Volume 225, number 1,2, 151-158 FEB 05343 December 1987

On the possible mechanism of recognition of DNA base sequence by steroid hormones

B. Chandrasekhar, Mrigank and V. Kothekar

Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110029, India

Received 2 October 1987

Geometry of the complex of a steroid hormone, dexamethasone, with a hexanucleotide sequence from the glucocorticoid responsive element d(TGTTCT)$_2$, is optimised here using computer aided geometry simulation with an energy minimization technique. We have also optimised its geometries with genetically modified and arbitrarily chosen DNA sequences. The drug molecule is considered to have both intercalative as well as non-intercalative binding. Comparison of energetics and stereochemical aspects, as well as the H-bonding scheme, is used here to bring out salient features about the mechanism of DNA sequence recognition by steroid hormones.

Steroid hormone; Dexamethasone; DNA recognition; Intercalation; Major groove binding

1. INTRODUCTION

Steroid hormones are believed to enter the cytoplasm and bind to specific receptors. The complex migrates to the nucleus and affects transcription. This classical 'two-step' model based on the early work of Toft and Gorski [1] and extended by Baxter et al. [2] to glucocorticoids has been severely challenged. In vitro physicochemical studies showed that steroids can directly interact with DNA [3–5] whereas recent evidence has shown that steroid receptors alone can bind to DNA hormone responsive elements (HRE) [6,7]. The specificity of the receptor recognition is not unique, in that other steroid receptors can also bind to the same sequence of HRE [8–10]. In contrast, the presence of hormone is mandatory for steroid receptor-HRE interaction to occur in vivo [11]. To date there is no conclusive evidence of whether steroid hormone receptor alone can transcribe a hormone-dependent gene [12–14]. It has only been demonstrated that the consensus core sequence d(TGTTCT)$_2$ found in 5′- and 3′-flanking ends of glucocorticoid responsive elements of several genes can play a significant role in gene regulation. It has proved to be an enhancer.

The above data show the current status of the knowledge on the mechanism of steroid action and indicate the important role played by steroid-DNA interaction. However, the nature of steroid-DNA interaction remains obscure.

Several stereochemical models have been proposed in the literature for steroid hormone-DNA interaction. On the basis of molecular modelling (CPK space filling and Kendrew skeletal models) Hendry et al. [15] suggested that steroids may be capable of inserting between DNA base pairs and forming a highly specific transient intercalation complex with right-handed double helical DNA. Contrary to this, Duax and co-workers [16], on the basis of X-ray crystallographic studies on steroid hormones, proposed that steroids may recognise the DNA base sequence through the D-ring alone via specific H-bonds. Receptors can interact at the A-ring (see fig.1 for nomenclature). The OH$_{11}$ on the β-surface has been shown to be stereospecific

Correspondence address: V. Kothekar, Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110029, India

Published by Elsevier Science Publishers B.V. (Biomedical Division)

for the glucocorticoid receptor [17]. Edelman [18] studied the effects of ethidium bromide, proflavin sulfate, actinomycin D and netropsin on steroid receptor-DNA interaction and concluded that binding occurred in the DNA major groove.

In the present paper we have tried to analyse possible modes by which a bulky steroid molecule with 'domed' structure can interact with DNA as well as bind simultaneously with the receptor which can interact with DNA and enhance specificity.

2. EXPERIMENTAL

2.1. *Intercalation model*

The possibility of intercalation was tried with two dinucleotide sequences $d(TG)_2$ and $d(CG)_2$. We generated the first base pair (TA or CG) following data of Arnott et al. [19]. Cartesian coordinates of the steroid were obtained on the basis of X-ray crystallographic data [20]. The

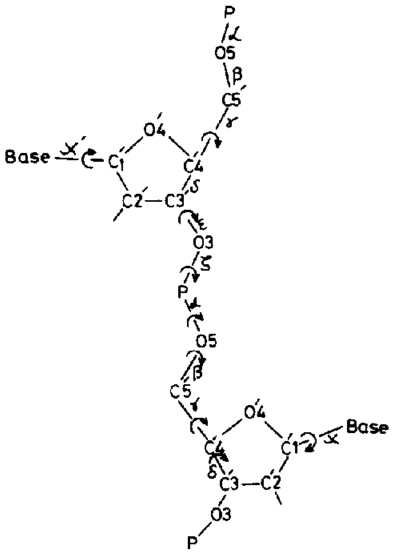

Fig.2. Nomenclature for rotational angles of the dinucleotide backbone.

