elongate base geometries toward the minor groove, thereby visualizing considerable differences among the base pairs in opening and propeller twist. Distances between pairs of idealized donor atoms range from 3 to 7 Å. The average distance found in AT base pairs is 4.55 Å (stdev: 0.58) and 4.90 Å (stdev: 0.44) for GC base pairs. The distances between the real hydrogen-bond donors are about 0.5 Å larger than the distances between the idealized hydrogen-bond donor centers.

With alignments in Scheme 3 in mind, it would not be too surprising to find deviations of base pair parameters from average values for the B-DNA structure in a direction toward the ligand interaction points. This would represent an increased opening and a negative propeller twist⁵⁹ and as a consequence a larger C1–C1 distance (D_{C1-C1}) and deviations from the default base–base hydrogen-bond lengths. The statistical average does however not confirm this impression. D_{C1-C1} is 10.6 Å (stdev: 0.34) for AT base pairs (default: 10.7 Å), and for GC base pairs, it is 10.7 Å (stdev: 0.25; default: 10.8 Å).⁶⁰ Base–base hydrogen-bond lengths overall correspond to values from high-resolution crystal structures.^{60,61} In thermodynamic experiments, DNA conformation was also found not to change significantly upon ligand binding.⁶² We conclude that for DNA it is energetically less favorable to leave the B-DNA conformation than to optimize the fit of hydrogen bonds.

Obviously, the ligand atoms' positions depicted in Scheme 3 do not meet expectations on the basis of geometric hydrogen-bond considerations represented by the black spheres. The top-left image with the AT base pair shows that hydrogen-bond donor atoms addressing the hydrogen-bond acceptor nitrogen prefer not to lie in the base plane but are displaced toward the 5' direction. The area covered by hydrogen-bonding partners and therefore deviations from the inferred optimal positions observed in the complexes studied are much larger than one might expect. In particular, this is true for the few oxygen atoms that appear at the boundaries of the point clouds.

Moreover, there is a clear tendency of donors opposite the AT base pair toward the middle of the minor groove and to the right of the black spheres. This aspect of ligand positioning contrasts expectations about the differences between purine and pyrimidine bases derived from geometric considerations for hydrogen bonding. These considerations would suggest that hydrogen-bonding partners of purine bases should rather lie near the backbone, whereas hydrogen-bonding partners of pyrimidine bases should prefer positions next to the middle of the minor groove. However, our analysis shows that this is not the case. The positioning is rather independent of the base type involved but depends on the backbone limiting the space and forcing the ligand to lie in a balanced position between the walls. This finding supports other reports of difficulties in distinguishing AT from TA base pairs.^{16,17,63–65} It is also in good agreement with earlier work where we did not succeed in describing

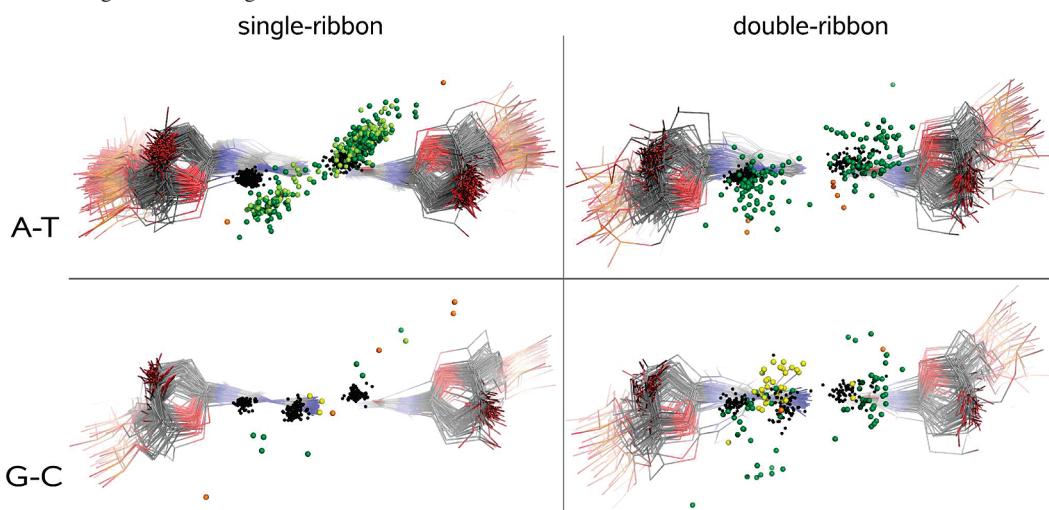
sequence-specific properties of minor groove interactions without taking into account the steric situation on the groove floor.⁶⁶ Similarly, others found van der Waals terms to be important for binding energy calculation.^{67–69} In contrast to other studies, we do not find that ligands around N3 (purine bases) are significantly more compact than around O2 (pyrimidine bases).⁷⁰

