Volume 281, number 1,2, 209–211 FEBS 09591 April 1991

ADONIS 001457939100301I

Geometrical correlations useful for design of sequence-specific DNA narrow groove binding ligands

A.S. Zasedatelev

Engelhardt Institute of Molecular Biology, Academy of Sciences of the USSR, 117984 Moscow B-334, USSR

Received 30 January 1991

Isohelical geometry of sequence-specific DNA narrow groove binding ligands was analyzed in terms of H-bond donor/acceptor complementarity between the base pair atoms facing into the narrow groove and the corresponding H-bond donating atoms regularly disposed along the ligand molecule. Spatial correlations found in analytical form were applied to analysis of naturally occurring and hypothetical drug molecule structures. For the case of B-like isohelices the permitted values of the distance L_0 between each two neighboring H-bond donating atoms of the ligand as well as the bending angle τ_0 of the line subsequently connecting these atoms were estimated as follows: $L_0 \simeq (5.0 \pm 0.4)$ Å; $\tau_0 \simeq (26 \pm 2)°$.

DNA/drug interaction; DNA sequence-specificity; DNA ligand recognition; Lexitropsin, Isohelical ligand

1. INTRODUCTION

Successful investigations of the molecular basis of AT-binding specificity in netropsin, distamycin A and Hoechst 33258 binding to DNA [1–7] have continued with the design and synthesis of new types of sequence-specific DNA narrow groove binding ligands [3–8]. Original suggestions of hypothetical chemical structures and corresponding theoretical considerations (lexitropsins, isolexins, and so on) [4,5,9,10] have been made to stimulate the synthesis of ligands with more complicated AT/GC specificity [11,12]. Unfortunately, such promising names have not resulted in the desired DNA-binding properties of the real substances subsequently synthesized [11–13]. To simplify the choice of chemical structures necessary for synthesis, some geometrical rules are presented here for prediction and analysis of the potential capacity of ligand molecules to form sequence-specific DNA narrow groove binding isohelices.

2. MATERIALS AND METHODS

An elongated molecule of a regular sequence-specific ligand was supposed to consist of several repeating bulky elements connected by flexible chains. Imaginary connections of the centers of these elements are shown in Fig. 1 with a cylindrical broken line, which is inscribed in a DNA-like helix. In this model, each middle plane of the bulke element passes through the corresponding two-fold symmetry axis and tends to be most parallel to the walls of the narrow groove, thereby ensuring the largest number of possible van der Waals contacts between ligand and DNA. A general method for calculating of helical parameters of polymer chains from bond lengths, bond angles and internal-rotation angles was developed earlier [14]. Arnott's spatial structure of the B-DNA double helix [15] and standard values of van der Waals radii of atoms were taken for stereometric calculations.

Correspondence address: A.S. Zasedatelev, Engelhardt Institute of Molecular Biology, Academy of Sciences of the USSR, 117984 Moscow B-334, USSR. Fax: (7) (95) 135 1405

3. RESULTS

The geometry of a regular isohelical ligand (see Fig. 1) may be characterized by three dimensions (3.5 Å $\leq s \leq$ 8.5 Å [16], d, $l < L$) of its bulky elements and by parameters L, τ, ψ and α, the dependencies of which on usual helix-generating parameters R, h and t [15] are easily obtained as follows:

$$L = \sqrt{h^2 + 4R^2\sin^2(t/2)} \quad (1)$$

$$\tau = 180° - 2\arcsin\sqrt{\frac{h^2 + R^2\sin^2 t}{h^2 + 4R^2\sin^2(t/2)}} \quad (2)$$

$$\psi = \arcsin\frac{h\sqrt{h^2 + 4R^2\sin^2(t/2)} \times \sin t}{h^2 + R^2\sin^2 t} \quad (3)$$

$$\alpha = \operatorname{arctg}\frac{h}{R \times \sin t} \quad (4)$$

Eqns. 1, 2 and 3 give two additional useful correlations:

$$\tau = 180° - 2\arcsin\frac{\sqrt{h^2\sin^2(t/2) + L^2\cos^2(t/2)}}{L} \quad (5)$$