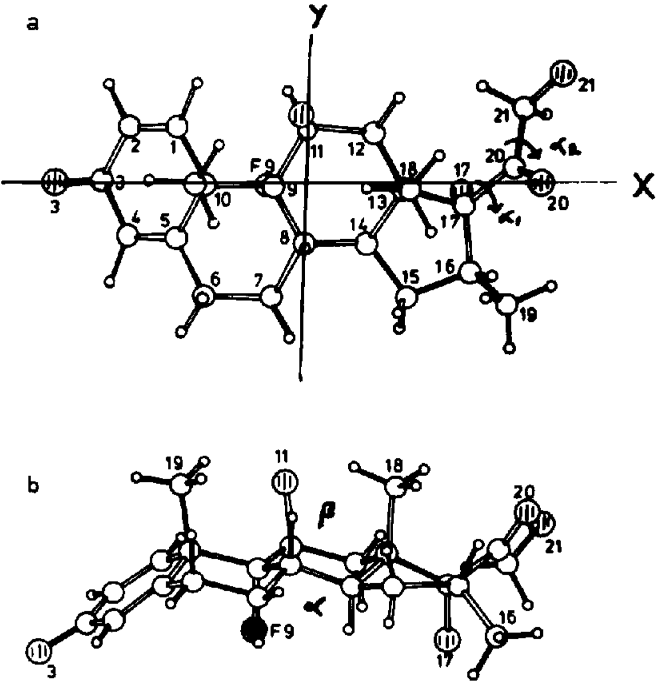

Fig.1. (a) Nomenclature and rotational angles for dexamethasone. The drug is shown in the plane. The X- and XY-axes are depicted. (b) View of the steroid dexamethasone in the XZ plane depicting α- and β-surfaces as well as protruding methyl groups at the 18th and 19th positions.

hydrogen atoms were inserted by a standard procedure [21]. A coordinate frame is defined with respect to the steroid (see fig.1 for details). The *X*-axis of the steroid is then aligned with the C_1'-C_1' vector and the *Y*-axis with the DNA dyad axis. The steroid molecule was allowed to rotate around the DNA helix axis from 0 to 360° in 10° intervals. It was simultaneously allowed to translate along it from 3 Å to 4 Å in 0.1 Å intervals. Small movements were also allowed along the *X*- and *Y*-axis. The drug was assumed to be rigid except for small freedom of rotation of the side chain around the bonds shown in fig.1. The position of the drug was optimised on the basis of total energy (interaction plus conformation) calculated using our previous atom-atom potentials [22,23].

The second base pair was generated in the same way as the first. Its position with respect to steroid-first base pair complex was optimised following the same procedure. The base separation was found to be 7.2 Å. We then generated the DNA backbone attached with the 1st and 2nd base pairs. Three additional atoms C_3'-C_4'-C_5' were attached to the backbone of the first base pair. Rotations were allowed around C_4'-C_5' (γ), C_5'-O_5' (β), P-O_5' (α), P-O_3' (ζ), O_3'-C_3' (ϵ) bonds (fig.2). A function consisting of strain (sum of squares of deviation of cartesian coordinates of the atoms C_3', C_4', C_5' with respect to same atoms attached to 2nd base pair) and conformation energies were calculated and minimised by the damped least square minimization procedure [24]. Because of large base separation, it was impossible to join backbones without increasing bond lengths and putting unusual strain on the DNA backbone. Since many steroids such as 19-nor-steroid compounds do not possess pro-

Table 1

Comparison of geometric parameters for the intercalating model of $d(TG)_2$ with B-DNA and average intercalation model