Observations discussed for the AT base pair concerning the deviation of ligand atoms from idealized positions are partially similar for the GC base pair. An important difference is the number of analyzed interactions; the sampling is considerably worse for the GC base pair. Furthermore, the ligand points do not accumulate around the positions most occupied in the AT base pair, but they are equally distributed in a diffuse cloud. In addition to other atom types in opposition to guanine, there are yellow spheres representing hydrogen-bond acceptor centers. They are also displaced from the base plane, but unlike donor atoms toward the 3' direction. Although not coinciding like the idealized interaction points, they are not as scattered as donor atoms.

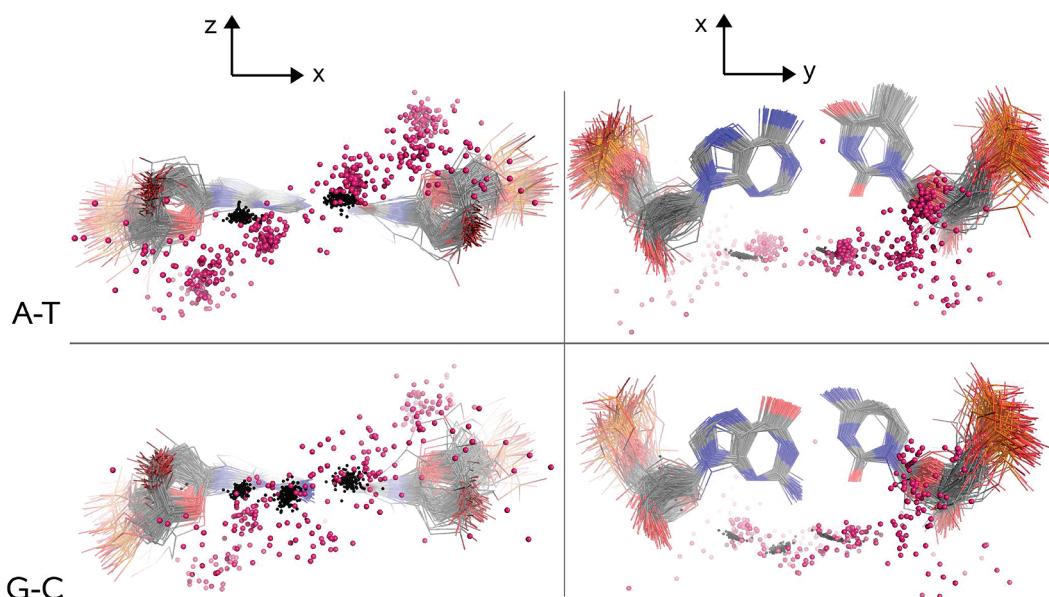
Differences between Interaction Types. The differences between single-ribbon and more sequence-selective double-ribbon minor groove binders are shown in Scheme 4. The latter interact with both GC and AT base pairs, whereas the former show a strong preference for AT base pairs. This preference is also evident from the DNA sequences chosen for structure analysis: In about two-thirds of the experiments, the Dickerson Drew dodecamer was used; half of the remaining sequences are similar in having a central AT tract. Numerous light-green spheres in the single-ribbon pictures represent nitrogen donor atoms in aromatic environments. These are missing in the double-ribbon pictures, indicating that the limited set of scaffolds used for unselective DNA binding is even more restricted for sequence-selective recognition.

The displacement of ligand atoms away from the base plane is most pronounced for single-ribbon minor groove binders near AT base pairs. They usually lie in the middle of the minor groove and form bifurcated hydrogen bonds with the acceptor atoms of one base and the next base in the opposite strand, therefore sitting between the two base planes. Better sequence selectivity and high affinity suggest that hydrogen bonds of double-ribbon ligands could be in some way more ideal than others. Indeed, a fraction of ligand spheres coincides with the black spheres of idealized hydrogen bonds. These are amide bonds from donors in central regions of the ligands, whereas terminal hydrogen-bond donors do not fit that well, as the distances between the donor groups do not correspond to the distances between two bases. However, a perfect hydrogen-bond geometry does not seem to be too important, as by far not all ligand points lie there; especially in case of the GC base pair, the distribution is quite diffuse.