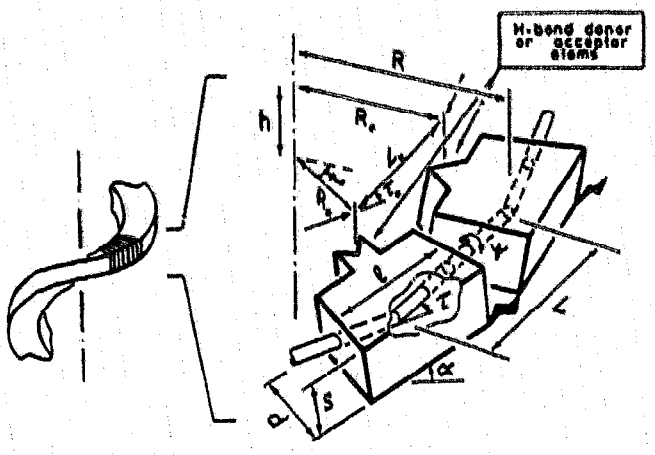

Fig. 1. Main parameters and model representation for two subsequent bulky elements of elongated regular isohelical ligand (right) which fit the free volume of DNA narrow groove (left).

$$\psi = \arcsin \frac{hL \times \sin t}{h^2\sin^2(t/2) + L^2\cos^2(t/2)} \qquad (6)$$

The plots of L and τ against R as well as that of τ against L are shown for the case of B-DNA in Fig. 2. Steric hindrance between H-bond donating atoms of the ligand and B-DNA narrow groove atoms as well as usual values of H-bond length (2.6-3.0 Å [17]) allow one to estimate the maximal interval of permitted R_o values: 4.8 Å $\leq R_o \leq$ 6.8 Å. These limitations determine the ranges of L_o and τ_o values shown in Fig. 2 with shaded areas. The approximate value of $R \simeq$ 8.5 Å characterizing the 'half-depth' of the B-DNA narrow

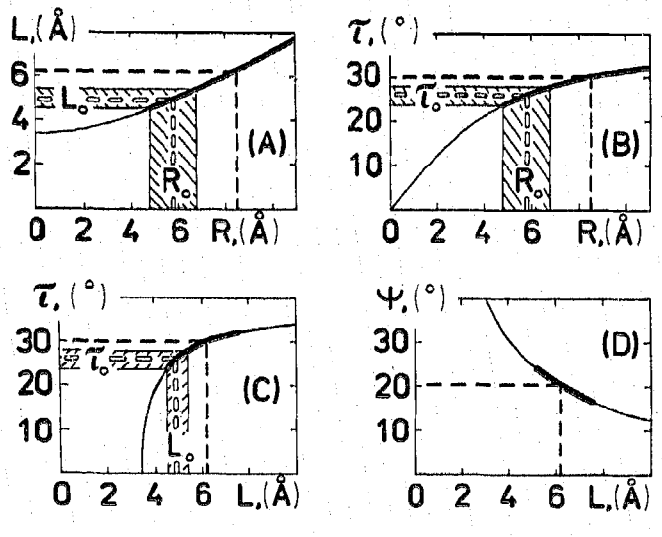

Fig. 2. Plots for eqns. (1) - (A), (2) - (B), (5) - (C) and (6) - (D) in the case of B-DNA (h = 3.38 Å, τ = 36°). Thickened parts of the curves show the sterically permitted regions for parameters. Inside the shaded areas, corresponding to possible variations of R_o, L_o, and τ_o, the open dotted lines determine the average values of these parameters (see text).