	B-DNA[a]	Average intercalation[b]	Model intercalation
1. Torsional angles (°)			
α	−39	−70(8)	−84
β	−151	−135(9)	150
γ	30	59(14)	160
δ	156	–	–
ϵ	159	−150(11)	−163
ζ	−98	−67(7)	−116
−5′ end	−95	−163(12)	−158
−3′ end	−95	−76	−95
sugar pucker	C_2' endo	C_2' endo 5′ end C_3'/C_2' endo 3′ end	C_2' endo
2. Interatomic distances (Å)			
Same strand			
C_1'-C_1'	4.9		7.0
P..P	6.4		6.2
Opposite strand			
C_1'-C_1'	10.7		11.6
P..P	18.4		16.2
3. Turn angle (°) between C_1'-C_1' 1st pair and C_1'-C_1' 2nd pair	36		16

[a] [19]

[b] [26]

truding methyl groups on the β-surface we decided to attempt the intercalation model with these groups removed. The base pair separation in the absence of methyl groups reduced to 6.6 Å. The backbone could join smoothly. This model is depicted in fig.3. The comparison of its geometric parameters and energetics for this model is shown in tables 1 and 2.

2.2. *Major groove binding model*

Because of steric limitations steroids cannot interact in the minor groove. The major groove binding was attempted following the same procedure as that for minor groove binding drugs (netropsin, distamycin, dye Hoechst 33258) [23,25]. The interaction of dexamethasone was studied with the consensus core sequence d(TGTTCT)$_2$ as well as genetically modified and arbitrarily chosen hexanucleotide duplexes (see table 3 for details). The model of interaction with the core sequence is depicted in fig.4. Comparison of the H-bonding pattern and energetics is shown in tables 3 and 4.

3. RESULTS AND DISCUSSION

3.1. *Intercalation model*

3.1.1. Stereochemical aspects

Several changes in backbone and glycosyl torsional angles are necessary for extending base pair separation from 3.4 Å to 6.6 Å (table 1). Rotation around phosphodiester linkages (α and ζ) was in -sc conformation. Rotation around C_3'-O_3' (ϵ) is in the *trans*-conformation. It is much different from the B-DNA model but it is close to the average intercalation model. Major differences between our model and the average intercalation model were noticed in the case of angles β and γ. The reason for this was the difference in sugar pucker between our model and classical intercalator models [26]. In the average intercalation model, the sugar pucker combination for adjacent base pairs had been C_2' endo-C_3' endo. In contrast we obtained a C_2' endo sugar pucker for both base pairs which agrees with the daunomycin-DNA complex [27,28]. In contrast to conventional DNA models, rotation around C_4'-C_5' (γ) was observed in *trans*-regions similar to Z-DNA [29] and justified by molecular orbital calculations [30]. The glycosyl angles were in *trans*- or high *anti*-region and were in the vicinity of average intercalation and B-DNA models.

We observed an increase in the C_1'-C_1' distance in the same strand by 2.1 Å. The C_1'-C_1' distance in opposite strands increased by 0.9 Å, and the P....P distance across the helix was reduced by 2.2 Å. There was unwinding of DNA base pairs by about 20°. Glycosyl bonds were in *cis*-conformation.