Single-ribbon ligands clearly prefer AT tracts that are well-known for their narrow minor groove. However, the shape or width of the AT base pair does not seem to impose any restriction on double-ribbon ligand binding. Both AT and GC base pairs are almost equally represented in Scheme 3, indicating that the small groove width is no intrinsic property of the AT base pair but arises from the accommodation of ligand or solvent molecules.⁷¹

Scheme 4. Base Pair Alignment with Ligand Atoms^a

^a For description of the color scheme, see Scheme 3. Left: Coordinates extracted from single-ribbon ligand–DNA structures. Right: coordinates from double-ribbon ligand–DNA structures.

Scheme 5. Base Pair Alignment of Ligand-Free DNA^a

^a Pink spheres represent water positions. Small black spheres mark idealized positions of hydrogen-bonding partners. Interaction positions in opposition to the base edges are in good agreement with those found for minor groove binding ligands, whereas positions in opposition to ribose O4' are occupied by water but remain largely vacant in ligand–DNA complexes.

Positions of Water Molecules. Scheme 5 shows aligned base pairs together with adjacent water positions extracted from ligand-free DNA structures. Again, small black spheres represent idealized geometries of hydrogen-bonding partners. Overall, water interaction sites next to the bases are slightly more compact than ligand sites. We did not find a difference in compactness near pyrimidine versus purine bases as reported by others.⁷⁰ The similarity between interaction positions in opposition to the base acceptors in AT base pairs (upper half of Scheme 5) confirms the reliability of water molecules as probes for ligand–DNA interactions⁵⁷ and binding sites at protein–DNA interfaces.⁷² The similarity between ligand and water positions is less pronounced for the GC alignments. Both ligand and water positions do not show such a well-defined pattern as they do next to AT base pairs. Still, ligand and water positions are similar to each other in the way that they are distributed over a comparable region in the bottom-left pictures of Schemes 3 and 5.

Obviously, guanine's exocyclic amine with its hydrogen-bond donor function has a disruptive effect on both ligand and water interaction patterns.⁷¹ The highly ordered water molecules next to AT base pairs cause an important gain of entropy if replaced during ligand binding.⁷¹

Apart from similarities next to the base edges, there is a major discrepancy between water and ligand positions: whereas the interaction site in opposition to the ribose O4' atoms is as well occupied by water in ligand-free DNA structures (Scheme 5), there are only some outliers in the ligand alignments in Scheme 3. Obviously, ligands presently known do not exploit the hydrogen-bonding potential next to ribose O4' atoms.

CONCLUSIONS

We have performed a statistical analysis of hydrogen-bond interactions in minor groove-binder complex structures to

revise the importance of hydrogen-bond patterns for the discovery of novel ligand scaffolds. Neither does the DNA have to adapt to fit to the ligand positions nor do the ligand's hydrogen-bond centers have to perfectly match the DNA acceptor pattern. This contrasts with previous findings by Moravek et al.⁷⁰ about the stability of binding positions and their conclusion that spacing between hydrogen-bond donors is crucial for effective minor groove binding. We found a rather large space accessible to individual hydrogen-bond partners, although they still do not exploit the possibilities demonstrated by water molecules or the diversity we found in the intra-DNA distance plot.

Obviously, effects apart from hydrogen bonds strongly influence the structural fit of ligands. This corresponds to experimental findings confirming the importance of van der Waals contacts and therefore shape complementary⁷³ and the influence of hydrophobic driving forces.⁷⁴ Therefore, we discourage simple reproduction of known hydrogen-bond patterns during scaffold development or discovery. Instead, we propose placing more emphasis on shape complementarity, although the flexibility of the backbone makes this task quite challenging. However, widening the scaffold frame for unspecific minor groove binders could help to increase the chance of new input for the development of specific sequence readout.

ACKNOWLEDGMENT

This work was supported by the Austrian Science Fund (Project P19756).