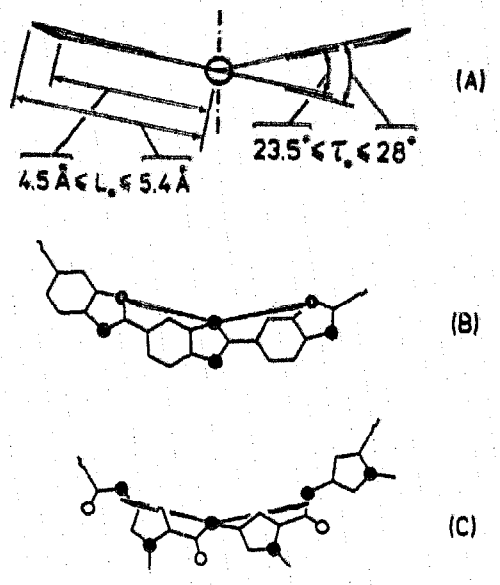

Fig. 3. Polar L_o, τ_o plot providing the stencil (A) helpful for checking of the three subsequent positions of H-bond donating atoms in the ligand molecule. (B) and (C) Corresponding comparisons of Hoechst 33258 and distamycin A backbones with the stencil. Only σ-bonds and not all hydrogens are drawn. (•) nitrogen; (○) oxygen atoms.

groove (see legend to Fig. 2) determines the corresponding values of the next dependent variables: L = 6.2 Å, τ = 30°, $\psi \simeq$ 20°. These values roughly seem to be the most preferable for the corresponding structures of designed isohelical ligands.

Polar L_o, τ_o plots (Fig. 3A) can be used as the stencil which determines the three subsequent positions of H-bond donating atoms in the ligand molecule. When the middle atom is imposed on the circle, the centers of both neighboring ones should fit the short segments shown with thickened lines. Corresponding comparisons obviously should be done with projections of isohelical structures on the stencil's plane. Practically, this procedure is simplified by using a planar representation of molecular structures, taking into account that the resulting deviations of atom positions under consideration appear to be of 0.1 ÷ 0.2 Å, if $\psi \simeq$ 20°, as seen from Fig. 2D.

Comparison of the stencil with such widely known AT-specific DNA narrow groove binding ligands as Hoechst 33258 [2,5-7] and distamycin A [1,3,4,7] is presented in Fig. 3. Taking into account possible rotations around single bonds, one may conclude that the backbone of the former ligand ensures a better fit. At the same time, both structures are characterized by those L_o values which are nearly out of the uppermost permitted limit. This fact can explain the reason of the observed winding of the DNA helix by about 2° caused by distamycin A binding [18].

Inspection of two isolexin structures suggested by Goodsel et al. [9] (Fig. 4A,B) clearly proposes some characteristic properties for each of the ligands. Thus

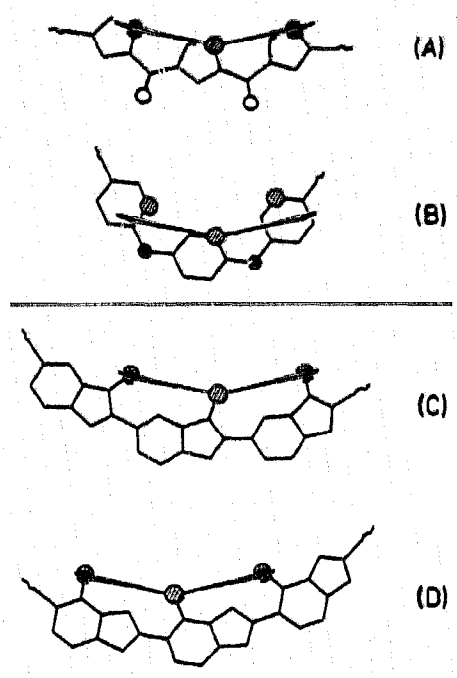

Fig. 4. Application of the stencil for checking the possibility to use ligands as sequence-specific DNA narrow groove binding agents; (A) and (B) isolexins suggested by D. Goodsel and R.E. Dickerson [9]; (C) and (D) new structural motifs, which are suggested here as showing the best fit with the stencil. The shaded circles are H-bond donating atoms (nitrogen or oxygen). Other heteroatomic positions are not detailed.

pyrrol-ketone-like backbones (A) are rather good, but nearly out of the lowest limit of permitted values for L_o. This implies that binding of these sequence-specific ligands should probably be accompanied by some unwinding of the double helix. The pyridyl-amine-like backbones (B) clearly demonstrate a low chance for future DNA drug design.