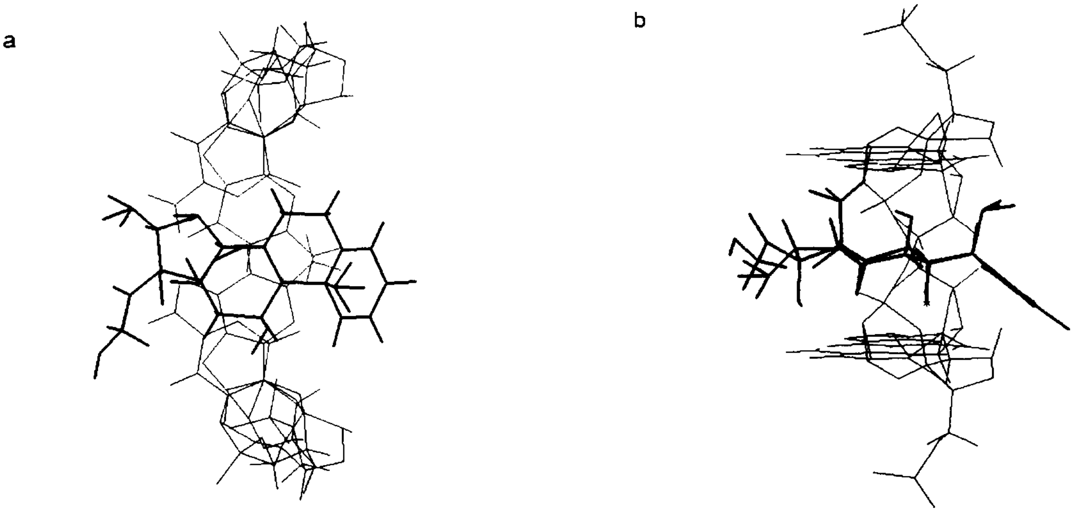

Fig.3. Intercalation model showing position of the drug with respect to DNA base pairs in (a) the *XY* plane and (b) the *YZ* plane. Intercalation complex of d(CG) with dexamethasone.

Table 2

Energy partitioning of intercalation models

No.	Base sequence	E_{non}	$E_{ele+pol}$	$E_{hydg.}$	$E_{int.}$
1.	d(TG)$_2$	−37.80	−16.06	−1.85	−55.72
2.	d(CG)$_2$	−23.05	−12.20	−3.79	−39.00

E_{non}, nonbonding interaction energy calculated using Lennard Jones 6–12 potential; $E_{ele+pol}$, electrostatic (monopole-monopole, dipole-induced dipole) interaction calculated at dielectric permeability value 4.0 using CNDO/II charge distribution; $E_{hydg.}$, H-bonding energy calculated by Momany et al. 10–12 potential [32]; $E_{int.}$, total interaction energy of DNA with sequence. All energies are in kcal/mol

The steroid molecule is almost perpendicular to the DNA base pair (fig.3a) with its A-ring penetrating the minor groove. The D-ring with attached side chain lies in the major groove. Its OH_{17} proton lies within the H-bonding distance of O_6 of guanine. Its distance in d(TG)$_2$ and d(CG)$_2$ was 2.9 and 2.6 Å, respectively. This model agrees generally with X-ray crystallographic data on the daunomycin d(CGATCG) complex [27].

3.1.2. Energetic aspects

Comparison of total interaction energy shows (table 2) that d(TG)$_2$ is preferred over d(CG)$_2$. The interaction is stabilized by stacking which contributes 71.4 and 56.9%, respectively. Steroid has slightly more interaction with the 5′-strand in d(TG)$_2$ and with the 3′-strand in d(CG)$_2$. Preference for d(TG)$_2$ is mainly because of changes in the stacking interaction (table 2). Electrostatic and H-bonding contributions show a small preference for d(CG)$_2$.

3.2. *Major groove model*

3.2.1. Stereochemical aspects

The model for the interaction of dexamethasone in the major groove of the core sequence is depicted in fig.4a. The drug molecule is seen to occupy a region of over three base pairs. In the first