Supporting Information Available: Lists of PDB codes for all PDB structures used for this study. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES AND NOTES

- (1) Neidle, S. DNA minor-groove recognition by small molecules. *Nat. Prod. Rep.* **2001**, *18*, 291–309.
- (2) Stafford, R. L.; Arndt, H.-D.; Brezinski, M. L.; Ansari, A. Z.; Dervan, P. B. Minimization of a protein-DNA dimerizer. *J. Am. Chem. Soc.* **2007**, *129*, 2660–2668.
- (3) Uil, T. G.; Haisma, H. J.; Rots, M. G. Therapeutic modulation of endogenous gene function by agents with designed DNA-sequence specificities. *Nucleic Acids Res.* **2003**, *31*, 6064–6078.
- (4) Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. Regulation of gene expression by small molecules. *Nature* **1997**, *387*, 202–205.
- (5) Dervan, P. B.; Edelson, B. S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, *13*, 264–299.
- (6) Dervan, P. B.; Doss, R. M.; Marques, M. A. Programmable DNA binding oligomers for control of transcription. *Curr. Med. Chem. Anticancer Agents* **2005**, *5*, 373–387.
- (7) Melander, C.; Burnett, R.; Gottesfeld, J. M. Regulation of gene expression with pyrrole-imidazole polyamides. *J. Biotechnol.* **2004**, *112*, 195–220.
- (8) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem. Biol.* **2007**, *2*, 561–571.
- (9) Nickols, N. G.; Dervan, P. B. Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 10418–10423.
- (10) Mallena, S.; Lee, M. P. H.; 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.
- (11) James, P. L.; Merkina, E. E.; Khalaf, A. I.; Suckling, C. J.; Waigh, R. D.; Brown, T.; Fox, K. R. DNA sequence recognition by an isopropyl substituted thiazole polyamide. *Nucleic Acids Res.* **2004**, *32*, 3410–3417.
- (12) Nelson, S. M.; Ferguson, L. R.; Denny, W. A. Non-covalent ligand/DNA interactions: minor groove binding agents. *Mutat. Res.* **2007**, *623*, 24–40.
- (13) Puckett, J. W.; Muzikar, K. A.; Tietjen, J.; Warren, C. L.; Ansari, A. Z.; Dervan, P. B. Quantitative microarray profiling of DNA-binding molecules. *J. Am. Chem. Soc.* **2007**, *129*, 12310–12319.
- (14) Briehn, C. A.; Weyermann, P.; Dervan, P. B. Alternative heterocycles for DNA recognition: the benzimidazole/imidazole pair. *Chem.—Eur. J.* **2003**, *9*, 2110–2122.
- (15) Chenoweth, D. M.; Viger, A.; Dervan, P. B. Fluorescent sequence-specific dsDNA binding oligomers. *J. Am. Chem. Soc.* **2007**, *129*, 2216–2217.
- (16) Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. A structural basis for recognition of A.T and T.A base pairs in the minor groove of B-DNA. *Science* **1998**, *282*, 111–115.
- (17) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **1998**, *391*, 468–471.
- (18) White, S.; Turner, J. M.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B. Affinity and specificity of multiple hydroxypyrrrole/pyrrole ring pairings for coded recognition of DNA. *J. Am. Chem. Soc.* **1999**, *121*, 260–261.
- (19) Doss, R. M.; Marques, M. A.; Foister, S.; Chenoweth, D. M.; Dervan, P. B. Programmable oligomers for minor groove DNA recognition. *J. Am. Chem. Soc.* **2006**, *128*, 9074–9079.
- (20) Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B. Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J. Am. Chem. Soc.* **2008**, *130*, 6859–6866.