New structural motifs suggested in Fig. 4C,D, have much better coincidence with the stencil. Therefore these motifs are more adaptable to local deviations of real DNA conformation from the ideal B one.

The stencil suggested in this paper may appear to be a useful instrument (probably by means of computer program) for organic chemists elaborating the strategy for synthesis of sequence-specific DNA binding drugs. The same approach can also be developed for the design of ligands which should demonstrate binding specificity to A- and Z-DNAs or to double stranded RNAs.

Acknowledgements: I thank Drs. G.V. Gursky and A.L. Zhuze (Moscow) for stimulating discussions and appreciate helpful corrections by Prof. R.H. Shafer (San-Francisco).

REFERENCES

[1] Zasedatelev, A.S., Zhuze, A.L., Zimmer, Ch., Grokhovsky, S.L., Tumanyan, V.G., Gursky, G.V. and Gottikh, B.P. (1976) Dokl. Akad. Nauk SSSR 231, 1006–1009.
[2] Zasedatelev, A.S., Mikhailov, M.V., Krylov, A.S. and Gursky, G.V. (1980) Dokl. Akad. Nauk SSSR 255, 756–760.
[3] Gursky, G.V., Zasedatelev, A.S., Zhuze, A.L., Khorlin, A.A., Grokhovsky, S.L., Streltsov, S.A., Surovaya, A.N., Nikitin, S.M., Krylov, A.S., Retchinsky, V.O., Mikhailov, M.V., Beabealashvili, R.S. and Gottikh, B.P. (1983) Cold Spring Harb. Symp. Quant. Biol. 47, 367–378.
[4] Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E. (1985) Proc. Natl. Acad. Sci. USA 82, 1376–1380.
[5] Pjura, P.E., Grzeskowiak, K. and Dickerson, R.E. (1987) J. Mol. Biol. 197, 257–271.
[6] Teng, M., Usman, N., Frederick, Ch.A. and Wang, A.H.-J. (1988) Nucleic Acids Res. 16, 2671–2690.
[7] Gursky, G.V. and Zasedatelev, A.S. (1984) Sov. Sci. Rev. Physicochem. Biol. 5, 53–139.
[8] Youngquist, R.S. and Dervan, P.B. (1987) J. Am. Chem. Soc. 109, 7564–7566.
[9] Goodsell, D. and Dickerson, R.E. (1986) J. Med. Chem. 29, 727–733.
[10] Zakrzewska, K., Randrianarivelo, M. and Pullman, B. (1988) J. Biomol. Struct. Dyn. 6, 331–344.
[11] Lown, J.W., Krowicki, K., Bhat, U.G., Skorobogaty, A., Ward, B. and Dabrowiak, J.C. (1986) Biochemistry 25, 7408–7416.
[12] Borodulin, V.B., Zasedatelev, A.S., Gursky, G.V., Grokhovsky, S.L., Gottikh, B.P. and Zhuze, A.L. (1987) Dokl. Acad. Nauk SSSR 294, 1243–1247.
[13] Burcrhardt, G., Luck, G., Zimmer, Ch., Storl, J., Krowicki, K. and Lown, J.W. (1989) Biochim. Biophys. Acta 1009, 11–18.
[14] Sugeta, H. and Miyazawa, T. (1967) Biopolymers 5, 673–680.
[15] Arnott, S. and Hukins, D.W.L. (1973) J. Mol. Biol. 81, 93–105.
[16] Yoon, C., Prive, G.G., Goodsell, D.S. and Dickerson, R.E. (1988) Proc. Natl. Acad. Sci. USA 85, 6332–6336.
[17] Snounou, G. and Malcolm, D.B. (1983) J. Mol. Biol. 167, 211–216.
[18] Bercovitch-Yellin, Z., Ariel, S. and Leiserowitz, L. (1983) J. Am. Chem. Soc. 105, 765–767.