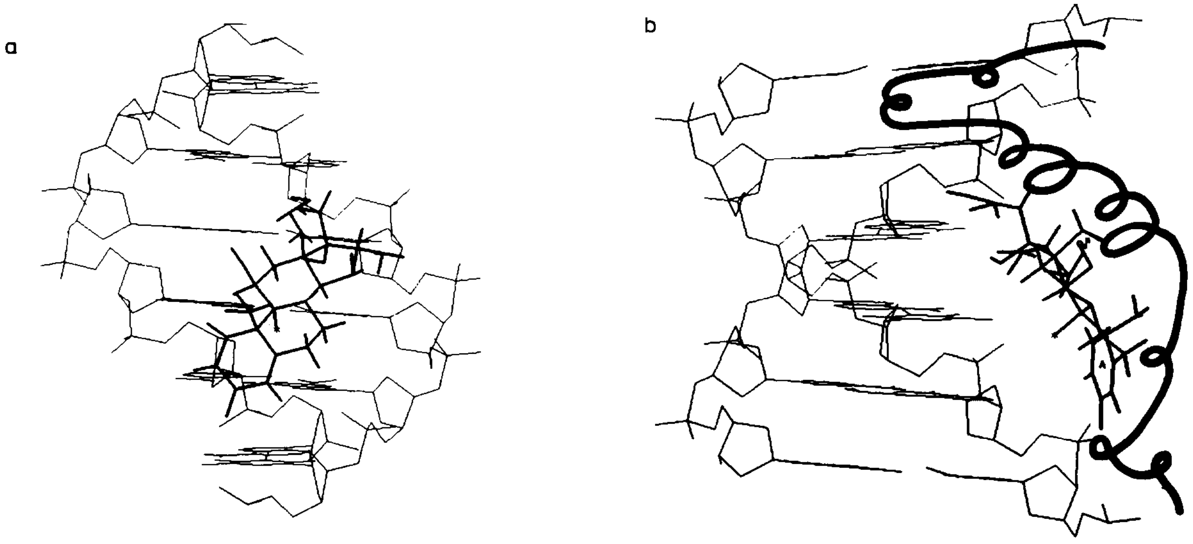

Fig.4. (a) Interaction of dexamethasone with the major groove of the core sequence d(TGTTCT)$_2$. Figure clearly shows interaction with 3rd, 4th and 5th base pairs. (b) The proposed model for the interaction of the receptor (shown by very thick line) with steroid in the major groove of DNA. The figure shows points of contact of the receptor with the A-ring, OH_{11} and phosphate and bases of DNA.

Table 3

Hydrogen bonding scheme (major groove binding model)

No.	Sequence	Atom pair distance (Å)	Atom pair distance (Å)	Atom pair distance (Å)	Atom pair distance (Å)
1.	$d(TGTTCT)_2$	OH_{17} $TO_4(3)$ 2.79	OH_{17} $TO_4(4)$ 2.89	OH_{21} $TO_4(4)$ 2.28	OH_{21} $GO_6(-5)$ 1.94
2.	$d(TG\underline{G}TCT)_2$		OH_{17} $TO_4(4)$ 2.90	OH_{21} $TO_4(4)$ 2.24	OH_{21} $GO_6(-5)$ 1.97
3.	$d(TG\underline{A}TCT)_2$		OH_{17} $TO_4(4)$ 2.90	OH_{21} $TO_4(4)$ 2.24	OH_{21} $GO_6(-5)$ 1.97
4.	$d(TG\underline{C}TCT)_2$		OH_{17} $TO_4(4)$ 2.90	OH_{21} $TO_4(4)$ 2.24	OH_{21} $GO_6(-5)$ 1.97
5.	$d(T\underline{A}TTCT)_2$	OH_{17} $TO_4(3)$ 3.04	OH_{17} $TO_4(4)$ 2.60	OH_{21} $TO_4(-4)$ 2.28	OH_{21} $GO_6(-5)$ 1.94
6.	$d(TGT\underline{A}CT)_2$		OH_{17} $TO_4(-4)$ 2.95	OH_{21} $TO_4(-4)$ 2.64	OH_{21} $GO_6(-5)$ 2.04
7.	$d(TGT\underline{G}CT)_2$	OH_{17} $TO_4(3)$ 2.79	OH_{17} $GO_6(4)$ 2.93	OH_{21} $GO_6(4)$ 1.92	OH_{21} $GO_6(-5)$ 2.91
8.	$dG_6 \cdot dC_6$		OH_{17} $GO_6(4)$ 2.92	OH_{21} $GO_6(4)$ 1.98	OH_{21} $GO_6(5)$ 2.44
9.	$d(TATATA)_2$		OH_{17} $TO_4(-4)$ 2.56	OH_{21} $TO_4(-4)$ 2.19	OH_{21} $AN_7(-5)$ 2.77

The first sequence is the core sequence. Other sequences have a mutation (shown by underlining) at 2nd, 3rd or 4th position (positions involved in H-bonding). Numbers in parentheses represent base pair position. The distance between proton (H) and electron donor (x) is given here

five sequences considered by us (table 3) thymine at the fourth position and guanine at position (−5) are invariant. As a result OH_{17} and OH_{21} form H-bonds with thymine O_4 or guanine O_6 at 4th and 5th positions. Mutations at the fourth position alter the H-bonding pattern. Replacement of guanine for thymine does not affect H-bonds as O_6 can be involved in H-bonding in place of O_4. In the case of replacement of thymine by adenine the drug readjusts its position and forms an H-bond with O_4 of T-4.