- (21) Ismail, M. A.; Arafa, R. K.; Wenzler, T.; Brun, R.; Tanious, F. A.; Wilson, W. D.; Boykin, D. W. Synthesis and antiprotozoal activity of novel bis-benzamidino imidazo[1,2-a]pyridines and 5,6,7,8-tetrahydroimidazo[1,2-a]pyridines. *Bioorg. Med. Chem.* **2008**, *16*, 683–691.
- (22) Mundt, M.; Lee, M.; Neidle, S.; Arafa, R.; Boykin, D. W.; Liu, Y.; Bailly, C.; Wilson, W. D. Induced fit conformational changes of a “reversed amide” heterocycle: Optimized interactions in a DNA minor groove complex. *J. Am. Chem. Soc.* **2007**, *129*, 5688–5698.
- (23) Athri, P.; Wenzler, T.; Ruiz, P.; Brun, R.; Boykin, D. W.; Tidwell, R.; Wilson, W. D. 3D QSAR on a library of heterocyclic diamidine derivatives with antiparasitic activity. *Bioorg. Med. Chem.* **2006**, *14*, 3144–3152.
- (24) Baraldi, P. G.; Bovero, A.; Fruttarolo, D.; Preti, F.; Tabrizi, M. A.; Pavani, M. G.; Romagnoli, R. DNA minor groove binders as potential antitumor and antimicrobial agents. *Med. Res. Rev.* **2004**, *24*, 475–528.
- (25) Kageyama, Y.; Sugiyama, H.; Ayame, H.; Iwai, A.; Fujii, Y.; Huang, L. E.; Kizaka-Kondoh, S.; Hiraoka, M.; Kihara, K. Suppression of VEGF transcription in renal cell carcinoma cells by pyrrole-imidazole hairpin polyamides targeting the hypoxia responsive element. *Acta Oncol.* **2006**, *45*, 317–324.
- (26) Hurley, L. H. DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer* **2002**, *2*, 188–200.
- (27) Pandolfi, P. P. Transcription therapy for cancer. *Oncogene* **2001**, *20*, 3116–3127.
- (28) Mapp, A. K. Finding their groove, bifunctional molecules arrest growth of cancer cells. *Chem. Biol.* **2004**, *11*, 1480–1482.
- (29) Dyatkina, N. B.; Roberts, C. D.; Keicher, J. D.; Dai, Y.; Nadherny, J. P.; Zhang, W.; Schmitz, U.; Kongpachith, A.; Fung, K.; Novikov, A. A.; Lou, L.; Velligan, M.; Khorlin, A. A.; Chen, M. S. Minor groove DNA binders as antimicrobial agents. 1. Pyrrole tetraamides are potent antibacterials against vancomycin resistant Enterococci and methicillin resistant Staphylococcus aureus. *J. Med. Chem.* **2002**, *45*, 805–817.
- (30) Burli, R. W.; Kaizerman, J. A.; Duan, J.-X.; Jones, P.; Johnson, K. W.; Iwamoto, M.; Truong, K.; Hu, W.; Stanton, T.; Chen, A.; Touami, S.; Gross, M.; Jiang, V.; Ge, Y.; Moser, H. E. DNA binding ligands with in vivo efficacy in murine models of bacterial infection: optimization of internal aromatic amino acids. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2067–2072.
- (31) Burli, R. W.; McMinn, D.; Kaizerman, J. A.; Hu, W.; Ge, Y.; Pack, Q.; Jiang, V.; Gross, M.; Garcia, M.; Tanaka, R.; Moser, H. E. DNA binding ligands targeting drug-resistant Gram-positive bacteria. Part 1: Internal benzimidazole derivatives. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1253–1257.
- (32) Burli, R. W.; Jones, P.; McMinn, D.; Le, Q.; Duan, J.-X.; Kaizerman, J. A.; Difuntorum, S.; Moser, H. E. DNA binding ligands targeting drug-resistant Gram-positive bacteria. Part 2: C-terminal benzimidazoles and derivatives. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1259–1263.
- (33) Kaizerman, J. A.; Gross, M. I.; Ge, Y.; White, S.; Hu, W.; Duan, J.-X.; Baird, E. E.; Johnson, K. W.; Tanaka, R. D.; Moser, H. E.; Burli, R. W. DNA binding ligands targeting drug-resistant bacteria: structure, activity, and pharmacology. *J. Med. Chem.* **2003**, *46*, 3914–3929.

- (34) Rodriguez, F.; Rozas, I.; Kaiser, M.; Brun, R.; Nguyen, B.; Wilson, W. D.; Garcia, R. N.; Dardonville, C. New bis(2-aminoimidazoline) and bisguanidine DNA minor groove binders with potent in vivo antitrypanosomal and antiplasmoidal activity. *J. Med. Chem.* **2008**, *51*, 909–923.
- (35) Anthony, N. G.; Breen, D.; Clarke, J.; Donoghue, G.; Drummond, A. J.; Ellis, E. M.; Gemmell, C. G.; Helesbeux, J.-J.; Hunter, I. S.; Khalaf, A. I.; Mackay, S. P.; Parkinson, J. A.; Suckling, C. J.; Waigh, R. D. Antimicrobial lexitropsins containing amide, amidine, and alkene linking groups. *J. Med. Chem.* **2007**, *50*, 6116–6125.
- (36) Suckling, C. J. Molecular recognition and physicochemical properties in the discovery of selective antibacterial minor groove binders. *J. Phys. Org. Chem.* **2008**, *21*, 575–583.
- (37) Marini, N. J.; Baliga, R.; Taylor, M. J.; White, S.; Simpson, P.; Tsai, L.; Baird, E. E. DNA binding hairpin polyamides with antifungal activity. *Chem. Biol.* **2003**, *10*, 635–644.
- (38) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.
- (39) Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235–249.
- (40) Tiberghien, A. C.; Evans, D. A.; Kiakos, K.; Martin, C. R.; Hartley, J. A.; Thurston, D. E.; Howard, P. W. An asymmetric C8/C8'-trityrrole-linked sequence-selective pyrrolo[2,1-c][1,4]benzodiazepine (PBD) dimer DNA interstrand cross-linking agent spanning 11 DNA base pairs. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2073–2077.
- (41) Neidle, S.; Thurston, D. E. Chemical approaches to the discovery and development of cancer therapies. *Nat. Rev. Cancer* **2005**, *5*, 285–296.
- (42) Renneberg, D.; Dervan, P. B. Imidazopyridine/Pyrrole and hydroxybenzimidazole/pyrrole pairs for DNA minor groove recognition. *J. Am. Chem. Soc.* **2003**, *125*, 5707–5716.
- (43) Trauger, J. W.; Baird, E. E.; Dervan, P. B. Extended hairpin polyamide motif for sequence-specific recognition in the minor groove of DNA. *Chem. Biol.* **1996**, *3*, 369–377.
- (44) Herman, D. M.; Baird, E. E.; Dervan, P. B. Tandem hairpin motif for recognition in the minor groove of DNA by pyrrole-Imidazole polyamides. *Chem.—Eur. J.* **1999**, *5*, 975–983.
- (45) Halby, L.; Ryabinin, V. A.; Sinyakov, A. N.; Boutorine, A. S. Functionalized head-to-head hairpin polyamides: synthesis, double-stranded DNA-binding activity and affinity. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3720–3724.
- (46) Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B. Cycle polyamide motif for recognition of the minor groove of DNA. *J. Am. Chem. Soc.* **1999**, *121*, 1121–1129.
- (47) Melander, C.; Herman, D. M.; Dervan, P. B. Discrimination of A/T sequences in the minor groove of DNA within a cyclic polyamide motif. *Chem.—Eur. J.* **2000**, *6*, 4487–4497.
- (48) Rinehart, K. L.; Holt, T. G.; Fregeau, N. L.; Keifer, P. A.; Wilson, G. R.; Perun, T. J.; Sakai, R.; Thompson, A. G.; Stroh, J. G.; Shield, L. S. Bioactive compounds from aquatic and terrestrial sources. *J. Nat. Prod.* **1990**, *53*, 771–792.
- (49) Carter, N. J.; Keam, S. J. Trabectedin: a review of its use in the management of soft tissue sarcoma and ovarian cancer. *Drugs* **2007**, *67*, 2257–2276.
- (50) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (51) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S. H.; Srinivasan, A. R.; Schneider, B. The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.* **1992**, *63*, 751–759.
- (52) Taberner, L.; Bella, J.; Aleman, C. Hydrogen bond geometry in DNA-minor groove binding drug complexes. *Nucleic Acids Res.* **1996**, *24*, 3458–3466.
- (53) Auffinger, P.; Hashem, Y. SwS: a solvation web service for nucleic acids. *Bioinformatics* **2007**, *23*, 1035–1037.
- (54) Dickerson, R. E. Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res.* **1989**, *17*, 1797–1803.
- (55) Schneider, B.; Berman, H. M. Hydration of the DNA bases is local. *Biophys. J.* **1995**, *69*, 2661–2669.