3.2.2. Energetic aspects

The net difference in the interaction energy of the core sequence and genetically modified sequences was small. All three types of interactions (nonbonding, electrostatic and H-bonding) contributed towards it (table 4). The core sequence (sequence 1) and one mutated sequence where guanine at the second position was replaced by adenine (sequence 5) had almost the same energy and were the most preferred sequences. The maximum difference in the electrostatic interaction was observed between $dG_6 \cdot dC_6$ and $d(TATATA)_2$. This was due to difference in the electrostatic potential in the major groove of these two sequences [31]. Maximum contribution due to H-bonding interaction was observed in the case of

Table 4

Interaction energy break ups (kcal/mol) for major groove binding model

No.	Hexanucleotides	E_{non}	$E_{ele+pol}$	$E_{hydg.}$	$E_{int.}$
1.	$d(TGTTCT)_2$	−23.60	−17.93	−9.67	−52.04
2.	$d(TG\underline{G}TCT)_2$	−20.47	−16.54	−8.60	−45.61
3.	$d(TG\underline{A}TCT)_2$	−21.01	−17.77	−8.16	−47.01
4.	$d(TG\underline{C}TCT)_2$	−23.24	−17.37	−7.15	−47.77
5.	$d(T\underline{A}TTCT)_2$	−24.95	−18.28	−9.54	−52.77
6.	$d(TGT\underline{A}CT)_2$	−21.33	−18.25	−8.24	−47.83
7.	$d(TGT\underline{G}CT)_2$	−25.31	−16.07	−8.93	−50.33
8.	$dG_6 \cdot dC_6$	−22.59	−15.20	−7.37	−45.03
9.	$d(TATATA)_2$	−17.94	−21.22	−8.00	−47.18

Energy parameters E_{non}, $E_{ele+pol}$, $E_{hydg.}$, $E_{int.}$ have the same meaning as in table 2 core sequence. This as we believe is the primary reason for the preference of core sequence in 5′- and 3′-flanking ends of GRE of many organisms [12–14].

3.3. *Receptor binding*

In the case of the intercalation model OH_{11} is deeply buried within DNA base pairs and the receptor molecule cannot interact with it. The drug can interact with the receptor through the A-ring in the minor groove. However, the minor groove of the B-form of DNA is too small to accommodate the receptor.

In the case of the major groove binding model OH_{11} on the β-surface as well as the A-ring are free to interact with the receptor in the major groove. Not only can the major groove accommodate the steroid receptor complex but there can also be specific contacts between receptor and DNA (fig.4b) which would enhance specificity.

4. CONCLUSION

The results presented here clearly demonstrate preference of the major groove binding model over intercalation for dexamethasone. In the case of other steroids such as 19-nor-steroid compounds where methyl groups are absent the intercalation model may be feasible stereochemically. Although the specificity in the case of the intercalation model is higher, it is difficult to explain the interaction of the receptor-steroid complex with DNA on the basis of this model.

ACKNOWLEDGEMENT

The authors are thankful to the Department of Science and Technology, Government of India, for financial assistance.

REFERENCES

[1] Toft, D.O. and Gorski, J. (1966) Proc. Natl. Acad. Sci. USA 55, 1574–1581.

[2] Baxter, J.D., Rousseau, G.G., Benson, M.C., Garecea, R.L., Ito, J. and Tomkins, G.M. (1972) Proc. Natl. Acad. Sci. USA 69, 1892–1896.

[3] Kidson, C., Thomas, A. and Cohen, P. (1970) Biochemistry 9, 1571–1576.

[4] Arya, S.K. and Yang, T.J. (1975) Biochemistry 14, 963–969.

[5] Blackburn, G.M., Kellard, B., Rashid, A. and Thompson, M.H. (1983) Biochem. Pharmacol. 32, 2571–2575.

[6] Willmann, T. and Beato, M. (1986) Nature 324, 688–691.

[7] Bailly, A., Lepage, C., Rauch, M. and Milgrom, E. (1986) EMBO J., 3235–3241.

[8] McKnight, G.S., Penneguin, P. and Schimke, R.T. (1975) J. Biol. Chem. 250, 8105–8110.

[9] Cattelli, M.G., Binart, N., Elkik, F. and Baulieu, E.E. (1980) Eur. J. Biochem. 197, 165–172.

[10] Von der Ahe, D., Renoir, J.M., Buchou, T., Baulieu, E.E. and Beato, M. (1986) Proc. Natl. Acad. Sci. USA 83, 2817–2821.

[11] Becker, P.B., Gloss, B., Schmid, W., Strable, U. and Schutz, G. (1986) Nature 324, 686–688.

[12] Govindan, M.V., Spiess, E. and Majors, J. (1982) Proc. Natl. Acad. Sci. USA 79, 5157–5161.

[13] Payvar, F., Defranco, D., Firestone, G.L., Edgar, B.E., Wrange, O., Orket, S., Gastafsson, J.A. and Yamamoto, K.R. (1983) Cell 35, 381–392.

[14] Scheidereit, C. and Beato, M. (1984) Proc. Natl. Acad. Sci. USA 81, 3665–3669.

[15] Hendry, L.B., Bransome, E.D.J., Lehner, A.F., Muldoon, T.G., Hutson, M.S. and Mahesh, V.B. (1986) J. Steroid Biochem. 24, 843–852.

[16] Duax, W.L., Grittin, J.F. and Ebright, R. (1985) in: Molecular Basis of Cancer, Part B (Rein, R. ed.) pp.263–273, Alan Liss, New York.

[17] Rousseau, G., Baxter, J.D. and Tomkins, G. (1972) J. Mol. Biol. 67, 99–115.

[18] Edelman, I.S. (1975) J. Steroid Biochem. 6, 147–160.

[19] Arnott, S., Smith, P.J.C. and Chandrasekharan, R. (1976) in: CRC Handbook of Biochemistry and Molecular Biology (Fasnan, G.D. ed.) vol.2, pp.411–422, CRC, Cleveland.

[20] Griffin, J.F. and Duax, W.L. (1985) in: Atlas of Steroid Structure, vol.2, PR 96A1, p.553.

[21] Wiberg, K.B. (1965) J. Am. Chem. Soc. 87, 1070–1075.

[22] Kothekar, V., Mrigank, Royyuru, A.K. and Kalia, A. (1986) Int. J. Quant. Chem. Quant. Biol. Symp. 13, 175–183.

[23] Gopalkrishna, K., Kalia, A., Royyuru, A.K., Mrigank, Adreya, M., Rao, M.V.R. and Kothekar, V. (1987) FEBS Lett. 215, 95–99.

[24] Vitek, A. (1968) Collect. Czeck. Chem. Commun. 33, 1601–1607.

[25] Mrigank, Royyuru, A.K. and Kothekar, V. (1986) FEBS Lett. 195, 203–208.

[26] Shieh, H.S., Berman, H.M., Dabrow, M. and Neidle, S. (1980) Nucleic Acids Res. 8, 85–97.

[27] Wang, A.H.-J., Ughetto, G., Quigley, G.J. and Rich, A. (1987) Biochemistry 26, 1152–1163.

[28] Nuss, M.E., Tame, T.L., Apple, M.A. and Kollman, P.A. (1980) Biochim. Biophys. Acta 609, 136–147.

[29] Wang, A.H., Quigley, G.J., Kolpak, F.J., Van der Marel, G., Boom, J.H.V. and Rich, A. (1981) Science 211, 171–176.

[30] Pack, G.R., Muskavitch, M.A. and Koew, G. (1977) Biochim. Biophys. Acta 78, 19–22.

[31] Lavery, R. and Pullman, B. (1984) Nucleic Acids Res. 9, 4673–4679.

[32] Momany, F.A., Carruthers, I.M., McGuire, R.F. and Scheraga, H.A. (1974) J. Phys. Chem. 78, 1595–1620.