- (56) Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.* **2001**, *29*, 2860–2874.
- (57) Ge, W.; Schneider, B.; Olson, W. K. Knowledge-based elastic potentials for docking drugs or proteins with nucleic acids. *Biophys. J.* **2005**, *88*, 1166–1190.
- (58) DeLano, W. L. *The PyMOL Molecular Graphics System version 1.0*; DeLano Scientific LLC: Palo Alto, CA, 2008. <http://www.pymol.org> (accessed Feb 06, 2009).
- (59) Urbach, A. R.; Love, J. J.; Ross, S. A.; Dervan, P. B. Structure of a beta-alanine-linked polyamide bound to a full helical turn of purine tract DNA in the 1:1 motif. *J. Mol. Biol.* **2002**, *320*, 55–71.
- (60) Olson, W. K.; et al. A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* **2001**, *313*, 229–237.
- (61) Clowney, L.; Jain, S. C.; Srinivasan, A. R.; Westbrook, J.; Olson, W. K.; Berman, H. M. Geometric parameters in nucleic acids: nitrogenous bases. *J. Am. Chem. Soc.* **1996**, *118*, 509–518.
- (62) Chaires, J. B. A thermodynamic signature for drug-DNA binding mode. *Arch. Biochem. Biophys.* **2006**, *453*, 26–31.
- (63) Kielkopf, C. L.; Bremer, R. E.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. Structural effects of DNA sequence on T-A recognition by hydroxypyrrrole/pyrrole pairs in the minor groove. *J. Mol. Biol.* **2000**, *295*, 557–567.
- (64) Nguyen, D. H.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B. Alternative heterocycles for DNA recognition: an N-methylpyrazole/N-methylpyrrole pair specifies for A/T/T/A base pairs. *Bioorg. Med. Chem.* **2001**, *9*, 7–17.
- (65) Urbach, A. R.; Szewczyk, J. W.; White, S.; Turner, J. M.; Baird, E. E.; Dervan, P. B. Sequence selectivity of 3-hydroxypyrrrole/pyrrole ring pairings in the DNA minor groove. *J. Am. Chem. Soc.* **1999**, *121*, 11621–11629.
- (66) Spitzer, G. M.; Wellenzohn, B.; Laggner, C.; Langer, T.; Liedl, K. R. Sequence specific feature based pharmacophores for DNA minor groove description. *J. Chem. Inf. Model.* **2007**, *47*, 1580–1589.
- (67) Anthony, N. G.; Huchet, G.; Johnston, B. F.; Parkinson, J. A.; Suckling, C. J.; Waigh, R. D.; Mackay, S. P. In silico footprinting of ligands binding to the minor groove of DNA. *J. Chem. Inf. Model.* **2005**, *45*, 1896–1907.
- (68) Laughton, C. A.; Jenkins, T. C.; Fox, K. R.; Neidle, S. Interaction of berenil with the tyrT DNA sequence studied by footprinting and molecular modelling. Implications for the design of sequence-specific DNA recognition agents. *Nucleic Acids Res.* **1990**, *18*, 4479–4488.
- (69) Dolec, J.; Borstnik, U.; Hodosek, M.; Koller, J.; Janezic, D. An ab initio QM/MM study of the conformational stability of complexes formed by netropsin and DNA. The importance of van der Waals interactions and hydrogen bonding. *THEOCHEM* **2005**, *718*, 77–85.
- (70) Moravek, Z.; Neidle, S.; Schneider, B. Protein and drug interactions in the minor groove of DNA. *Nucleic Acids Res.* **2002**, *30*, 1182–1191.
- (71) Spitzer, G. M.; Fuchs, J. E.; Markt, P.; Kirchmair, J.; Wellenzohn, B.; Langer, T.; Liedl, K. R. Sequence-Specific Positions of Water Molecules at the Interface between DNA and Minor Groove Binders. *ChemPhysChem* **2008**, *9*, 2766–2771.
- (72) Woda, J.; Schneider, B.; Patel, K.; Mistry, K.; Berman, H. M. An Analysis of the Relationship between Hydration and Protein-DNA Interactions. *Biophys. J.* **1998**, *75*, 2170–2177.
- (73) Czarny, A.; Boykin, D. W.; Wood, A. A.; Nunn, C. M.; Neidle, S.; Zhao, M.; Wilson, W. D. Analysis of van der Waals and electrostatic contributions in the interactions of minor groove binding benzimidazoles with DNA. *J. Am. Chem. Soc.* **1995**, *117*, 4716–4717.
- (74) Fishleigh, R. V.; Fox, K. R.; Khalaf, A. I.; Pitt, A. R.; Scobie, M.; Suckling, C. J.; Urwin, J.; Waigh, R. D.; Young, S. C. DNA binding, solubility, and partitioning characteristics of extended lexitropsins. *J. Med. Chem.* **2000**, *43*, 3257–3266.